CHAPTER II
VISIONS AND VIEWS ON MEASURING AND MANAGING SOCIAL SECURITY

CHAPTER II
VISIONS AND VIEWS ON MEASURING AND MANAGING SOCIAL SECURITY

II.1 Defining a Vision of Programs, Agencies, and Systems

The first step in defining a proper evaluation strategy is recognizing the basic principles and goals of the programs under consideration. Contemporary societies often disagree so strongly on what should be encompassed within social security programs that consensus is rarely reached on all points. However, the persistence and strength of social security programs confirms that there is widespread agreement on the need to achieve basic social protection regarding income, health insurance, childcare and LTC, as well as other issues typically covered by these institutions. This is why the word *visions* appears in the title of this chapter: We believe there remain shared values across nations that allow us to constructively compare their social protection systems.

We also desire to highlight the word *views* to recognize the alternative ways of measuring and evaluating a program, including the economic, actuarial, administrative, fiscal, and political views, and because the way these views are applied is highly context dependent. Actuarial studies can be useful in evaluating a national collective pension system, such as the U.S. Social Security Administration (SSA) plan or the plan of the Social Security Institute of Nicaragua (*INSS*). They can also be useful to examine capitalized options, such as the Chilean Pension Fund Manager (*AFP*) system, or a mixed system, such as the Argentinean Social Security-Argentina Pension Fund Managers (*ANSES-AFJP*) program. As each actuarial study will have idiosyncratic elements, none can be considered superior to the others; each serves a purpose, and the challenge is balancing them in a manner to be useful to the heterogeneous set of actors that participate in the systems.

Because social security is a broad concept, it is helpful to distinguish among programs, agencies, and systems. All of the following discussions should be understood to refer to the social insurance environment, which includes the three concepts. At the same time, it must be recognized that applications can vary significantly, depending upon whether an agency, a program, or a system is to be evaluated. For the purposes of this study, we define a *program* as a set of rules and regulations created to manage a social risk and the relevant budgetary appropriations necessary to manage that risk. Thus, we have health insurance programs, retirement programs, and general disability programs, among many others. We define *agencies* as the administrations in charge of the programs and *systems* as collections of programs and agencies. Only by exception is there a neat nesting of programs into agencies and systems. Usually, programs and agencies are the result of historical processes tempered by the pressures of political events; private and communitarian initiatives; boom and bust periods; and reform processes to centralize or decentralize, privatize or nationalize, and merge

or separate agencies or programs, as well as other goals that can be contradictory over time or across contemporary programs. Thus, nationals of most countries can claim exceptionalism to be an attribute of their social security systems, with institutional settings being very different across nations.

On the other hand, social security programs exhibit a remarkable consistency across nations in terms of the basic goals and tools employed in reaching their diverse goals. Mulligan and Sala-i-Martin pointed out the commonalities in pension programs when they explained, "At least 166 countries have public old age pension programs. In some of the countries, public old age pensions can be dated back at least a hundred years. Although each of the programs is unique in many respects, they also tend to have many common features" (1999, 3). Similarly, Chernichovsky (1995a and 1995b) has pointed out the strong correspondence in the basic functions performed by health insurance programs across countries.

To explore these similarities, it is useful to think of programs as originating from a single stream of legislation and link them to the agencies in charge of them. An agency is often in charge of several programs. However, at the level of national systems, it is more common to have a multiplicity of programs and agencies influencing the administration of a social risk. Sometimes, a country will have a comprehensive and, hopefully, consistent strategy to define the operating model of the system, while in other countries the system will be the result of the action of a multiplicity of actors.

In principle, these structural issues are not in themselves beneficial or harmful; that is, we cannot argue that social insurance functions in a better manner if it follows a centralized plan, because it is clear that decentralized solutions often produce desirable results. Conversely, we cannot promote a view critical of centralized solutions in general when it is clear that a number of programs work well under such an approach. When addressing evaluation, we must focus upon the actual results of social insurance. In addition to avoiding prejudice, this focus is required for conducting comparisons across programs, agencies, and systems.

