

4



Forecasting Education

Forecasting tools and models can be exploratory, normative, or both.

There are many ways to approach the discussion of tools and models for forecasting the expansion of global education. It is possible to consider them in terms of their coverage and aggregation—whether they are country-specific or multicountry; whether they focus on primary education or look also at other levels of education; whether they consider only enrollment levels or also the underlying intake and survival patterns; whether they attend to student flows and/or to adult education attainment; or whether they forecast for ten years, twenty-five years, or more.

It is also possible to talk about tools and models in terms of their concern with, and treatment of, related issue areas—whether they consider demographics and economics explicitly and dynamically in interaction with education, whether they consider primarily the impact of these other systems on education, or whether they also look to the implications of education for other aspects of human

development. And it is possible to talk about tools and models in terms of their basic methodological characteristics—whether they are largely extrapolative of select variables or more broadly structural in their representation of multiple, interacting facets of educational systems; whether they tend primarily to be accounting systems with exogenously (externally) provided assumptions about change; or whether they more dynamically represent households, governments, and other potential agents in interaction.

As important as all such characteristics are, perhaps the most fundamental distinguishing characteristic is something else—namely, the purpose and desired outcome of the use of the tool or model. Most broadly, forecasting tools and models are organized around two purposes. *Exploratory* tools seek “simply” to understand the path of a system, whereas *normative* tools identify a desired future and then assess the likelihood of attaining that future and/or identify means by which the path toward the

desired outcome might be accelerated, redirected (if the current path is not congruent with the goal), or otherwise enhanced.¹

This volume combines exploratory and normative purposes. As Chapter 1 indicated, the questions we seek to address are: (1) What path does the formal global education system, as a collection of countries, appear to be on as we look forward fifty years? (2) Is an aggressive but still reasonable acceleration of that path possible? (3) What might be the broader consequences of such a normative but attainable acceleration? Given those purposes and questions—and understanding that all models are simplifications of reality and therefore fall short of being ideal tools—what are some of the general characteristics of the “ideal” tool that we might want for such investigation?

Characteristics of Ideal Education Forecasting Models and Tools

There is a considerable distance between the characteristics of the simplest possible exploratory and normative education forecasting tools and the characteristics that would be found in an ideal model or tool for our purposes in this volume. Here, we list some desirable elements of a tool with a mid- to long-term temporal reach, beginning with the characteristics that an exploratory model would include:

- An *accounting system* that tracks student flows by education level across all levels and grades as well as education attainment in the adult population, with as much detail as possible regarding elements that vary from one component of a population to another (e.g., sex, age, rural-urban residence, income status, and ethnicity).
- Representation of the *dynamics* that are the immediate drivers of student flows (and hence ultimately of attainment levels), including separate representations of demand and supply dynamics and constraints.
 - With respect to demand, the ideal exploratory system represents the dynamics of enrollment patterns (intake and survival) in the context of family circumstances and demographic and economic trends.
 - With respect to supply, the ideal tool has the capability to estimate the costs

and resource requirements associated with various enrollment dynamics and demographic patterns and to forecast the likely need for, and availability of, public, private, and international funds.

- Representation not only of demographic and economic impacts and constraints on education but also of education’s impacts on demography and on economic systems, as well as bidirectional feedback loops between education and other aspects of human development systems, such as poverty reduction and the characteristics of sociopolitical systems.²
- Transparency of structures, equations, algorithms, and data; availability to others for use and analysis; flexibility and simplicity of use.

These same elements would characterize an ideal mid- to long-range normative forecasting model or tool. However, the ideal normative tool would also include the following:

- Specification of points of intervention and an assessment of their reasonableness.³
- Evaluation of the impacts of the interventions not just on education participation and attainment but also on broader systems (demographic, economic, and sociopolitical).
- At least some elements of a cost-benefit analysis.

This chapter describes the IFs modeling system, which can be used for both exploratory and normative analyses, and considers its particular strengths and limitations relative to the ideal. First, however, we briefly introduce significant global education modeling and forecasting approaches that others have developed in recent years.

Recent Education Modeling and Forecasting Approaches

Over the last several years, a number of models have been developed by others with similar interests in understanding the education transition and the transition’s likely continued unfolding, requirements, and/or consequences. We identify the models or tools here that have informed our efforts and comment very briefly on some of their features that had special

■ *Certain characteristics are associated with “ideal” models, even though no model can attain them all.* ■

■ *In recent years, a number of models have been developed to help in understanding, forecasting, and facilitating the education transition.* ■

■ *The IFs project has built upon features of other models and has also independently developed new features and modeling capabilities.* ■

relevance for our own work; more information about these other models and approaches appears in the Appendix to this chapter. Later, in Chapter 5, we will consider forecasts from some of these tools in comparison with those from IFs.

McMahon (1999) focused heavily on exploring the social benefits of primary and secondary education, thus connecting education's expansion to economic, demographic, and sociopolitical change. He used a cross-sectional approach to drive much of the dynamic analysis in his econometric model and developed a base case as well as two normative scenarios to forecast the impact of specific education policy changes.⁴

Delamonica, Mehrotra, and Vandemoortele (2001), using an accounting-centric approach and UN population projections, conducted a normative analysis of the incremental costs of moving to universal primary education by 2015. Their estimates of costs included measures intended to enhance education quality as well as capital costs for needed increases in capacity. Bruns, Mingat, and Rakotomalala (2003) also explored the costs of meeting universal primary education by 2015. They extended previous analyses by developing "best practice" expenditure and resource mobilization guidelines based on the education policies and practices of low-income countries making the best progress toward universal primary education, and their framework included the normative concept of "minimum adequate cost."

Clemens (2004) used historical data to analyze transition paths in net primary enrollment rates and found an S-shaped curve that could be used to extrapolate the number of years countries and regions might need to reach 90 percent primary net enrollment. He also compared the earlier experience of currently high-income countries with the recent experience of low- and middle-income countries, and found that the speed of advance in enrollment rates has accelerated considerably. Wils, O'Connor, and Somerville, as reported in a paper authored by Wils, Carrol, and Barrow (2005), also found S-shaped patterns in the advance of primary education and used them to project growth in primary entry rates and completion rates separately (rather than aggregate enrollment rates). They also used data

gathered in household surveys to provide more extensive and longer estimates of historical patterns of school participation.

Lutz, Goujon, and Wils (2005) built on the multistate demographic methodology of the International Institute for Applied Systems Analysis (IIASA) to explore the future across all levels of education, including tertiary. Moreover, the principal focus of their work was on adult attainment levels differentiated by age and sex, so as to build the foundation for looking at the relationship between those levels and characteristics and other aspects of global change. Similarly, Hilderink (2007) added attention to adult attainment to his forecasting of flows across all levels of formal education. Also of interest to us, his formulations, like ours, relate education demand and education supply to GDP per capita.

An EPDC paper (2007b) reported the work of Wils, Barrow, Oliver, Chaluda, Goodfriend, Kim, and Sylla in the development and early use of ProEnrol, a country-level, cohort-projection model for use at the primary and secondary levels. The paper described the model as the first effort to make cohort or grade-by-grade projections in an international, global series, including representations of promotion and repetition.

The IFs project, as will be seen in the discussion that follows, has both independently developed and also built upon many of the features of these other models: the cross-sectional analysis and attention to sociopolitical impacts of education found in McMahon (1999); the computation of costs required to meet goals of Delamonica, Mehrotra, and Vandemoortele (2001), as well as some of the attention to best practice that Bruns, Mingat, and Rakotomalala (2003) built into their analysis; the S-shaped expectations for education's advance that both Clemens (2004) and Wils, O'Connor, and Somerville cited in Wils, Carrol, and Barrow (2005) found and used, as well as their explicit recognition of country-specific circumstances; the attention to adult attainment of education by Lutz, Goujon, and Wils (2005); the use of GDP per capita to drive the formulation of education demand and supply made by Hilderink (2007); and the country-level, cohort analysis of Wils et al. (2007b). This is not to say, of course, that one model can do everything as fully or as well as more specialized studies and approaches, but

we do believe there is also value in our more comprehensive approach. We will return to a description of the IFs approach to modeling of education after providing an introduction to the larger IFs system.

The IFs Modeling System

As stated previously, the particular strengths of IFs derive from the combination of its extended time frame, its extensive geographic coverage with capability to flexibly group countries for analysis and display, and its dynamic integration of multiple human systems. In addition, its global education model is the only one we know of that represents all three levels of formal education in grade-by-grade student flows or cohorts, as well as the only one that represents lower and upper secondary education separately. IFs can be used both for exploratory analyses of dynamic trends and patterns and for the creation of normative scenarios and explorations of their respective impacts.

In the sections that follow, we will first provide a brief overview of the broader IFs forecasting system and then discuss the IFs education model in more detail. In the process of that discussion, we will attempt to identify the strengths and weaknesses of IFs in comparison with the characteristics of an ideal global education forecasting tool.

General design considerations

International Futures is a large-scale, long-term, integrated global modeling system. It represents demographic, economic, energy, agricultural, sociopolitical, and environmental subsystems for 183 countries interacting in the global system.⁵ The central purpose of IFs is to facilitate exploration of global futures through alternative scenarios.

The issues of interest that motivated the design of IFs fall generally into three categories: human development, social fairness and security, and environmental sustainability (see Table 4.1). Across these domains, the project especially looks to Sen (1999) for his emphasis on freedom and individual development, Rawls (1971) for his emphasis on fairness within society, and Brundtland (UN 1987) for her seminal definition of sustainability. These emphases, in combination, provide a philosophical

Table 4.1 Human systems and issues of interest to the IFs project

Humans as individuals	Personal development/freedom
Humans with each other	Peace and security/social fairness
Humans with the environment	Sustainable material well-being

framework for the exploration of human beings as individuals, of human beings with each other, and of human beings with the environment.

Human systems fundamentally involve different types of agents (economists tend to focus on households and firms; political scientists add governments) interacting with each other in various structures (economists focus on markets; political scientists look to action-reaction systems and international regimes; sociologists add societies and demographic structures; anthropologists focus on cultures; and physical scientists extend the reach to ecosystems). In general, scientists seek to understand the complex cocreation and evolution of agent behavior and the structural characteristics of human and social systems.

IFs attempts to capture *some* of that complexity and richness by being rooted in the theory of various disciplines and subspecializations. It is a structure-based, agent-class-driven, dynamic modeling system. That is, it tries to represent typical behavior patterns of major agent classes (households, governments, firms) interacting in a variety of global structures (demographic, economic, social, and environmental) with extensive representation of underlying accounting systems.⁶ IFs draws upon standard approaches to modeling specific issue areas whenever possible, and then, as necessary, it extends and integrates these. For instance, the IFs demographic model uses a typical “cohort-component” representation, tracking country-specific populations over time by age and sex, further differentiated in IFs by education. Within that structural or accounting framework, the model represents the fertility decisions of households (influenced by income and education) as well as mortality and migration patterns.

