A Knowledge-Based Framework for Development Policy

By: Massoud Moussavi, Noel McGinn, Kent Lewis

Our Approach to Development Policy

In evaluating policy options, say, developing new textbooks or creating more vocational schools, policymakers must deal with a series of interrelated questions: Which options are likely to be the most effective in our country? Which relate most closely to the needs and wishes of the stakeholders? What are the relative costs? How have other countries addressed similar problems? The international research community has produced a wealth of studies that can provide guidance, yet policymakers are unable to incorporate development knowledge effectively into the policymaking process. As a result, policy decisions are often based on anecdotal information and the opinions of a few experts and consultants rather than solid evidence and the corpus of available knowledge and experience. Moreover, consultation with the stakeholders themselves is often limited, and in many cases it is not clear whether the right questions are being asked. These shortcomings in the policy process are not unique to developing countries but their effects are most keenly felt there. We aim to help policymakers in developing countries overcome these process shortcomings.

Typically, policy analysis focuses on, at most, a few input variables, rather than the full array of variables affecting a given public or social system. For example, in education, much of the policy research is based on the "production function" approach, in which inputs such as school funding, physical facilities, family attributes, teacher attributes, etc. are linked to student achievement using regression analysis. While these efforts have contributed to the understanding of the factors associated with student learning, research on education production functions simply has not shown a clear, systemic relationship between resource inputs and student outcomes.[1]

Not surprisingly, analytical approaches that focus on just a few inputs often lead to single-factor policy interventions rather than more robust interventions that take advantage of the existing knowledge of the full array of factors that influence the outcome. For example, in the education sector, policymakers may

[1] Hanushek, E. A. (2008). Education production functions. In *The New Palgrave Dictionary of Economics.*

offer incentives to increase teachers' level of effort, but take no action to increase teachers' knowledge or skills. One might liken this approach to traffic planning for a city, in which only traffic volume across the bridges into and out of the city is taken into account. In such a situation, policymakers will limit their interventions to measures that either restrict or expand traffic across bridges. But our approach is to identify intermediate and contextual factors as well; in other words, we look not only at the bridges in and out of the city but also the major intersections within the city.

In the last two decades, a growing body of research has focused on understanding and developing causal as well as diagnostic models and their applications to policy analysis.[2] Our approach to development policy analysis is to build on these advances, develop theoretical models, and link them to a specific body of development knowledge. We use a Bayesian networks[3] modeling approach (described in the following sections of this note) to organize the development knowledge into expert systems that enable policymakers in developing countries to analyze and formulate policy more effectively. Bayesian networks are particularly suited for specifying causal relationships among numerous variables in a complex system to analyze their interrelationship and the performance of the entire system. This modeling technique is now used in disparate fields, such as medical diagnosis, genetics, banking, and oil production. However, it has not yet been widely used in the development context.

Using a Bayesian network model as our inference engine, we have built a software tool, called Policymakers' Workbench. Through this tool, policymakers can run unlimited iterations of various policy scenarios and connect existing knowledge to effective action. While much of the discussion in this paper is focused on education, the framework and the tool that we discuss is general and could be employed in policy analysis in other development areas, such as health, the environment, economic growth, migration, and so forth.

[2] For example, see books such as *Causalityby* Judea Pearl (2000), and *Causation, Prediction, and Search* by Spirtes, Glymore, and Scheines (2000). Harvard economists Ricardo Hausmann and Dani Rodrik have recently advocated a diagnostic approach to economic growth. Hausmann, et al. (2008), "Doing Growth Diagnostics in Practice," Center for International Development, Harvard University, September 2008.

[3] A Bayesian network is a probabilistic graphical model that represents dependencies among variables.

Policymakers' Workbench

Policymakers' Workbench© is a knowledge-based software engine based on the approach described above. It is founded upon *heuristics*, *empirical cases*, and *information from stakeholders*. Policymakers' Workbench allows decisionmakers to iteratively observe various policy scenarios, better understand and anticipate the system-wide implications of specific policy measures, and identify the areas in which it is most important to get the knowledge and views of local stakeholders.