The Americas Social Security Report 2009 addresses the evaluation of social insurance programs by reference to pensions, healthcare, and social services, their three main components, subdividing social services into the two main components of LTC and childcare programs. Most countries' vision of social insurance follows this classification, with each class having internal complexities. As we know, a pension system is actually a combination of retirement, general disability, and workmen's compensation programs and a healthcare system is a combination of programs for active workers, retirees, children, the disabled, and other groups. LTC programs bring together elements of the pension and healthcare systems while childcare programs are financed by social security programs to offer benefits similar to those offered by the general educational system. More often than not, programs complement and overlap one another.

The following is a summary of the three subsequent chapters, each of which focuses upon the visions of the pension, healthcare, and social services systems in detail:

- *Pension systems* aim to prevent poverty in old age and provide families with a relative standard of living, similar to that which they experienced while active in the labor force, after retirement or disability.
- *Healthcare systems* aim to provide families with access to healthcare services in a holistic, integral, and equitable manner.
- *Childcare systems* aim to aid families that face obstacles to participating in the labor force due to a lack of the resources necessary to care for young children.
- *LTC systems* aim to provide permanently disabled individuals and their families with the economic and social resources necessary to

perform vital activities and maintain their standard of living.

This Report aims to find common ground across countries and agencies in the principles and goals of their social security programs by identifying approaches that can be used for alternative purposes. Given the large amount of institutional diversity, we began by defining *programs, agencies, and systems*. Now we turn our attention to identifying the means of defining the right approach for each concept.

One consideration is that the evaluation of the adequacy of the pension levels provided by a national pension system should be largely independent of the financial model followed in each country, and is certainly one of the main evaluation indicators that can be identified for a pension system. However, agencies and programs within a pension system serve different populations, and although their goals should certainly relate to the overall system, these goals do not have to be the same. Adding to the complexity is the fact that an evaluation agency may have policy, administrative, control, and other goals. An agency may be most concerned with evaluation of the management of a program while the Congress or Ministry of Finance may be most concerned with the evaluation of a policy. Clearly, it is not possible to define a single model to fulfill all purposes. The approach followed in the previous CISS Report (CISS 2007) was particularly beneficial because it allowed for the division of the components of social security and the rebuilding of applications for specific cases.

Our current social protection programs are usually the result of long historical processes that have left national legacies. Thus, each system (*e.g.*, healthcare or pension) typically mixes a variety of programs and agencies, likely sharing visions regarding basic goals, such as providing retirement income, but conforming to varied organizational models that cannot be easily compared on significant dimensions, such as the cost of provision, the meaning of coverage, or the risks of each model to the population or the national budget. Therefore, it is useful to identify which components agencies and programs share before measuring the performance of different agencies and programs in a systematic manner.

II.2 Approaches: Social and Economic, Operations Research, Administrative, Fiscal, and Actuarial

It is typical and highly convenient for the agents involved in evaluation, including actuaries, administrators, auditors, economists, and operations researchers, to maintain alternative approaches to evaluation. It is therefore useful to draw a map of the alternative approaches to find the links between them and develop useful bridges that will allow meaningful comparisons. The use of each approach varies across programs according to its nature. For example, the actuarial approach is more often used to evaluate pension programs than healthcare programs, and applications from OR approaches are increasingly found to be important in the provision of healthcare services and social insurance.

Similarly, evaluation must consider the needs of different social agents. For example, administrations are typically very interested in an approach that provides them with elements with which to make decisions on day-to-day operations so they can focus upon achieving the goals mandated by their statutes. On the other hand, a ministry developing policies or an academic institution may be interested in an evaluation of a program and agency that considers its effects upon society, the labor market, the national budget, and other variables that may not be directly of interest to the operating agency.