The database underlying IFs (and integrated with the system so it can be used by others) includes a vast range of data for 183 countries, represented over as many years since 1960 as possible on a country-by-country basis.⁷

■ *The desire to explore issues of human development, social fairness and security, and environmental sustainability motivated the design of IFs.* ■

■ *IFs draws upon standard modeling approaches to specific issue areas whenever possible and then, as needed, extends and integrates them.* ■

■ *Extensive documentation of the IFs system of models is at www.ifs.du.edu, as is the model itself in both online and downloadable forms.* ■

The model system itself runs in annual time steps from its initial year (currently 2005),⁸ and the model interface facilitates user interventions flexibly across time, issue area, and geography.

The models of IFs

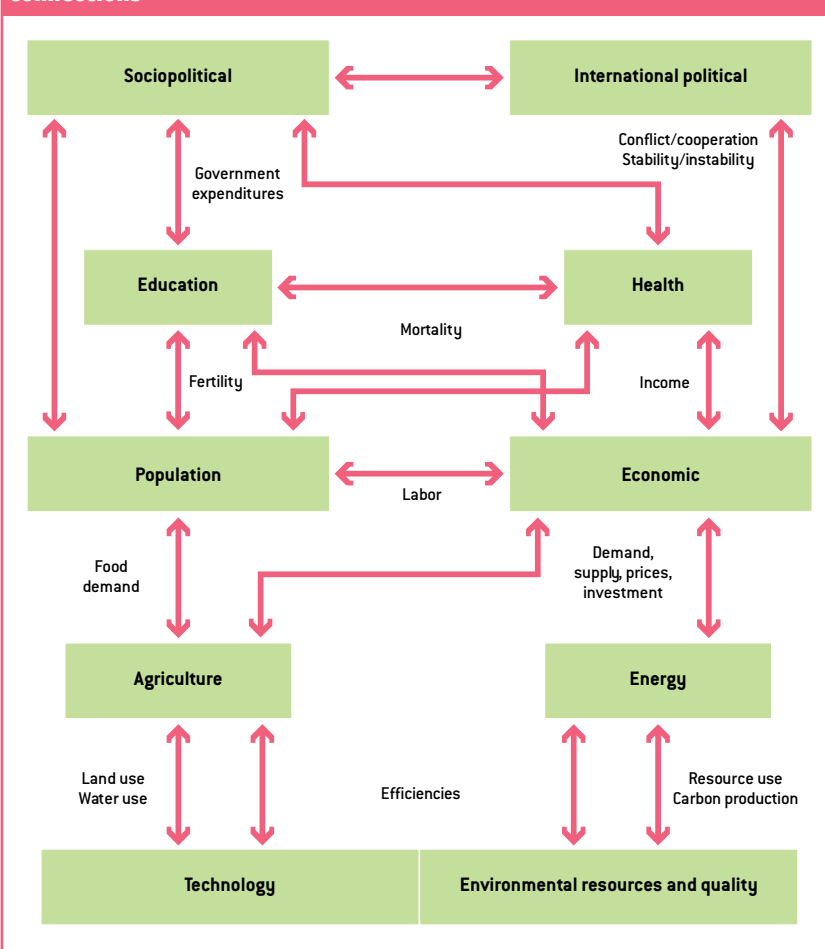
Figure 4.1 shows the major conceptual blocks. Full issue-specific models represent most of the blocks, including education. The elements of the technology block are actually dispersed throughout the system, and the named linkages between blocks (and the identified linkages themselves) are a small illustrative subset, by no means an exhaustive listing.

The two models within the IFs system that interact mostly closely with the education model are the population and economic models. In our representation of the human development system, the sociopolitical model also interacts quite closely with the education model (as well

as with the economic and demographic models). We cannot here provide technical details of these or other models that collectively make up the IFs forecasting system that fully integrates those models. We do, however, provide certain summary information in the discussion ahead, which will be too much technical detail for some readers and far too little for others. Those who want more information about the IFs system (including the education model) will find extensive documentation at www.ifs.du.edu. In particular, see Hughes (2004b, 2006; Hughes et al. 2004) for a structural overview and for discussions of scenario analysis and validation. The model system is also freely available there in both online and downloadable forms.

The demographic model has the standard cohort-component structure that the UN and other institutions use in population forecasting, but it represents fertility and mortality as functions of other variables in IFs. Some of the key characteristics of the population model are that it

Figure 4.1 The major models in the IFs modeling system and example connections



- Represents twenty-two age-sex categories to age 100+ in the cohort-component structure (but computationally spreads the five-year cohorts initially to one-year cohorts and calculates change in one-year time steps)
- Calculates change in cohort-specific fertility in response to income, income distribution, education levels, and contraception use
- Calculates change in mortality rates in response to income, income distribution, education, and assumptions about technological change's impact on mortality
- Separately represents the evolution of HIV infection rates and deaths from AIDS
- Computes literacy rates, average life expectancy at birth, and an overall measure of human development (the Human Development Index, or HDI)
- Represents migration and ties it to flows of remittances.

The economic model has the multisector equilibrium structure of models that most forecasters of development processes use, but it has extended representation of the production side so as to facilitate long-term analysis and to link productivity to other variables in IFs, including education. Some of the most important characteristics of the economic model are that it

- Represents the economy in six sectors: agriculture, materials, energy, industry, services, and information/communications technology, or ICT; other sectors could also be configured because the system uses raw data from the Global Trade and Analysis Project (GTAP)
- Computes and uses input-output matrices that change dynamically with development level
- Is a general equilibrium-seeking model that does not assume exact equilibrium will exist in any given year; rather, it uses inventories as buffer stocks and to provide price signals so that the model chases equilibrium over time
- Contains a Cobb-Douglas production function that (following insights of Solow and Romer) endogenously represents contributions to growth in multifactor productivity from human capital (education and health), social capital and governance, physical and natural capital (infrastructure and energy prices), and knowledge development and diffusion (research and development and economic integration with the outside world)
- Uses a Linear Expenditure System to represent changing consumption patterns
- Utilizes a “pooled” rather than bilateral trade approach for international trade
- Is embedded in a social accounting matrix (SAM) envelope that ties economic production and consumption to representation of intra-actor financial flows (it represents, however, only the skilled and unskilled households from GTAP).

Few sociopolitical models exist except in the form of highly specialized representations (such as the forecasting of state failure). The model in IFs has a relatively extensive treatment of sociopolitical variables, including government budgeting, which is important in representing constraints upon expansion of education. Some of the sociopolitical model’s relevant features are that it

- Represents fiscal balances through taxing and spending decisions
- Shows six categories of government spending: military, health, education, research and development, foreign aid, and a residual category (as well as representing transfer

- payments for pensions and social welfare)
- Represents changes in social conditions of individuals (such as fertility rates, literacy levels, or poverty), attitudes of individuals (such as the level of materialism/postmaterialism of a society from the World Values Survey), and the social organization of people (such as the status of women)
- Represents the evolution of democracy
- Represents (in very basic fashion) the prospects for state instability or failure.

The use of IFs

Although initially developed as an educational tool, IFs is increasingly used in research and policy analysis. For instance, it was a core component of the TERRA project sponsored by the European Commission to explore the New Economy. More recently, forecasts from IFs supported Project 2020 of the National Intelligence Council (NIC) (USNIC 2004) as well as NIC’s subsequent study, *Global Trends 2025: A Transformed World* (USNIC 2008). IFs also provided driver forecasts and some integrating analysis for *Global Environment Outlook-4* of the United Nations Environment Programme (2008).

The menu-drive interface of the International Futures software system allows display (in tables and standard graphical formats) of historical data values since 1960, in combination with forecasts from the base case and from alternative scenarios over time horizons from 2005 through 2100. It includes a Geographic Information System (GIS), or mapping capability, and also provides specialized display formats, such as age/sex and age/sex/education cohort structures and social accounting matrices.

The system facilitates scenario development and policy analysis via a “scenario tree” that simplifies changes in framing assumptions and agent-class interventions. Users can save scenarios for development and refinement over time, including the normative education scenario developed and analyzed in Chapter 6 and Chapter 7. Standard framing scenarios, such as those from the United Nations Environment Programme’s *Global Environmental Outlook-4*, are available with the model for users to explore and potentially to develop further.

■ The IFs system facilitates explorations of interventions and scenarios. ■

The Education Model

The education model of IFs simulates patterns of education participation and attainment in 183 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.

In Figure 4.2, we display the major variables and components that directly determine education demand, supply, and flows in the IFs system. We emphasize again the interconnectedness 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 sociopolitical systems provide funding for education, and levels of educational attainment affect economic productivity and growth and therefore also education spending.

Table 4.2 summarizes the most important aspects of the accounting system, the dominant relationships, and the key dynamics that our education model represents. At the *accounting*

level, the major flows within the model are student and budgetary flows, and the major stock is that of gender-differentiated educational attainment of the adult population. The model structurally represents the formal education system from the primary through tertiary levels, and it further divides the secondary level into lower secondary and upper secondary levels and into general and vocational categories within each of the secondary levels. It tracks students by grade and by sex. Intake (or transition to a higher educational level) and persistence or survival rates are the two variables that most immediately determine the patterns of student participation and progression through the grades.