Policymakers' Workbench has two main components: a "theoretical engine" and a computer user interface. The theoretical engine or model is built upon a set of underlying relations or functions, such as "if attendance goes up, achievement goes up, but only if there is time on task", or "if the quality of the health services increases so does their use, but only if they are affordable." This model is elaborated as a Bayesian belief network that specifies dependencies among the variables, which are based on the results of previous research, direct experience, and the testimony of experts and stakeholders in the particular domain or field. The computer interface enables users to evaluate the effect of changes in specific variables; perform diagnostics; and analyze the impact of various interventions or policies.

The organization of beliefs in networks makes it possible to update or improve the entire system of beliefs by introduction of new evidence about individual variables. This feature makes the Bayesian network an ideal method for training policymakers to learn from the evidence available to them (and to identify gaps in the evidence). In effect, a Bayesian network simulates the process of learning through reasoning and explaining.

Policymakers' Workbench is designed to help policymakers formulate policy and make decisions based on:

- Knowledge of similar cases in other countries, randomized evaluations, the extensive experience of international development institutions, etc.;
- The concerns and opinions of stakeholders; and
- Improved understanding of the implications of a policy, acquired by working through various scenarios.

The way in which Policymakers' Workbench operates can be illustrated in greater detail by describing its application to educational policy.

Applications of Policymakers' Workbench to the Education Sector

The focus of this work has been to model the teaching and learning process as it takes place in schools.[4] This requires synthesizing the results of empirical research and field experience, consistent with advances in organizational and learning theory, and presenting in a coherent fashion the accumulated knowledge and understanding of many educators and researchers, including those at work within the system being examined. Policymakers' Workbench attempts to represent aggregated evidence-supported beliefs in a systematic way that describes how all inputs into schools are interrelated and how they are transformed into learning outputs.

The Bayesian Network Model

A Bayesian (or belief) network is a directed graph that represents dependencies among variables. The full specification of such a network is as follows:

- A set of variables makes up the nodes.
- A set of links or arrows connects pairs of nodes. If there is an arrow from node X to node Y, X is said to be a *parent* of Y. The intuitive meaning of an arrow between nodes X and Y is usually that X has a direct influence on Y.
- Each node has a conditional probability distribution, based on research literature and experience, which quantifies the effect of parents on the node.
- The graph has no cycles.

Consider the following simplified example: *teacher motivation and classroom discipline influence the amount of classroom time spent on instruction, and small classes have a better level of classroom discipline.* The Bayesian network for this example is shown in Figure 1.

[4] This model is the result of collaboration of Noel McGinn and Massoud Moussavi.