Each of the following specialized chapters maps the existing knowledge with regards to the different approaches, pointing out to the ways in which they can be applied in alternative institutional contexts. Table II.1 summarizes the key concepts applied in different approaches to evaluation.

Table II.1
Key Concepts in Evaluation Approaches

Economic	Key concept: causality. Economists use behavioral models to identify how the environment or an intervention affects a variable. For example, an economic model may evaluate whether an increase in old-age economic benefits reduces the average age of retirement or whether a change in a formula to finance healthcare increases utilization of services.
Actuarial	Key concept: actuarial balance. Actuaries use numerical models to restrict the evolution of a system over time, given the rules and assumptions regarding behavior. Thus, predictions on demographics and financials are obtained and systems are evaluated in terms of their financial balance over time.
Administrative	Key concepts: customer satisfaction, efficient operations, consistent financials, and effective personnel management. In comparison with the other approaches listed in this table, administrative approaches do not strongly relate to a specific academic discipline. Administrative approaches are highly influenced by the training of administrators and the products offered by vendors for evaluation (*e.g.*, the software solutions).
Fiscal	Key concept: solvency. Governments are interested in the feasibility of social security programs, and as such combine actuarial and other statistical models to evaluate the impact of programs on public expenditures and deficits.
Operations research	Key concept: optimization of a complex system. OR engineers utilize statistics, optimization, stochastics, queuing theory, game theory, graph theory, decision analysis, and simulation to measure the efficiency of a system and the sources of inefficiency and recommend ways to find optimal solutions.

It must be made clear that there is not one preferred view. Depending upon the application, user, program, or agency under evaluation, each approach can add value to an assessment exercise. This is pointed out in Table II.2, which lists the weaknesses and strengths of each approach, which are then illustrated with examples in the following subsections.

II.2.1 Economic Approach

The economic approach presumes to follow the scientific method. Because economic models propose cause-and-effect relations, they naturally view programs as "interventions" that affect the behavior of individuals, and measure how and how much the programs affect such behavior. They may also simulate how program operations themselves are fed back into the program design.

A prototypical economic evaluation of social security relates the behavior of workers to the incentives offered by a pension program. In this case, economists model the decision of workers to supply hours of work, considering market wage schedules and worker saving and leisure options. A pension program affects this decision, usually through making it more costly for the individual to keep working once he or she is entitled to a pension benefit. Similarly, a mandatory pension program reduces the incentive to save private income. An economic evaluation typically attempts to measure the reduction in hours worked and savings as a consequence of receiving social security benefits. Several recent references to this application are contained in *Social Security Programs and Retirement around the World: Fiscal Implications of Reform,* an excellent book edited by Gruber and Wise (2007).

Many economic models that evaluate social programs are based upon the human capital theory, which argues that time has a value to individuals and that variations in its value influence their decisions. Once they had adopted this theory, many economic approaches began to be applied to almost any decision related to work, education, or participation in social programs. In childcare programs, economists study the value that mothers assign to their time working in relation to the value

that they assign to time spent caring for their children. This makes it possible to study the impact of a cash subsidy or an in-kind benefit on a mother's decision to use childcare services while working and on her decision to work. Similarly, economics models are used to study a family's decision to participate in the formal economy and obtain health insurance against the option of remaining in a job without such protection and paying for healthcare services out of pocket.