The *dominant relationships* in the model are those that determine the intake (or transition) and survival rates and the costs of education per student, all three of which are closely connected to per capita income. The model also takes into account the long-term nonincome drivers of education in an aggregate fashion. As the model simulates the gradual expansion of education, the intake and survival rates saturate following an S-shaped pattern.⁹

With respect to *key dynamics*, the processes of the demographic and economic models, as sketched earlier, significantly affect the forecasting of education. Similarly, the dynamics

Figure 4.2 Direct drivers of education demand and supply in IFs

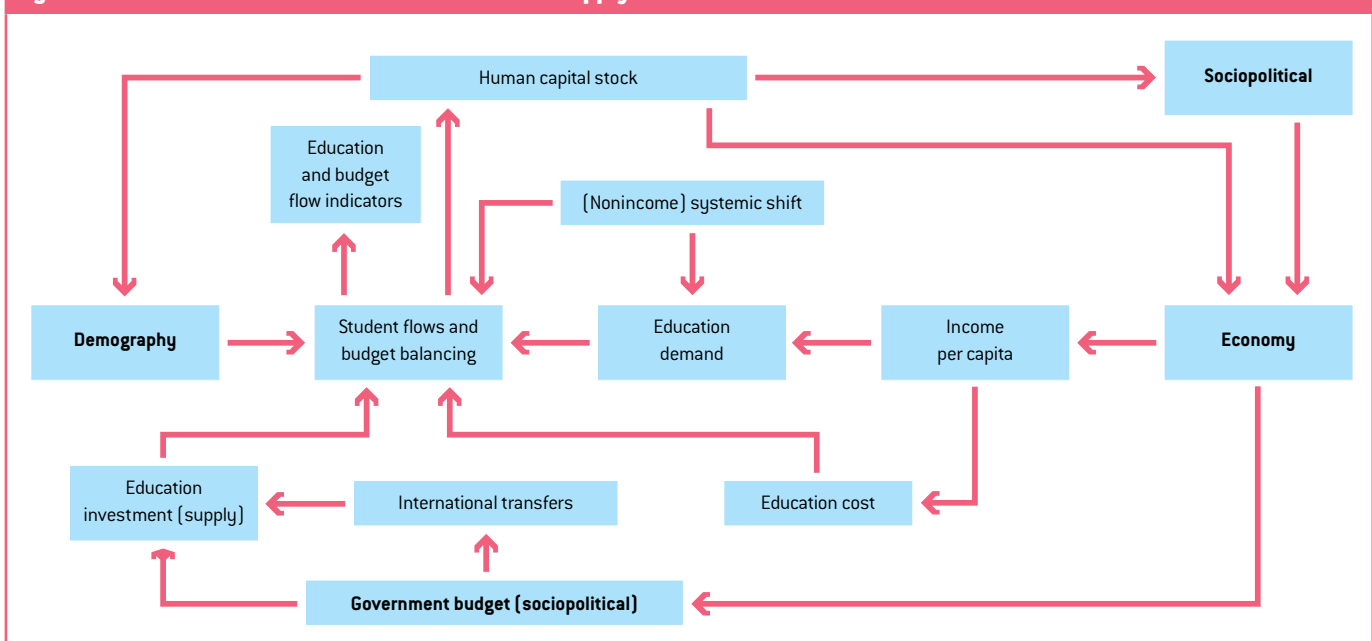


Table 4.2 Foundational elements of the IFs education model

Education model aspect	Key elements
Accounting system	Flows of students into, through, and out of schools Flows of public spending into education system Stocks of adults with different levels of education attainment
Dominant relationships	Intake demand is driven by household income and nonincome systemic factors and follows an S-shaped pattern toward a saturation point Survival rate is driven by the same factors as intake with income being the most dominant Education cost is driven by per capita income with a different cost function at each level of education
Key dynamics	Demographic change Economic development Public education spending constrained by revenue receipts, government consumption, and demands from other public sectors Equilibration between the demand and supply of education funds

of the government budget process in the sociopolitical model produce a key variable for the education model, namely, the availability of funding for education. Within the education model itself, a central dynamic is finding the balance between the demand for education and its availability or supply and then adjusting growth in intake (or transition) and survival rates, as well as spending per student, to be consistent with that balance.

Modeling of complex, integrated, dynamic systems for long-term forecasting is seldom a matter simply of specifying equations. It generally requires development of *algorithmic structures* (logical procedures for integrating calculations and maintaining accounting systems), as well as equations. In the IFs educational model, one such algorithm manages student progression through the grades. On the budgetary side, another balances the forecasted funding demand and funding availability in order to shape the enrollment and spending levels. And still another addresses the flows of graduates into and through the adult population. Subsequent sections of this explication of the education model provide basic information on these processes.¹⁰

Accounting system

As Chapter 2 outlined, a conceptual description of student flows begins with entry in the first grade of primary school. At the end of each year, students either progress to the next grade, repeat the current grade, or drop out.

Eventually, some proportion of the entering cohort reaches the beginning of the final grade; that proportion constitutes the “survival rate.” Further, of those who persist to the beginning of the last grade of primary school, most subsequently graduate and become eligible to continue to the lower secondary level. The “transition rate” identifies the portion of those completing the primary level that actually continues into general programs at the lower secondary level,¹¹ following which a new pattern of grade-level progression, repetition, and dropout ensues. Conceptually, similar flows (albeit at different rates) take place at the upper secondary level and at the tertiary level.

IFs accounts for education participation by simulating gender-specific grade-by-grade student flows, using country-specific entry ages and years of schooling at each level to represent enrollments and to distinguish gross and net flow indicators. We dynamically forecast intake rates (or transition rates to general programs at the lower and upper secondary levels) and survival rates, and we calculate enrollment rates as the combined result of those flows, tracking students through grades. Clearly, this approach provides more useful information than a focus on enrollment rates alone would, as the same enrollment rate might result from different combinations of intake and survival rates. It also provides points for representing interventions that shape the actual dynamics of enrollment.

■ Long-term dynamic forecasting often requires development of algorithms as well as the specification of accounting structures and equations. ■

IFs simulates gender-specific grade-by-grade student flows, using country-specific entry ages and years of schooling.

Student flows

A truly full representation of student flows (see, again, Figure 2.3) would represent movement across grades over time with grade-by-grade and country-by-country specification of repetition rates, dropout rates, and rates of return and late entry by overage students. It would culminate with rates of completion and transition by some to higher levels. In addition to being very intensive with respect to initial data and ongoing computations, there is a limited basis for forecasting idiosyncratic patterns of repetition, dropout, and reentry by grade.

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 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 gradewise enrollments are then used to determine the total enrollment at the particular level of education.

There are some obvious limitations to our simplified approach. Although our model effectively includes repeaters, we represent them implicitly (by including them in our grade

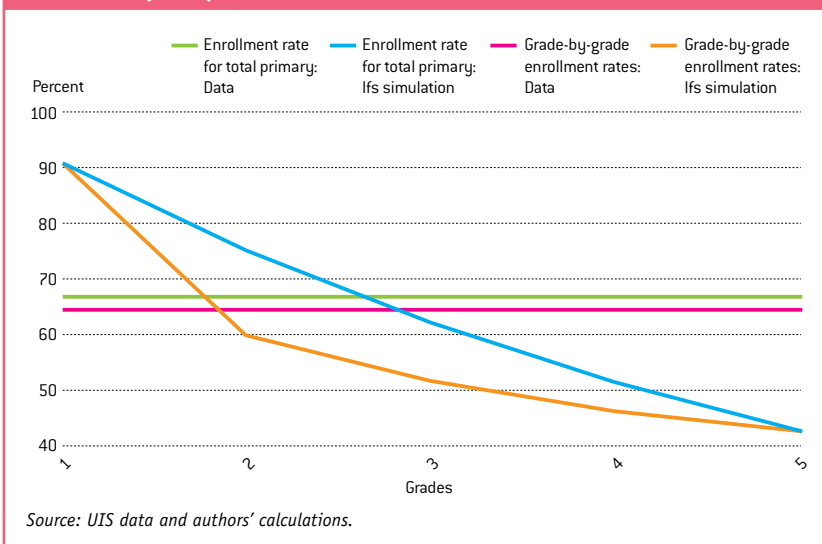
progression) rather than representing them explicitly as a separate category. Moreover, by setting first-grade enrollments to school entrants, we exclude repeating students from the first-grade total. On the other hand, the assumption of the same grade-to-grade flow rate across all grades might somewhat overstate first-grade enrollment in a typical low-education country, where first-grade dropout rates are typically higher than the dropout rates in subsequent grades. Since our objective is to forecast enrollment, attainment, and associated costs by level rather than by grade, we do not lose much information by accounting for the approximate number of school places occupied by the cohorts as they proceed and by focusing on accurate representation of total enrollment.

Figure 4.3 juxtaposes the primary grade-by-grade enrollment data in Bangladesh in 1988 as constructed from UIS-reported intake and gradewise survival rates against the grade-by-grade pattern that IFs simulates. It illustrates that the net effect of our simplifying assumptions generally produces reasonable results with respect to overall enrollment rates, headcounts, and hence also resource requirements. The initialization of the model, discussed later, further protects initial data on enrollment. Even so, our approach results in an (usually small but occasionally quite significant) initial discrepancy between reported and calculated enrollment, as seen in Figure 4.3. IFs computes that differential as an additive factor so as to assure that our computation and the data are consistent; the model carries forward the additive factor but causes it to converge to zero over time.

A separate algorithmic structure helps represent gross enrollment patterns at the primary level. Specifically, 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 it brings some back as students (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 student flows at the lower and upper secondary levels, including country-specific entrance ages and durations at each level. Two adaptations were necessary. First, UIS provides only gross enrollment data for lower and upper

Figure 4.3 Example comparing primary grade flow data and IFs simulation: Bangladesh (1988)



secondary enrollment rates, so our core lower and upper secondary forecasts are also gross rates only (a relationship estimates total secondary net enrollment from the gross values). Second, although UIS provides transition rate data from the primary to the secondary level—which in effect is the transition rate into lower secondary—it does not provide transition rates from the lower secondary to the upper secondary level. However, UIS does provide grade-by-grade secondary headcount time-series, from which the IFs model calculates historical lower to upper secondary transition rates as the starting point for forecasts of future rates.

In the ISCED taxonomy of educational programs (see the Appendix of Chapter 2), tertiary education displays the greatest complexity. Not only are there two categories (programs that lead to an “advanced research qualification” and programs that do not), there are also two subcategories within the programs that do not lead to an advanced research qualification. One subcategory—itsself quite broad—encompasses theoretically based programs and programs that prepare students for practice in high-skill professions; the second category includes programs that are practical, technical, and “occupationally specific.”

To cut through some of this complexity, UIS in some treatment of data and IFs in its representation of student flows both make simplifying assumptions at the total tertiary level.¹² For example, rather than using country-specific and tertiary category-specific program durations to calculate flows, both UIS and IFs base calculations of tertiary flow rates on an assumed five-year program period. To initialize the model, we first use the total UIS headcount of graduates of all programs to calculate an overall gross tertiary graduation rate based on the assumed five-year program period. We then use our calculated overall graduation rate with the total tertiary gross enrollment rate from UIS to calculate an overall tertiary gross intake rate.¹³

Education attainment

The algorithm for the tracking of education attainment is very straightforward. 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 one-year time steps after subtracting deaths. The primary weakness of the approach, common to many but not all other models, is that it does not represent differential mortality rates associated with different levels of education attainment (generally lower for the more educated).¹⁴ This leads, other things being equal, to a modest underestimate of adult education attainment, growing with the length of the forecast horizon. The method that IFs uses to advance adults through the age/sex/education categories also slightly misrepresents the level of education attainment in each five-year category.¹⁵

Financial flows

In addition to student flows, and interacting closely with them, we want to track financial flows. In IFs, we conceptualize those flows as being the result of the interaction of demand- and supply-side forces, a dynamic to which the discussion will return. The accounting side is relatively simple. Given forecasts of spending per student by level of education and given enrollments by level, an estimate of the total “demand” for education funding is simply the sum across education levels of the products of spending per student and student numbers. This so-called demand for educational funding is, however, a crude conceptualization. The flow structure of the model does not truly represent a demand for education (see, again, Figure 4.2) because initial conditions clearly reflect historical financial constraints. As with any other collective good, societies tend to underprovide education relative to the point at which expenditures would truly equal their potential benefits.