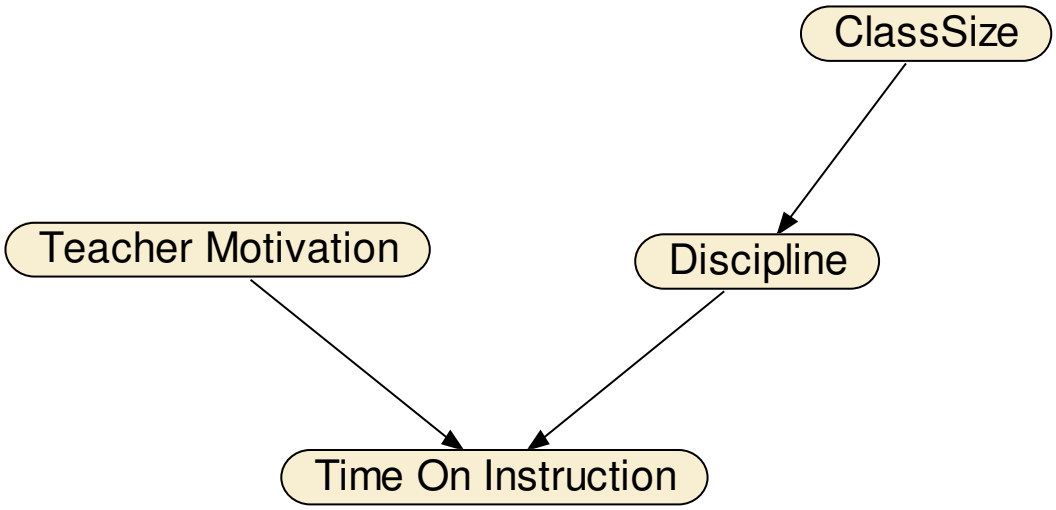

Figure 1. Simplified Example of a Belief Network

Because a belief network represents variables and their relationships of interdependence, it can be used to answer queries about those relationships using probabilistic inference. For example, given the evidence that class sizes are small, we can query the model about the probability of having good discipline in classes or the probability that a high portion of the class time is spent on instruction. Now if we additionally observe that teachers are highly motivated, this latter probability would go up. Alternatively, it would go down if the observation was that the teachers were not highly motivated. This simplified example involves just a few variables; the education model of Policymakers' Workshop has 60 some linked variables, with built-in probabilities based on the research literature.

The model predicts educational system performance and explains learning outcomes in terms of the choice and implementation of practices that affect those outcomes. At the level of individual schools the tool explains performance as a joint function of national policies and practices as well as local contextual conditions and characteristics (including teachers, students, and community). In addition, the tool can be used to diagnose non-observable aspects of school operation that require improvement. Policymakers' Workbench can make the relationship between research, on the one hand, and educational practice and policy, on the other, a productive and mutually supportive one.

The User Interface

The computer interface enables users to analyze the impact of various interventions or policies and also to examine the trade-offs among different factors, taken individually and in various combinations. Figures 2 and 3 on the following page provide sample screen shots of the system's user interface. Figure 2 shows how the user can decide to look at historical cases or perform a diagnostic. If the option "Perform Diagnostics" is chosen, the user is provided with the screen shown in Figure 3. On this screen, the user can change the value of any variable and observe its impact on all the other variables. As can be seen in the top section of Figure 3, in the educational application the input variables are classified in four general categories: **School Organization** (Principal, Textbooks, Library, etc.), **Family** (Family Socioeconomic Status, Family Involvement, etc.), **Student** (Academic History, Health, Nutrition, etc.), and **Teachers** (Academic Education, Variety in Methods, Training, etc.) The bottom section of the screen in Figure 3, displays the computed or inferred variables (e.g., Total Amount of Learning, Expected Reading Score, etc.) Here again variables are grouped in four categories.

In the context of the needs of a particular country or region, the tool may be customized to include:

- *a database* of basic social and economic information about the country, its labor market, etc.;
- a *case library* containing historical information about the reform experience in the country and other countries;
- *general domain knowledge* in the form of best practices and heuristics;
- *local knowledge* gathered from stakeholders; and
- *cost information* associated with each action.