Table II.2
Strengths and Weaknesses of the Major Evaluation Approaches

	Strengths	Weaknesses
Economic	Potential for analysis of behavioral responses.	No existing theories of behavior for important cases.
	Use of more sophisticated statistical (econometric) techniques.	Too much "faith" in theory; higher risk of political bias.
	Ability to use the very large databases available today.	Requires information very often unavailable.
		Lack of standardization; comparability of studies usually only possible after long research periods.
Actuarial	Hard use of numerical models.	Difficulties in modeling programs in which individual family members change behavior as a consequence of the program.
	Consistent use of sampling theories and methods.	Risk of assigning permanent status to predictions based upon time-specific assumptions.
	Regulations on the profession and report to ensure consistency and comparability across time and organizations.	Social security agencies and other social programs may have little control over access, rates, and benefits in the midterm, and behavioral responses to programs can be very large and fundamentally alter assumptions.
	More reliable to the extent that administration has control over rules of access to programs, premiums, and benefits.	
Fiscal	Capacity to provide government with information on feasibility of public programs.	Ignores considerations of individual welfare to evaluate aggregate results.
	Ability to present programs in a form amenable to understanding by legislators.	Emphasis on cash flows; accrual accounting viewed as secondary.
Operations research	Careful modeling of administrative process.	Requires quality and flow of data not supplied by many organizations.
	Ability to evaluate operations' costs by process.	Assumes adequate definition of process in the organization.
	Very useful for established high volume, repetitive processes.	
Administrative	Ability to obtain balanced views.	Risk of bias due to administrative "conveniences".
	Provides useful tools to make day-to-day decisions.	Tunnel vision; lack of criticism for deviation of social goals.
	Provides tools designed to evaluate and manage.	Risk of falling into disorganized state due to excessive volume of information and criteria and the presence of many contradictory indicators.
	Ability to communicate with personnel in charge of the agency or program.	

The human capital approach is not the only model upon which the economic approach is based. Models on the behavior of organizations are also important for social security programs. Important examples are the studies that link the financing of social security to healthcare services and examine the operation of these services. For example, social security sometimes finances healthcare services through reimbursement or "line-item" budgets. This means that service providers make decisions on inputs and their costs, and then pass the bill onto a financial fund that pays for them.

On the other hand, some health insurance programs have transitioned to some form of "prospective payment" policy in which the financial fund pays providers according to an "expected cost" rule that transfers the risk of deviations in cost to service providers. Examples of prospective payment policies are those based upon *capitated budgets* and *diagnostic payments.* Capitated budgets base a healthcare service provider's budget upon the number of persons covered and their actuarial risk. The provider is responsible for providing services and must absorb the risk of cost overruns. Similarly, diagnostic payments compensate providers on the basis of the average cost of a diagnostic procedure, with cost savings or overruns absorbed by the provider. Many social security health insurance funds have moved towards this type of policy to promote cost control and saving among providers. Thus, an economic model may try to evaluate the impact of the change in policy on the behavior of providers. For example, it has been observed that diagnostic payments reduce the length of hospital stays and hospital mortality, while capitated budgets give providers incentives to reduce the number of high-cost physician consultations and restrict the access of individuals to costly hospital beds. In one study, Cutler (1995) found that after the U.S. Medicare program adopted a prospective payment program, hospitals observed changes in the timing of deaths and rate of readmission. This may be attributed to the fact that when hospitals see changes in the financial compensation for a specific patient diagnosis, they may decide to modify the way in which they classify patients, particularly patients with dual diagnoses. For example, a complex case may be "divided" into two diagnostic categories, generating an additional payment from the social insurance fund and a readmission event in the statistics. This sort of behavioral response to programs may be modeled with an economic approach.

While economic models often promise that they are based upon a robust understanding of the behavior of individuals and organizations, they face difficult challenges, primarily due to a lack of theories to explain important phenomena and the paucity of good data to apply existing theories. As in any field in academia, economists often study that which they best understand to solve the problems that they can. Although it would be ideal if all policy and administrative decisions were based on strong and reliable research, the real world cannot always wait for research solutions, and must often resort to the untried and unproven.

It should be noted that as large international financial institutions are dominated by economists, their evaluations of programs are generally biased towards economic approaches.

II.2.2 Actuarial Approach

The actuarial approach has a strong tradition in social security, in particular in the pension field. Most national pension programs were founded following actuarial studies that calculated the taxes and benefits that could be supported and, based upon these calculations, the investment policies that should be followed. Actuarial science has been experiencing a period of increased demand for its services as a consequence of the development of IT that has made it possible to apply its methods to a widening array of applications. The actuarial profession is subject to regulations that standardize its practice and allow comparisons of studies across time and organizations. Actuarial studies are often an integral

component of the financial information of public and private corporations. The use of well-established numerical methods and the intensive use of statistical samples allow actuaries to model the regular behavior of systems.