Hence, a more accurate conceptualization is that IFs represents a demand-driven, supply-constrained system. In the future, the extent of supply constraint may wax or wane (sometimes even providing largesse), and the differing budget situations will affect both expenditures per student and enrollment levels.

Turning to the budget, governments provide most education funding. Public expenditures on education as a portion of GDP vary greatly across countries (see Figures 3.12 and 3.13),

■ IFs represents population structures by age, sex, and years and levels of completed education. ■

■ *The government budget submodel of IFs determines the amount of funding available for education.* ■

averaging around 5 percent in recent years and ranging from under 2 percent to over 13 percent. Although fewer than 25 percent of all countries report data on private funding, it appears that private funds account at most for about one-fourth of all education expenditures and that they are concentrated at the upper secondary and tertiary levels (see Chapter 3 for further discussion). 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. In reality, the picture is more complicated. At the primary level, and perhaps also at the lower secondary level, it is more likely that tuition fees charged for public education will be phased out over time. In fact, the targeted increases in our formulations of per student costs in those countries with low current per student expenditures may be a proxy for a shift to public support. However, at the upper secondary and tertiary levels, private funding in the form of tuition fees may increase in some countries in order to expand capacity.

Dominant relationships

Before turning to the formulations of the model for forecasting intake and survival, it is useful to note that two alternative methodologies frame effectively all long-term forecasting. The first is extrapolation, and the second is causal analysis. Each has a variety of advantages and disadvantages, and our earlier review showed that both have been used to forecast the spread of education.

A key advantage of extrapolation is the relative simplicity involved in fundamentally univariate analysis, relying only upon the history of a variable in order to determine its future. It should be noted, though, that the frequent use of specialized formulations in extrapolations (such as the S-curve) implicitly builds in the effects of other variables, such as the constraint of bringing difficult-to-reach populations into school and the shift of resources to higher levels of education, which are both implicit in the slowing of growth as an enrollment rate approaches 100 percent.

The fact that it is univariate is also an important disadvantage of extrapolation. For instance, when historical series are short and

especially when they are sparse, both of which tend to be true for intake or transition and survival rates, the basis for extrapolation is liable to weaken (although turning to household survey data can extend data series relative to purely administrative data).

Chapter 2 already reviewed some of the most significant issues associated with causal analysis, including problems in sorting out the direction of causality in bivariate analysis, the possibilities of spurious relationships and complex interaction effects in multivariate analysis, and complications introduced by long lag effects.

However, among the important advantages of a causal approach such as that in IFs is the ability to “play with” the driving variables in a causal analysis, allowing development of a range of scenarios linked to important drivers, some of which may in turn be linked to potential policy levers. In long-term analyses, causal approaches can sometimes more clearly represent the structure of a system, incorporating interaction effects and constraints such as that between the supply of funds and the demand for funds in the IFs education model. Education systems are, in fact, subject to a variety of such interactions and constraints. For instance, students cannot enter higher levels of education unless they complete lower ones. Moreover, there tend to be patterns of relationships between intake and survival rates, as well as between enrollment rates at different levels of education, that purely extrapolative formulations might, in long-term forecasting, not reproduce.

Even much of traditional causal analysis, if it were undertaken purely on the basis of independent formulations for intake and survival at different education levels, would strain to maintain such relationship patterns. Instead, causal analysis embedded in algorithmic (rule-representing) logic and attentive to the patterns of causal or dominant relationships across levels of education can be useful, and that is fundamentally the approach of the IFs education model.

Intake and survival

As the discussion of student flow accounting emphasized, the rates of intake of students into primary education (or the rates of transition of primary students to higher levels), the patterns

■ *A causal approach, such as that of IFs, allows development and analysis of a range of scenarios linked to drivers of change.* ■

of grade-by-grade progression, and the rates of survival through the grade progression to the final grade (as well as rates of completion of that final grade) collectively determine enrollment rates and numbers. The relationships that underlie those forecasts of intake or transition and of survival are especially important or dominant ones.

The forecasting of adjusted primary net intake rates begins to illustrate the IFs causal approach (later in this chapter we explain the concept and use of adjusted primary net intake). It has several elements. The first is use of cross-sectional analysis to specify the relationship between gross domestic product per capita and adjusted net intake at the primary level (see Figure 4.4).¹⁶ Such cross-sectional representations, looking at relationships between variables across countries at a given time point, help us understand something about the typical long-term developmental patterns of countries globally and thus give basic insight into likely longitudinal dynamics.¹⁷ There is a clear tendency for primary intake rates to increase with GDP per capita, particularly at lower levels of GDP per capita (below about \$5,000 at purchasing power parity). This relationship reflects, in part, changing economic structures and changing demand for the skills acquired through education, as well as the growing ability of richer societies to provide education. The specific basic function in Figure 4.4 is

$$ANIR_t = 41.8 + 5.77 * \ln GDPPCP_t$$

where

ANIR is adjusted primary net intake rate
GDPPCP is GDP per capita at purchasing power parity

Although GDP per capita is a powerful driver and/or correlate of a great many aspects of social change (Hughes 2001), the relatively low R-squared values in Figure 4.4 (and for most such relationships between GDP per capita and intake, transition, and survival rates, the R-squared values fall in the range of 0.15–0.35) suggest there is much room for extended analysis of potential dynamics of intake. We have explored the addition of other factors, such as the education of women as

captured in the percentage of women fifteen and older who have completed secondary education. On the whole, that factor tends to be comparable in power to GDP per capita (and is highly correlated with it), but such factors tend not to add a great deal to the multiple R-squared. Still, we know from many empirical analyses that parents' education is a key determinant of intake rates (Clemens 2004: 4), and the omission of its explicit treatment from the IFs formulation for intake is almost certainly a weakness.

We estimated, and the IFs model uses, a full set of gender-specific, cross-sectional functions (see Figure 4.5) as the first step in forecasting the flow rates at different levels of education. The functions for intake rates show the expected progression with GDP per capita. That is, at lower levels of income, countries show higher typical rates of primary net intake than they do rates of lower secondary gross enrollment, which in turn exceeds rates of upper secondary gross enrollment. The patterns for survival rates are more complex, and it is, of course, possible that countries have higher survival rates at the secondary level than at the primary level. The functions also show the advantages that females tend to develop even in middle-income countries at the tertiary level, both in intake and survival rates.

Forecasting rates of intake or transition and of survival is foundational to IFs forecasts of enrollment.

Figure 4.4 Relationship of female adjusted primary net intake and GDP per capita at PPP

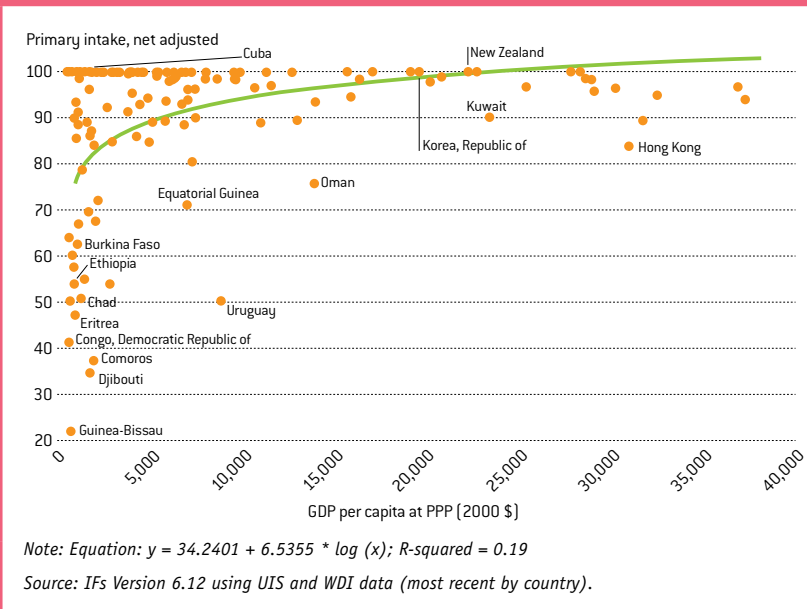
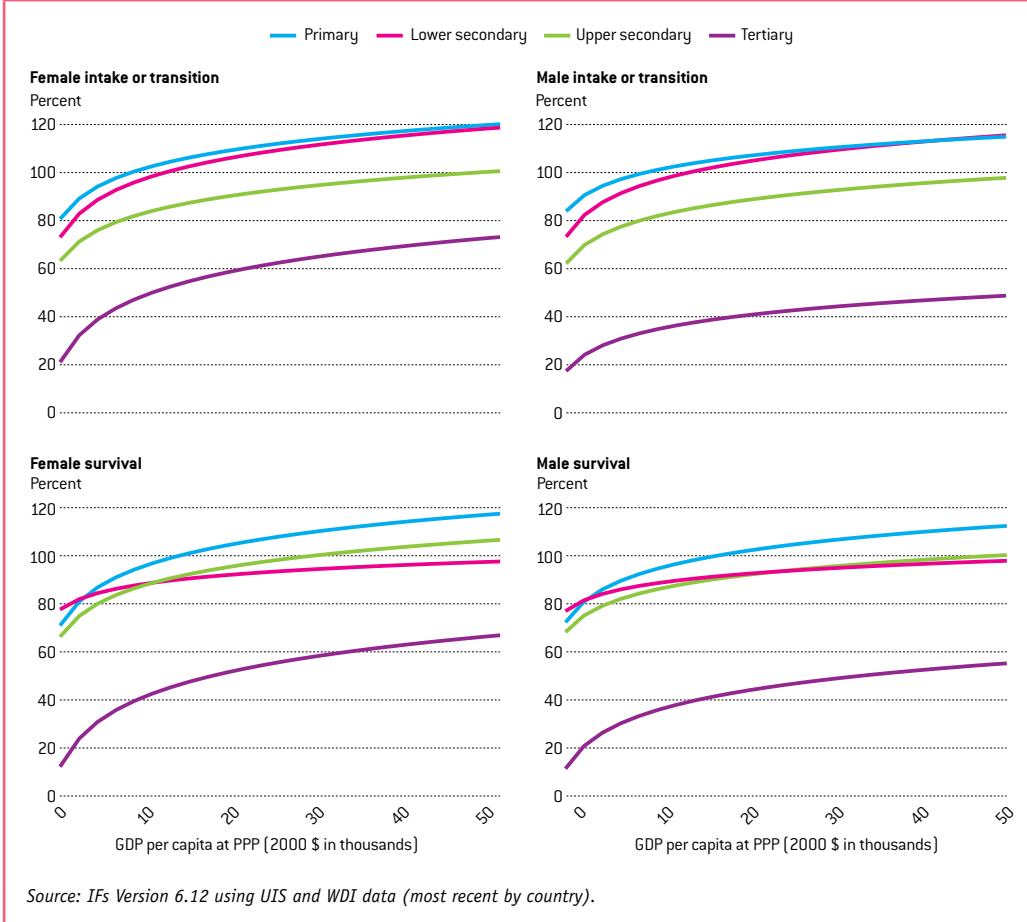


Figure 4.5 Relationship of intake or transition and survival rates at different levels of education with GDP per capita at PPP



This set of functions provides a fundamental group of expectations for intake and survival that provides an initial foundation for forecasts. In addition, these functions help maintain relational integrity of forecasts across education levels. Again, however, the great variation of countries around the functions makes clear the need for additional steps.