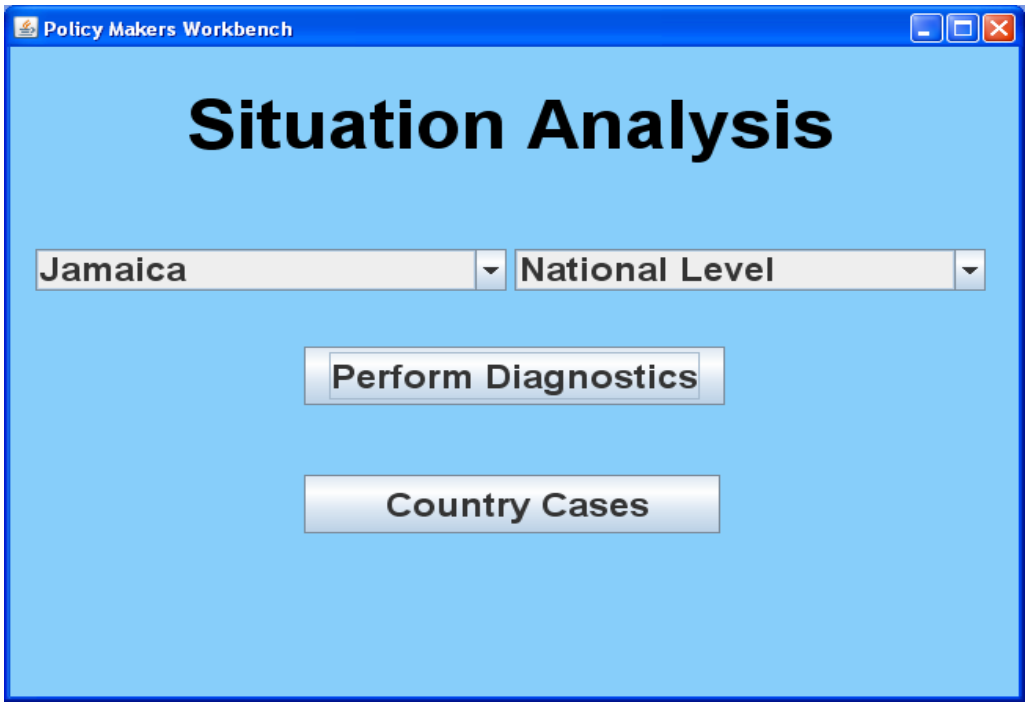

Figure 2: Situation Assessment Screen

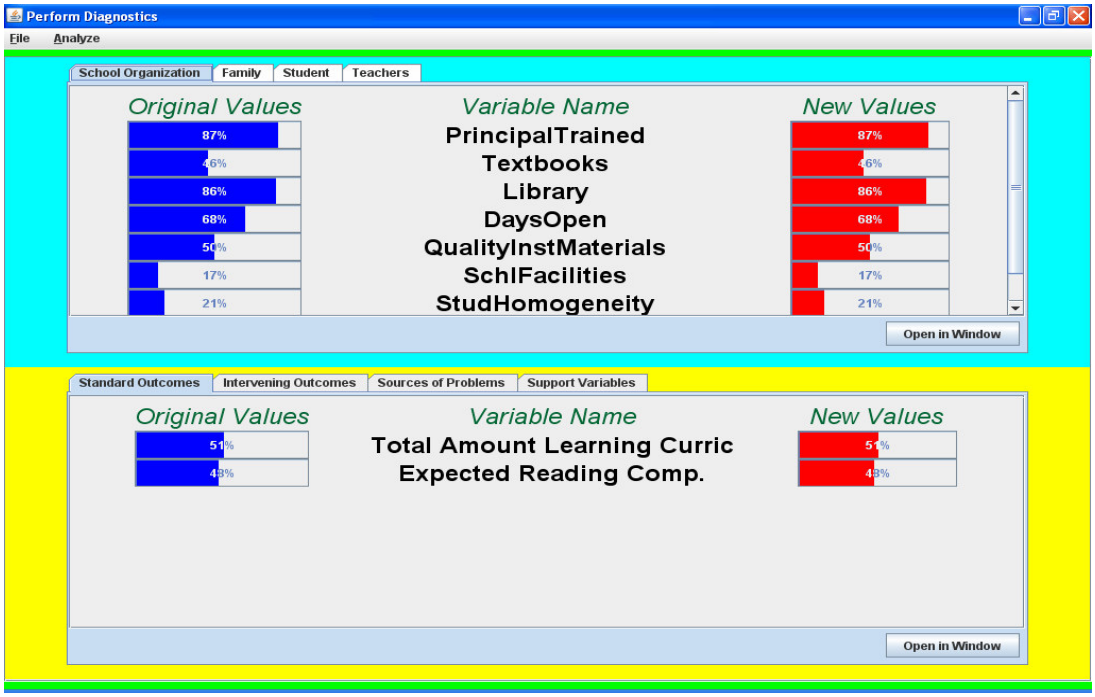

Figure 3: The input and Output variables