For social security, the prototypical application is the evaluation of a pension plan. An actuarial study uses demographic and biometric micro-data and tables that provide information on fertility and mortality; contributions and benefits; and assumptions regarding the expected behavior of external variables, such as interest and inflation rates. Its products are income, expense, and deficit projections over time measured in terms of cash flows and other accounting metrics, such as the liabilities of the pension plan at different points in time.

To the extent that the assumptions used in an actuarial model are accurate, its projections tend to be correct. For a private pension plan in which the insurer can control the admission of individuals, prices charged, administrative costs, and other variables, the studies can be quite accurate. Similarly, an actuarial study for a national pension plan in a country with stable demographic variables, low and predictable inflation, financial stability, and nearly universal coverage will yield fairly accurate data. On the other hand, an actuarial study of a country experiencing large demographic transformations, an unstable relationship between wages and inflation, and a large informal economy that can sap the growth of social security programs is likely to yield large projection errors.

In the health insurance area, actuarial studies have been increasing in size and complexity as health expenditures have reached record levels. These studies can make use of the very detailed information on diagnostics and unit costs provided by contemporary information systems used by hospitals and healthcare funds. The main challenges faced by the actuarial approach have been that the costs of treatment can change greatly after the introduction of new drugs and treatments and that the behavior of individuals and healthcare service providers can change substantially over relatively short periods.

Is there a guideline for when to use actuarial studies in policy analysis? In a sense, the answer is simply that we can always do so, and therefore the large databases of contemporary social security administrations and healthcare providers should be accessed intensively. However, whenever we expect important changes in behavior due to new conditions, such as new rules for participation in programs, new means of financing providers, significant economic instability, or other events that can alter behavior (*e.g.*, very large increases in the price of drugs), we must be careful in interpreting the results of these studies.

II.2.3 Fiscal Approach

The fiscal approach is relatively simple: It measures the cash flow produced by programs financed by the government budget and the public debt. Because it makes use of both the actuarial and economic models, it can be seen as a specialized application of these two approaches.

It is useful to review in detail the way in which countries evaluate the fiscal impacts of social security. Although not the only criteria used by governments to make decisions, they are very important. The main producers of information on the fiscal approach are the Organisation for Economic Co-operation and Development (OECD) and international financial organizations. These organizations have produced reports that indicate increasing concern about the impact of social security on public finances. When we review the main existing reports in the following chapters, it will be seen that most of their interest concerns the role of pension plans, but there is a growing awareness that health insurance may become an even bigger concern for national finance.

II.2.4 Operations Research Approach

Operations research (OR), a subfield of the engineering field, models administrative processes in detail and measures their costs to redesign them in an optimal manner. To the extent that such processes can be adequately modeled and the financial accounting of agencies allows the measurement of costs, OR can provide interesting insights into improving the management of agencies.

The main OR applications of interest for social security are in the service provision arena. It is helpful to consider a situation to illustrate the applications. Consider two hospitals: Hospital A is located in an urban region populated by low-income families while Hospital B in an urban region populated by middle-income families. One has more physicians, but the other has more beds. One has a higher budget per user than does the other. How do we measure and compare their efficiency? Do we measure the number of surgeries per bed? Do we measure the number of surgeries per physician? Do we measure the number of consultations per person covered? One answer is measuring all of these indicators to provide insights into the operations of the hospitals, but it may be very difficult or even impossible to reach conclusions useful for decision making after doing so. Should additional budget resources be provided to one of these hospitals? Is one of them wasting resources? Do they have too many physicians? Box II.1 elaborates upon this dilemma.