Instead of adding additional variables to the regression analysis, our formulation turns to other factors and approaches. First, considerable path dependency exists at the country level. Returning to Figure 4.4, note that Cuba is positioned well above the regression line, due to socialist policies that support universal education. Oman falls well below the line, as other Middle East countries often do. Geographic factors, ethnic and religious patterns, and cultural traditions influence intake rates, helping to create such country-specific and region-specific path dependencies. IFs partially

protects those patterns by computing additive adjustment factors in the first forecast year that represent the position of empirical values relative to the relationship. As GDP per capita grows in forecasts, these adjustment factors continue to maintain the position of countries relative to the relationship.

At the same time, however, such differences can be idiosyncratic and temporary, and deviations from larger systemic patterns often erode. Thus, the model uses a convergence process to bring outliers gradually to the values of the function. We have estimated the duration of convergence periods subjectively based on model behavior, and these periods vary across levels of education. We generally anticipate convergence to be faster at lower levels of education, where emphasis is greater and enrollment rates on average are higher, than at higher levels.¹⁸ We allow primary adjusted net intake to converge only in an upward direction

on the assumption that demand for education in countries will very seldom actually decline (extreme budget pressures, as we shall see, can sometimes force a decline).¹⁹

A general upward bias in convergence patterns is also consistent with another factor built into the formulation for adjusted primary net intake, namely, a “systemic shift.” Figure 4.6 shows the upward systemic shift of intake rates for males in relation to GDP per capita during the 1990s, reflecting the additional emphasis that individuals and governments have placed on education in recent years regardless of income levels. Some of this shift may be due to a greater need for education in order to compete for jobs in an increasingly knowledge-based economy. Some, however, may also reflect simple competition for relative position by increasingly well-educated individuals—the “credentialism” or “sheepskin” effect (Hungerford and Solon 1987). It also seems reasonable to suspect that the greater emphasis on education in recent years has an ideational component, not solely material ones. We compute an ongoing systemic shift at the primary level based on the pattern of recent years, subject, of course, to saturation effects as levels move higher.

More generally, the education model uses this constellation of elements (GDP per capita, historic uniqueness of countries and their movement toward convergence over an extended time frame, and the representation of systemically shifting patterns) in its basic formulations for intake, transition, and survival rates. Illustratively, that for adjusted primary net intake is:

$$ANIR_t = F(GDPPCP, ANIR_{t-1}, Converge, SS)$$

where

- ANIR is adjusted net intake rate
- GDPPCP is GDP per capita at purchasing power parity
- Converge is a fractional movement toward the estimated function
- SS is systemic shift (upward) of the function across time

On top of these formulations, used in exploratory analysis, the model also makes it possible for the user to target growth rates for more normative analysis—for instance, by replacing

the function with specification of a 2 percentage point annual increase of intake rates across a forecast horizon (the 2 percentage points would be effective at the midlevel in the range of intake rates, tapering to zero as the intake rate approaches 100 percent, because of an S-shaped representation). Chapters 6, 7, and 8 reflect this type of normative use of the model.

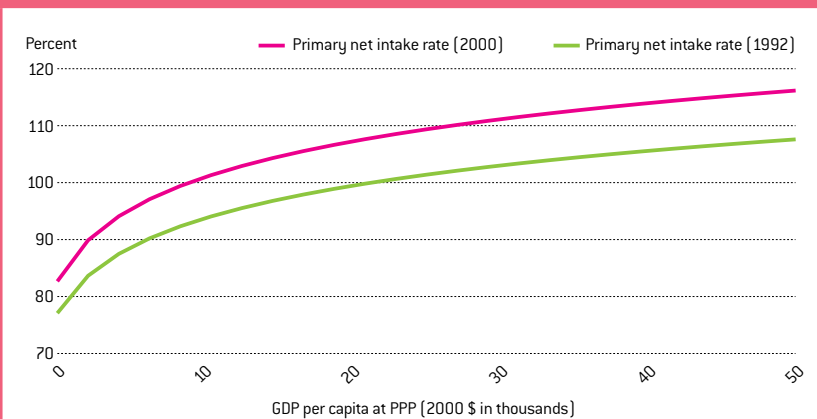
Costs of education and public spending

The education of each student has a cost, differing by level of education and generally rising across levels. Countries vary greatly, however, in what they spend per student at each level, and patterns also can change rather dramatically over time. Because spending per student is a key variable, its determination in forecasts is another relationship that greatly influences or dominates model behavior.

In the context of developing a normative scenario for education futures, Chapter 6 provides an extended discussion of spending on education, both per student and in the aggregate. With respect to spending per student, it attempts to tease reasonable target levels out of existing analyses of good practice and from cross-sectional patterns of spending by educational level as a function of GDP per capita (see Figures 6.5–6.7 for cross-sectional patterns). Typically, and especially at the upper secondary and tertiary levels, spending per student begins at quite high proportions of GDP per capita when income and enrollment levels are low, reflecting high cost structures. It falls as enrollment and

IFs reflects global patterns of change through gradual convergence of country-specific path dependencies to “typical” paths over long time frames. ■

Figure 4.6 The changing relationship of adjusted primary net intake and GDP per capita at PPP over recent years



Source: IFs Version 6.12 using most recent UIS and World Bank data available in IFs.

■ IFs uses analyses of good practice, cross-sectional patterns, and country-specific starting points to forecast per student spending at each education level. ■

income rise at higher levels of education, but it can rise with income at the primary level.

For the purposes of both exploratory (where we seem to be going) and normative (where we might like to go) forecasting, the IFs project draws upon these analyses of good practice and cross-sectional patterns to anticipate future levels of spending per student. For exploratory analysis, we assume that levels of spending per student will very gradually converge from empirical initial conditions to target levels. For normative analysis, we posit considerably more rapid convergence.

Key dynamics

Education and broader systems of human development interact closely over time. For that reason, the International Futures system integrates the education model with the detailed demographic, economic, and sociopolitical models. The dynamics of those other models or systems become, in essence, part of the dynamics in the education model. Earlier discussion sketched the structures and features that determine the dynamics in those other models. In this section, we return briefly to note some of the relationships across models.

With respect to important dynamics within the education model itself, sound accounting systems assure us that student flow patterns are internally consistent and connected well with tracking of adults' education attainment. They also help us to track government revenues and expenditures and to identify how much funding governments might direct to education. The key specifications of demand for education at all levels and of potential spending per student will, if resources are available, dominate the forecasting patterns and allow calculation of the funding required if societies were to meet that demand. In an integrated modeling system (and in the real world of competing demands for resources), however, initial calculations of supply and demand seldom do balance. Instead, there will be complex dynamics of interaction between the two at any point in time and across time, and this section sketches that process as well.

Linkages (backward and forward) of education to other systems

The discussion of the education model to this point has indicated the important linkages

from the economic model to education demand and to spending per student, both involving GDP per capita (backward linkages from the perspective of the education model). Enrollment rates translate, of course, into student headcount only with the help of age-sex structures from the demographic model. This volume most often presents and discusses forecasts of enrollment rates, but the tracking in the IFs system of enrollment numbers is essential to analysis of the ability of societies to meet demand for education.

The government budget submodel of IFs determines the amount of funding for education. That submodel forecasts total government revenues and expenditures, using a social accounting matrix to embed them in the larger system of domestic and international financial flows and to maintain accounting consistency. Expenditures include both transfer payments and direct spending on the military, health, education, research and development, and other programs. Spending on education tends to increase as societies become wealthier, but so does other spending, including that on health. The government budget model balances competing demands.

In addition to these backward linkages to other systems, the IFs system represents a number of forward linkages from education to demographic variables (for example, fertility), economic variables (for example, productivity), and sociopolitical variables (for example, democracy level). Chapter 8 will return to these linkages. The existence of both backward linkages from education to drivers in other systems and of forward linkages from education to other systems creates important feedback loops (see, again, Figure 4.2), which Chapter 8 will also explore.

The reconciliation of budget demand and supply

What if spending demand from the education model and spending supply from the government budget submodel do not match? The reality is, of course, that they will not, which requires an algorithmic process of reconciliation. Imbalances between supply and demand for funding set up a multistage problem of allocation in the model, as in budgeting systems. Moreover, they set up both the need for immediate balancing and the need for incremental changes in longer-term

patterns. The theoretical framework for handling budget reconciliation is fundamentally that of incremental decisionmaking and budgeting.²⁰

The first decision issue for immediate resolution is determination of the total public spending on education. Given the tendency for governments to spend on average just under 5 percent of GDP on education quite independently of GDP per capita (except for the poorest countries) and the persistence of that average level across time, it is clear that there is a substantial degree of top-down influence in determination of the ultimate budget. At the same time, however, the need for governments to be responsive to changing demographic patterns and enrollment growth is obvious. Thus, for exploratory analysis, we presume that the balance of forces is predominantly governmental but not exclusively so (with a roughly sixty-forty weighting). One immediate forecasting implication of this is that countries such as the Democratic Republic of Congo, which mobilize resources very poorly and direct very little of them to education, cannot be expected to have strong educational futures unless those patterns change.

A second-stage decision issue for resolution is the balance of spending across levels of education. The model does that (whether the budget is in deficit or surplus) proportional to the initial demand for funding (student numbers times cost per student). Thus, the forecasting of intake/transition and survival rates and the student numbers to which they give rise, in combination with the forecasting of costs per student, determine this allocation.

A third-stage decision issue is allocation of spending surpluses or deficits at each level of education between student numbers and costs per student. The algorithmic structure at this stage is somewhat complicated by the need to contend with imbalances already existing between enrollment and spending patterns and targets for them; for instance, if enrollment is already above the target and spending is below, it would make no sense to adjust both upward in case of a budget surplus. Some preliminary adjustments incrementally correct such imbalances, and then, on the whole, proportionally comparable adjustments change enrollment drivers (intake/transition and survival) and spending per student.²¹

Finally, in order to facilitate adjustments in future years, we compute moving average multipliers that carry forward the magnitude of adjustment to each term over time. These multipliers smooth the adjustment processes across time, allowing the overall system to chase equilibrium, even if it never completely finds it.

In normative analysis, the budgetary link between demand and supply sides can be turned off, and the demand can force the spending on the supply side. Chapters 6 through 8 explore this use of the model. Even in this situation, however, it is important that the required spending on the demand side be accounted for in the government budgeting model, thereby reducing funds available for expenditure in other areas and/or requiring additional government revenues. The social accounting matrix of the IFs economic and sociopolitical models makes it possible to do this and to trace the consequences (such as the impact on health spending) of higher or lower spending on education.

Initializing the model

Initializing forecasts when data are scarce is hard work. But some algorithm ought to be able to do it. The historical series that constitute the IFs database begin with 1960, whereas the base year of the IFs education simulations (and those of the larger IFs system) is 2005. Data values from 2005 initialize the model for forecasting. The base year values of student and education financial flows come from the UIS, base year demographic data are primarily from the United Nations Population Division, economic data come heavily from the World Bank, and sociopolitical data are from many disparate sources.²² Before the model can use these data, however, significant data extension, cleaning, and reconciliation are necessary, including the estimation of base year values when they are missing for a country.

Data extension

UIS provides student and financial flow rate data for many measures of education participation, particularly at the primary level. In some instances, we had to modify the data to make it meaningful for a long-term model. For example, our representation of net intake rates combines the very strictly of-age rates that the UIS reports with students one year over- and one year

● **Reconciling demand for education with availability of funding is a multistage and incremental process.** ●

● **When data are missing or are incongruent across related variables, algorithms help with data extension, estimation, and reconciliation.** ●

Table 4.3 Primary intake rates by age categories (2005)

	Of-age	Overage 1 year	Underage 1 year	Overage or underage 2 or more years
Arab States	62.9	9.4	17.7	6.6
Central and Eastern Europe	73.9	12.0	9.0	1.5
Central Asia	67.0	17.1	19.7	6.3
East Asia and the Pacific (Poorer)	59.8	18.8	14.9	12.8
Latin America and the Caribbean	69.6	16.7	14.4	7.6
South and West Asia	72.2	23.3	1.6	17.4
Sub-Saharan Africa	49.0	25.2	10.6	18.5
East Asia and the Pacific (Richer)	88.8	14.3	0.8	1.4
North America and Western Europe	78.9	16.4	1.8	2.2
World	64.3	17.8	12.4	9.9

Source: Compiled from UIS data (unweighted country averages).

underage (also reported by the UIS) to create an adjusted net intake rate. Box 4.1 provides more detail on the concept and the rationale for our use of the adjusted net intake rate.²³

In other instances—for example, with respect to the division between lower and upper secondary levels—UIS provides raw data (e.g., headcounts) but fewer flow rates than at the primary level. In the instances when UIS provides flow rates, we import them directly into the IFs historical series. When UIS provides only raw data (e.g., grade-by-grade headcounts or total expenditures and total number of students), we calculate rates from these series offline, using spreadsheet applications or auxiliary programs. We then enter the results into the IFs historical series. Examples of data series handled this way include the lower and upper secondary survival

rates, the transition rate from the lower to the upper secondary level, the tertiary intake rate and the overall tertiary graduation rate, and differentiated per student costs at the lower and upper secondary levels. We also use such auxiliary processing to create the data series for the adjusted primary net intake rate.

Data cleaning and reconciliation

Whenever possible, we use an important automated subsystem of IFs that we call the *preprocessor* to help prepare initial conditions. The preprocessor uses algorithms that simplify the preparation of initial conditions from the raw data. Among other benefits, the preprocessor makes possible rapid recomputation of initial conditions when a new data update becomes available. The two major functions of the preprocessor are (1) filling missing base year values, and (2) reconciling incongruent data or estimates.

Filling missing base year values

When 2005 data are missing for a country, the IFs system estimates 2005 values rather than excluding the country from forecasts. We apply the following estimation techniques, normally in the order listed: (1) using the most recent data point for the country if it is temporally proximate; (2) calculating an imputed data point from a longitudinal temporal regression if a recent data point is not available but a longer historical series exists; and (3) estimating the data value from a cross-sectional relationship stored in the system, most often as a function of GDP per capita at PPP.

There are some specialized algorithms in the preprocessor or first model year to handle particular issues. For instance, one

Box 4.1 Adjusted primary net intake

In developing countries, the difference between gross and net intake rates is often great, particularly in early stages of the transition to broader education participation as “overage” students take advantage of an increased emphasis on education and growing opportunities to enroll. However, as is strikingly evident from Table 4.3, far more entering children are just one year over or under system-defined entry ages rather than two or more years older or younger. This entry pattern often persists indefinitely, even as the rates of children two or more years away from the system-defined entry age decline over time.

Were our focus only on the entry of students of precisely official entry age, we would, in our view, discount the progress countries are making with

respect to the timely entry of “appropriate-age” students. Hence, in place of the conventional net intake measure, we simulate an “adjusted primary net intake rate,” which is the intake rate of children at the official system-defined entry age plus the children one year above or below that age. The difference between the adjusted net intake rate and the gross intake rate then becomes the indicator of divergence from age-appropriate universal primary intake. In our model, simulated primary gross entry rates gradually converge toward the adjusted primary net intake rate as more students enter “on time” and the pool of potential late entrants diminishes.

segment of code computes the annual size of an “overage” pool of out-of-school children potentially available for primary intake. A particularly important specialized process takes the education attainment data (which are not provided from original sources by age) and spreads the attainment levels across age categories in order to initialize the ongoing calculation of attainment described earlier. This spread process takes into account the percentage of the adult population with a certain level of education, the current completion rates at that level of education, and the age structure of the adult population. Knowing that completion rates almost always exceed the average attainment levels (that is, education participation is increasing over time and therefore decreases across progressively older cohorts), a factor for age-related decline in attainment levels can be computed in an iterative process.²⁴

Reconciling incongruent data or estimates

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 to fill holes. 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 straightforward than either intake or survival data; in turn, intake rates have fewer missing values and are arguably more reliable measures than survival rates.

Conclusion

The purposes of our modeling are to enhance understanding of where the global education transition appears to be taking education

systems and to create and explore a normative scenario that might accelerate that transition. Past models and tools provide a strong basis for insights into how to structure such a model, and in building the education model in IFs, we have drawn upon them. In the process, we have created a model that is structurally based and agent-class-driven, that represents formal education at all levels, and that is integrated quite tightly with models of demographic, economic, and sociopolitical systems.

Although we believe our education model to be a strong one, it is imperfect. Modeling and forecasting systems simplify reality, in part to allow us to better understand its dominant structures and dynamics. In fact, this is the fundamental reason we do modeling. Simplified representations help us clarify and extend our own mental models of the system of interest to us. They also allow us to think about how systems might be unfolding and therefore to produce forecasts with and without modeled interventions. However, we should never confuse forecasting with prediction.

In this spirit, we will turn in succeeding chapters to the use of the IFs modeling system and to its forecasts of further advances in global education, including some comparison with forecasts produced by other tools and approaches. Chapter 5 and subsequent chapters will look explicitly to the future, presenting both the base case of IFs and alternative forecasts developed around it.

Appendix to Chapter 4: Education Modeling and Forecasting Approaches *McMahon*

Education and Development: Measuring the Social Benefits describes McMahon’s development and implementation of an econometric “interactive macrodynamic model” (McMahon 2002: ix), the purpose of which was exploration of the social benefits of education (McMahon 2002: 179). Analyses of the social benefits of education included both direct and indirect impacts of primary and secondary education on economic development; on population growth via health and fertility; on democracy, human rights, and political stability; on poverty and inequality; on the environment; and on crime.

The book began with a “base case” or exploratory analysis that assumed the

■ *Our structurally based, agent-class-driven model represents formal education at all levels and is integrated with demographic, economic, and sociopolitical models.* ■

continuation of past education policies and the associated unfolding of education and its impacts. Simulations of two specific normative education policy changes and how they might enhance the extension of education and its impacts moved beyond the base case. The first normative scenario was built on a 2 percentage point increase in public investment in education as a percentage of gross national product (GNP); the second normative scenario assumed a 20 percentage point increase in male and female secondary education enrollment rates (McMahon 2002: 185–186).

The analysis used historical data for the period from 1965 to 1995, and its forecasts extended to 2035; the countries included (78) were those for which consistent data were available on all key variables. The model incorporated empirically tested varying time lags between changes in primary and secondary education and changes in other components of development; it included bidirectional feedback loops; its mathematical equations incorporated the concept of long-run equilibrium relationships; and its parameters were estimated from cross-sectional analyses examining relationships between variables across many countries at one point in time. The major contributions of the model were (1) the inclusion of both primary and secondary education, (2) the placement of the broad direct and indirect social impacts of education at the center of analysis,²⁵ and (3) an effort to calculate a net return associated with extending education participation.²⁶

Delamonica, Mehrotra, and Vandemoortele

In a study published by the United Nations International Children’s Emergency Fund (UNICEF), Delamonica, Mehrotra, and Vandemoortele (2001) projected the incremental costs associated with moving from the level of each developing country’s net primary enrollment rate in 2000 to universal primary education for all developing countries (128) by 2015. The project was undertaken to update global and regional cost estimates for the period from 2000 to 2015, in distinction from cost estimates based on enrollment patterns in the early to mid-1990s (Delamonica, Mehrotra, and Vandemoortele 2001: 2). The purpose and approach were normative.

Delamonica et al. used population projections from the UN Population Division as a basis for

their cost projections. They then assumed the increases needed to bring each country’s net enrollment rate to 100 percent by 2015 would occur in a linear fashion. They also made the uniform assumption that all countries would absorb any incremental costs arising from population trends at constant enrollment rates (i.e., that a country could find the resources to educate the then-current proportion of its population of school-age children, no matter what size that school-age population would be over the 2000–2015 period). Their methodology held GDP per capita constant for the period from 2000 to 2015.

Delamonica et al. estimated costs in four discrete categories: (1) recurrent expenditures related to net enrollment rate increases; (2) quality improvements, as reflected by an adjustment in unit costs to allow 15 percent of recurrent costs for nonwage items (e.g., instructional materials) without a reduction in teacher salaries; (3) reducing pupil-to-teacher ratios to an average of 40; and (4) capital costs for those countries where the increase in students from the expanded net enrollment ratio would be greater than the decrease in the school population from trends in the decline of expected births. Items (1) and (4) were added to the costs as increases in the net enrollment ratio brought new students into the school system, whereas items (2) and (3) were added across the school population in the first year of estimated costs (Delamonica, Mehrotra, and Vandemoortele 2001: 12–13). Although the only dynamic element of the model was its use of school-age population projections as the basis for estimating costs of a linear increase to UPE, the model provided a framework for considering various cost components (including quality improvements and capital outlays) and overall resource requirements to be met domestically (through economic growth or reallocation of government funds) or by international donors.

Bruns, Mingat, and Rakotomalala

Bruns, Mingat, and Rakotomalala (2003) authored a seminal normative study entitled *Achieving Universal Primary Education by 2015: A Chance for Every Child*. The study utilized a simulation model developed by Ramahatra Rakotomalala and subsequently adopted for use by the countries selected to participate in the

Education for All Fast-Track Initiative sponsored by the World Bank, UNESCO, UNICEF, and the regional development banks.

The study provided a detailed analysis of the forty-seven low-income countries that were furthest in 2000 from the MDG goal of universal primary education, with an estimate added for Afghanistan (Bruns, Mingat, and Rakotomalala 2003: 20). The study focused on estimating, under certain normative targets or benchmarks, the following: (1) what it would cost to achieve the goal in terms of incremental funding between 2000 and 2015,²⁷ (2) the portion of that funding that developing countries could afford under the assumption of a 5 percent economic growth rate applied across all countries, and (3) where and how much international assistance would be needed.

World Bank task teams collected enrollment data for the then most recent year (usually 2000) directly from the education ministries of the forty-seven low-income countries included in the study. UNESCO-published data (usually for 1997) were used when more recent data were not available from the education ministries. Population data were from the United Nations/World Bank population database used by the World Bank (Bruns, Mingat, and Rakotomalala 2003: 39, 41).

The study began with an exploratory analysis of the characteristics of the low-income countries that were making accelerated progress toward UPE in 2000, compared to countries that were not (2003: 8). From this empirical analysis, a “best practices” or normative framework was created to provide guidelines for policy levers to achieve universal primary completion at “minimum adequate cost” (2003: 109). The framework included benchmarks or targets for quality improvements, for efficiency improvements, and for domestic resource mobilization (2003: 82). The financing benchmarks included a cap on the portion of educational expenditures from government revenues going to primary education in order to avoid stripping resources from secondary and tertiary education. The various benchmarks were combined in different ways to produce alternative scenarios, and incremental costs (including estimates of gaps in domestic funding capacity) were generated for each scenario. The approach connected education to broader

systems via population projections, benchmarks for funding and resource mobilization, and the inclusion of an economic growth rate assumption (albeit a single assumption for all countries). The model also moved toward inclusion of enrollment dynamics by measuring costs associated with a targeted upper limit on repetition rates and a targeted completion rate rather than an overall enrollment rate alone.

Clemens

In 2004, Michael Clemens, through the Center for Global Development, authored a provocative background paper for the Millennium Project Task Force on Education and Gender Equality. His focus was on understanding if there is a typical primary schooling “transition pathway” in developing countries and also on the degree to which the transition to mass primary education can be accelerated by government policies.²⁸ Clemens raised these questions in the context of exploring the feasibility of meeting the MDG goal of universal primary education by 2015; his approach was primarily exploratory, but it included some elements of normative analysis.

To explore these questions, Clemens developed an aggregated flow model that focused on transition speeds as measured by overall net primary enrollment rates, using administrative data compiled by UNESCO field offices from school registers for the years from 1960 to 2000 for over 100 developing countries. Based on the typical developing country experience between 1960 and 2000, he produced S-curve extrapolations²⁹ of the number of years individual countries and multicountry regions might need in order to reach 90 percent net primary enrollment (Clemens 2004: 42, 52). He used the same UNESCO data to model a typical “gender transition speed” in primary and secondary enrollment.

Clemens explored three other dimensions in his evaluation of the feasibility of achieving UPE by 2015: (1) he compared the 1960–2000 transition rates in developing countries with the rates of today’s rich countries during their earlier transitions to universal primary education; (2) he estimated the necessary transition speeds if today’s developing countries are to meet the 2015 goal of universal primary education (Clemens 2004: 55); and (3) he used cross-country data from 1980 to explore

relationships between education transition rates and a number of social, economic, and education policy variables at that point in time (Clemens 2004: 45). He also reviewed an extensive body of literature on drivers of education participation, pointing out in particular the importance of the relationship between household income and parental level of education in family schooling decisions (that is, the role of demand factors).

Clemens's approach did not include costs or the specifics of enrollment dynamics (e.g., intake, repetition, or completion). Instead, it focused needed attention on a number of other critically important aspects of the education transition: (1) the importance of policy attention not just to education availability and supply but also to circumstances that influence demand for education, (2) the importance of placing the transition to mass primary education in developing countries in a longer historical context in order to set aggressive but realistic goals with respect to time frames, and (3) the need to take individual countries' circumstances into account in setting goals.

Wils, O'Connor, and Somerville

A model developed by Wils, O'Connor, and Somerville, and published in a 2005 paper by Wils, Carrol, and Barrow for the Education Policy and Data Center,³⁰ focused on the concept of growth paths toward universal primary education. The context, as it was with Clemens's study, was an exploration of the feasibility of meeting the MDG goal of universal primary education by 2015 (Wils, Carrol, and Barrow 2005).

The model dealt exclusively with student flows at the primary level, and it operated by extrapolation of flow rates without regard to population dynamics or resource requirements and availability. However, it advanced the conceptualization and implementation of flow dynamics in modeling by using and comparing two measures of primary education coverage—entry and completion—rather than using a single overall enrollment rate. Average entry and completion rate patterns were estimated for each of seventy low-income countries for the historical period from 1950 to 2000, and then they were projected forward in S-shaped extrapolations. The descriptor used by Wils et al. for the trajectories of the paths was the number of years it will take each country to go from a primary completion

rate of 10 percent to a primary completion rate of 90 percent, represented as T10-90 (Wils, Carrol, and Barrow 2005: 10).

A significant difference between this and earlier models was its use of household surveys and population censuses (rather than administrative data) as preferred data sources (2005: 2).³¹ Another difference was its use of a "backward-looking" lens to establish historical entry and completion rates. The authors divided the population of 15- to 65-year-olds participating in the household surveys between 1999 and 2001 into single-year age cohorts, and they used the percentage of the cohort that was 14 years old in each year from 1950 to 2000 and that reported having at least some primary schooling in order to estimate primary entry rates for each of those years. They applied a similar methodology, using 19 as the age, to estimate primary completion rates for the same historical period. The use of household data also allowed identification of out-of-school children by various subcategories (e.g., in subpopulations within the country) as well as analyses of inequality across groupings, such as entry and completion gaps between urban males and rural females (2005: 39).

Lutz, Goujon, and Wils

A paper by Lutz, Goujon, and Wils, also published by the Education Policy and Data Center in 2005, elaborated the application of what is described as a "multistate demographic method" to forecast the extent of education attainment among adult populations.³² The model used in this paper was primarily exploratory but included some normative aspects, and it differed from the previously discussed models and tools in a number of important ways. First, its focus was not education system flows but rather the stock of human capital as reflected by the education attainments of a population by age and sex across four categories: no education, primary education, secondary education, and tertiary education.³³ Second, other than purely extrapolative enrollment trend projections produced by UNESCO in the 1980s and 1990s,³⁴ it was, so far as we know, the first model or tool to look across all levels of formal education. Third, by focusing on education levels in the adult population—and, further, on education levels by age and sex—it facilitated exploration of the relationships

between levels of education in a population and other human development systems (e.g., fertility, life expectancy, and economic growth).

The developers initialized the model with UN population data by age and sex extending back to 1937 and produced visual representations of population characteristics in three pilot countries (Guinea, Zambia, and Nicaragua) in 2000, in population pyramids by age and sex in five-year intervals. Education attainment levels from USAID Demographic and Health Surveys were superimposed upon the pyramids, as were estimated fertility levels and infant and child mortality levels by mothers' education from the Demographic and Health Surveys. Moving forward, projections of the population pyramids to 2030 reflected expected impacts of changes in education attainment, with both fertility and child mortality decreasing as education increases.

At the time the paper was written, the model used stylized rather than dynamically formulated assumptions to advance the pyramids to 2030 with respect to initial entry rates and transition rates between one level of education and the next. The authors created three scenarios, using differing stylized or normative assumptions as follows: (1) constant (current) entry and transition rates, (2) trend entry and transition rates, and (3) MDG goal fulfilling entry rates (with constant or trend transition rates).

The IIASA model and its pyramidal displays clearly illustrated that education is a long-term investment by showing the time lag between increases in education attainment among young members of a population and increases in the overall structure and pattern of "human capital stock" in the total population. Further, the authors pointed out that by using the distribution of educational attainment as an indicator of human capital (rather than a single population-wide measure of average years of schooling), it was possible to explore relationships between age, sex, levels of education, and other variables (e.g., health, poverty, and economic growth). They also pointed to the possibility of subnational forecasts, as the methodology can be applied to any population that is clearly defined and for which there is the necessary information by age, sex, and level of education (2005: 33). Work continues at IIASA to extend the model to a large number of countries and to further

exploration of relationships between age, sex, levels of education, and other variables; one example is a backward-reconstruction of populations by age, sex, and level of educational attainment for 120 countries for the period from 1970 to 2000 (Lutz et al. 2007).

Hilderink

A 2007 working paper authored by Hilderink described an exploratory education module being developed and embedded in the established PHOENIX dynamic population and health model at the Netherlands Ministry of Health and Environment (Hilderink 2007). At the time the paper was prepared, the PHOENIX education module used enrollment rates as the single measure of education flows; however, the author stated the plan was to use intake and drop-out rates in a subsequent phase. Education attainment levels and literacy are other components of the model. Geographic regions are the unit of analysis, and the model extends across primary, secondary, and tertiary levels.

The model was initialized with education data from UIS and economic data from the World Bank's World Development Indicators. Simulations for the period from 1950 to 2000 were being used to calibrate and validate the model; the paper provided forecasts for the period from 2000 to 2025, but it mentioned a simulation period extending to 2050. The model includes bidirectional connections between education levels, mortality, and fertility. It also introduces the concept of education demand and education supply by dynamically connecting both enrollment rates and education expenditures to GDP per capita, with the assumption that demand and supply are equal. Although still in a developmental stage, the model is being designed as a comprehensive tool for forecasting education dynamics in conjunction with a number of connections to broader systems.

Wils, Barrow, Oliver, Chaluda, Goodfriend, Kim, and Sylla

An EPDC background paper prepared for the 2008 Education for All Global Monitoring Report described and presented initial results from ProEnrol, a country-level cohort-projection model being developed by the Education Policy and Data Center (EPDC 2007b). The measures of student flows forecast in the paper were

primary and secondary enrollment rates, student headcounts, and gender parity indices, under the exploratory assumption of each country's continuation of its current education policies.

Cohort-projection models focus on the grade-by-grade dynamics of student flows. Countries often use them to meet education system operational planning needs by projecting numbers of students by grade and level and the school resources therefore needed.³⁵ Individual countries may also use them to understand the dynamics of their student flows and to assess education system functioning with respect to these dynamics (entry, promotion, repetition, dropout, reentry, survival, and completion).

ProEnrol, however, is intended for a larger-scale use. As the background paper noted: "The Cohort projection model developed by the EPDC is the first effort to make cohort projections in an international, global series and is done here on an experimental basis. The intention of the GMR [Global Monitoring Report] at this point is to test this method" (2007b: 69). The background paper included projections of primary net enrollment rates and student headcounts for 60 countries using ProEnrol, projections of primary gross enrollment rates and student headcounts for 129 countries using ProEnrol, and projections of secondary net enrollment rates and student headcounts for 82 countries using ProEnrol (2007b: 70). The EPDC made projections for two points in time, 2015 and 2030, and calculated a gender parity index for each projected series.

ProEnrol was initialized with historical enrollment data (1999 forward) from UIS

on pupils by grade (females and both sexes combined) at the primary and secondary levels. It used United Nations medium population projections for the period from 2000 to 2025 to calculate gross and net entering school populations (headcounts) by multiplying the projected population of school entry-age children by projected gross and net entry rates. Extrapolations from past trends in intake rates were used to project future intake rates, and country-specific constant values (equal to the most recent year data were available) were used for promotion and repetition rates.³⁶

The model's only linkage with systems outside education at the time of the background paper was its use of UN population projections of school entry-age children as a foundation for projecting school enrollments. For example, it did not calculate resource requirements or compare potential enrollments with estimates of resource availability. However, it would seem those components might rather easily be added, since the model calculates student headcounts, and per student costs are widely available from UIS. Perhaps more important at this stage, a protocol needs to be developed for projecting reasonable changes in promotion and repetition rates over time, since they are flow components subject to dynamic changes. However, the existence of possible future improvements should not detract from the contribution ProEnrol has already made by developing and now testing significant aspects of a grade-by-grade cohort projection methodology on a global scale, including the first specific representations of promotion and repetition.

1 Either type of model might employ a simple or a sophisticated methodology. At its simplest, an exploratory tool might forecast future enrollment trends by the extrapolation of recent patterns. Likewise, at its simplest, a normative tool might consist of a basic mathematical exercise, such as the calculation of how much primary intake and survival rates need to increase each year between now and 2015 to meet the MDG goal of universal primary education at that time.

2 The time lag between initial changes in intake rates and the possibility of impacts from increased education attainment on other aspects of human development systems is the reason we stipulate that a forecasting tool with a mid- to long-range time frame is critical.

3 In an ideal system, the points of intervention are sufficiently "actionable" that policymakers can readily discern implications for policy choices and implementation strategies. However, even more abstractly defined interventions (e.g., "focusing on increasing survival rates") can be helpful in providing a course for improved outcomes.

4 McMahon (2009) extended his analysis to the tertiary level in a book focused on the social and private benefits of higher education.

5 For an introduction to the character and use of the IFs modeling system, see Hughes and Hillebrand (2006).

6 We emphasize that IFs is not an agent-based modeling system because it focuses on the aggregated behavior of *agent classes*, rather than on the behavior of individual agents as agent-based models do.

7 The various member organizations of the United Nations family are a primary data source, but other sources, such as the World Bank's World Development Indicators, are also used extensively.

8 More technically, the model structure is recursive (it computes equations sequentially in each time step without simultaneous solution). It combines features of systems dynamics (notably, the accounting structures with careful attention to both flows and stocks) and econometrics (using estimated equations for the dynamic behavior of the agent classes).

9 We have an explicit representation of the S-shaped path in our normative scenario only. The more implicit saturation behavior in our exploratory base case results from the integration of various dynamic drivers of education flows.

- 10 Again, the interested reader is referred to www.ifs.du.edu for further documentation.
- 11 The rate is calculated as the proportion of those in the last primary grade who enter general programs at the lower secondary level the following year. At the secondary level, UIS enrollment rates include students in both general secondary programs and vocational secondary programs. However, UIS transition rates to lower secondary are for general secondary programs only, and our model follows that convention in our calculations and forecasts of both lower and upper secondary transition and survival rates. We maintain country-level gender-specific vocational enrollment rates as a constant percentage of official lower and upper secondary school-age populations, reflect them in enrollment data, and use them to bound transition and survival rates in lower and upper secondary general programs.
- 12 A future volume in this series will focus on infrastructure and will include more differentiated analyses and forecasts of tertiary education, making use of program-specific UIS tertiary data to initialize the model.
- 13 UIS has a data series for tertiary entry rates. However, we developed the procedure described earlier because the UIS series has data for only about 30 percent of all countries and for very few developing countries.
- 14 The multistate demographic method developed and utilized by IIASA does include education-specific mortality rates.
- 15 The current IFs education model tracks adult age-sex-education categories by five-year intervals rather than one-year intervals. In a model with a one-year time step, as IFs is, this means that one-fifth of each cohort advances annually. In an environment of increasing education participation and attainment, the process creates some degree of numerical diffusion as a portion of the educational attainment assigned to the youngest cohort advances too rapidly to the next cohort (a process sometimes called numerical diffusion). This means also that some of the stock of educational attainment ages and dies too rapidly, slightly exacerbating the underestimate.
- 16 IFs generates all such relationships for males and females separately in order to capture sex-related variations in education participation patterns vis-à-vis GDP per capita.
- 17 See McMahon (1999: 13–14) on the manner in which cross-sectional analysis helps represent patterns of long-term change.
- 18 Convergence periods in IFs range from 20 to 100 years; most are between 40 and 70 years.
- 19 We do not apply a similar constraint to primary gross intake rates, as they in fact typically “overshoot” 100 percent during a rapid education transition and then either rapidly or slowly decrease to just above or below 100 percent as adjusted net intake rates approach 100 percent.
- 20 As early as 1940, V. O. Key drew attention to the central question: “On what basis shall it be decided to allocate x dollars to activity A instead of activity B?” (1940: 1138). Later experts in the field of public finance (for example, Wildavsky [1988]) helped establish incrementalism as the dominant paradigm to explain budgeting processes and decisions. As Lindblom (1959: 81) put it, political decisions are made more through “successive limited comparisons” than through any “square one” comparison among possible alternatives. Allison (1971) contrasted three decisionmaking models, and his models of organizational process and bureaucratic politics are closer to reality for social and budgetary policy than is the rational actor model.
- 21 Even an algorithmic representation of an incremental decisionmaking process requires parameter specification. In contrast to the statistical estimation procedures used for functions such as those driving intake/transition and survival, as well as those setting targets for per student spending and total government spending on education, that algorithmic parameterization is done via analysis of the behavior of the model, a process that modelers commonly call *tuning*.
- 22 The maintenance of the IFs database is an ongoing process, and data from major sources are updated at regular intervals. In the preparation of this volume using IFs Version 6.12, our most recent download of UNESCO data followed UIS’s September 2008 update. We also used the 2008 version of the World Bank’s World Development Indicators and the 2006 Population Updates from the UN Population Division. The model was initialized with 2005 values from those sources. Data from those sources for more recent years, as available, were used in cross-sectional analyses, where our convention is to use the most recent year’s data available for each individual country.
- 23 At one extreme, in Indonesia about 60 percent of children one year below the official age enter every year, a higher entrance rate than for of-age children.
- 24 Weishuang Qu of the Millennium Institute provided information on that approach, used also in the T-21 model.
- 25 We note that McMahon chose not to follow economists’ more frequent convention of referring to “market” and “nonmarket” returns. By instead referring to “social returns” and including economic development among them, McMahon applied econometric analysis within a human development framework.
- 26 The model used gross enrollment rates as the single student flow measure in the 1999 publication. In a subsequent paper reporting the use of the model to assess social outcomes of education in Africa, primary completion rates were added as a second student flow measure (Appiah and McMahon 2002).
- 27 Like the study undertaken by Delamonica et al., this study focused on costs associated with achieving a normative target. Unlike the Delamonica study, however, Bruns et al. focused on incremental costs associated both with enrollment rate changes and with population dynamics and then compared these total incremental costs within a resource framework that, although simple in its assumptions, considered an impact from economic growth.
- 28 *Transition* in this context refers to the change from low to high rates of participation in primary education.
- 29 S-curves fit broad-scale social changes because change processes often start slowly, then build rapidly in a middle range, and slow as they approach a limit (such as 100 percent).
- 30 The Education Policy and Data Center was established in 2004 to contribute to global education policy and planning through data and analysis, and it has rapidly become very important in these roles. It is part of the Academy for Educational Development (AED) and is funded primarily by USAID and AED (Wils et al. 2007b: 6).
- 31 Most typically, these were the USAID-sponsored Demographic and Health Surveys (DHS) and the UNICEF-sponsored Multiple Indicator Cluster Survey (MICS) (Wils, Carrol, and Barrow 2005: 2).
- 32 Demographic multistate projection models reflect and project the distribution of various characteristics or “states” (such as levels of educational attainment) across a population (or subpopulations) segmented by age and by sex. The paper stated that demographic multistate projection methods were first developed at the IIASA in the 1970s (Lutz, Goujon, and Wils 2005: 9).
- 33 The definition of levels of education attainment used by Lutz et al. (2005: 16) differs from the definitions used by many other systems. Lutz et al. define “no education” as never having gone to school or completing less than one year of primary education. They place people in the category of primary education if they complete at least one year of primary school, in the category of secondary education if they ever entered secondary school, and in the category of higher education if they ever entered tertiary education after completion of secondary school. The use of these definitions produces a higher profile of education attainment than the use of completion measures would, and it needs to be taken into account when comparing their results with those of some other models and analyses, including IFs.
- 34 Although purely extrapolative, these earlier UNESCO projections were important and ambitious projects, particularly because they included all levels of formal education. One study released in 1989 provided trends and projections of enrollment by level of education and by age for the period from 1960 to 2025 (UNESCO 1989), and another in 1993 provided updated trends and projections for the same span of years (UNESCO 1993).
- 35 Porta and Wils (2007) described and compared four such detailed education system planning tools in a 2006 EPDC paper: (1) the World Bank tool associated with the Bruns, Mingat, and Rakotomalala project described earlier in this chapter; (2) the UNESCO Education Policy and Strategy Simulation Model (EPSSim); (3) the Modelo de Necesidades de Financiamiento (MNF) model used by Nicaragua and Guatemala; and (4) the EPDC Demo Ed Model.
- 36 The report noted the model could also have used trend values or user-set values for promotion and repetition rates. However, the analysis did not use trend rates because of the extreme projected values they sometimes produced (EPDC 2007b: 77). The analysis used grade-by-grade specific repetition rates when they were available and otherwise applied the average repetition rate to each grade (EPDC 2007b: 77).