An OR approach offers systematic solutions to these questions. By modeling the entire process of service provision, an OR approach allows comparisons across different hospitals regarding their efficiency in the use of resources. It also allows the measurement of inefficiency and the linking of the measure of inefficiency with the measure of resource inputs. An OR approach may yield statements such as the following: "Hospital A is 10% more efficient that Hospital B, and two thirds of that difference can be attributed to Hospital B's excess use of drugs and one third to its inefficient use of physician time." Statements such as this can be very useful to an administration but may be difficult to produce in the absence of a well-defined approach to measuring operations.

The OR approach requires detailed information on the processes that produce the services, the cost of inputs, and the measurement of outputs. In the hospital example, we might be forced to define outputs as the number of consultations and the number of surgeries because these are the only data provided by existing databases. Even though we might prefer to define output as the state of the health of the population, we may not be able to measure that variable or link it to data on inputs (*e.g.*, drugs, beds, and physician hours).

However, we should not dismiss using an OR approach because of the difficult example we have just posed. Social security administrations can achieve success by developing an OR strategy. For example, an agency can identify in a very detailed and accurate manner the process followed to pay pensions. Then it can evaluate the efficiency of each of its agencies or intermediaries (such as banks) that pay the pensions, which, because they could number in the thousands in a large country, could be very difficult to compare otherwise. However, even in the more difficult cases, OR methods can be helpful in organizing and understanding information.

II.3 Administrative Approaches

Unlike the other approaches, which have a body of knowledge and a community of practitioners that follow, define, evaluate them, there is no prototypical administrative approach. While there are certainly administrative schools, there is no one theory of administration based upon a defining hypothesis and management methods.

The goal of an administration is defined by the vision upon which it defines strategies, procedures, incentives, payment schemes, and other factors necessary to achieve its goals. Evaluation in management is naturally based upon an organization's goal. For example, a very competitive firm in the

Box II.1
Why the Usual Means of Measuring May Be Misleading

Perhaps the most common form of evaluation used by many social insurance programs is basing evaluation upon a list of indicators. For the OR approach, the use of quotient indicators poses a very basic problem. The term *quotient indicators* refers to the way that the indicators are obtained, which is by calculating the ratio of two variables, such as surgeries per surgery room, time required to process a disability claim, and waiting time with relation to the number of services provided.

The number of quotient indicators can be overwhelmingly large, which poses a great challenge. Consider two hospitals that provide the same interventions and are both evaluated by a list of possibly hundreds of indicators, as hospitals are in most countries. However, one is a large hospital that serves a large city and the other a small hospital that serves a small city. It would not be surprising that one would report consistently better results for some indicators and consistently worse results for others. For example, hospitals in large cities often have more crowded emergency rooms and cost overruns from attending more complex diseases; on the other hand, they have lower unit costs and higher occupancy rates. Against these systematic contradictions, we need a methodology to weigh the results according to the environment of each hospital.

But there is more. Consider a third hospital that, according to indicators, performs the best in some areas and the worst in others. It can be shown that even if this hospital is the least efficient of the three, it will appear to perform better than one of the others when a strategy of quotient indicators is used.

Fortunately, there are techniques to deal with these issues. A robust one advocated by engineers and economists has been termed *data envelopment analysis* (DEA). "Measuring the Efficiency of Decision-Making Units" by Charnes, Cooper, and Rhodes (1978) is considered seminal and *An Introduction to Data Envelopment Analysis: A Tool for Performance Measurement* by Ramanthan (2003) is a useful reference for practitioners. Applying DEA makes it is possible to achieve consistent evaluation of operating units (*e.g.*, hospitals, clinics, and customer service offices), measurement of the degree of inefficiency, and the linking of the general measure of inefficiency with the use of particular inputs.

As with any evaluation strategy, managers must understand application possibilities and limitations and be able to interact with the experts. As with other contemporary techniques, DEA requires a commitment by the administration to invest in database development and the use of the tools. However, any manager of a large agency that has been faced with an unending meeting regarding the evaluation of indicators that appear contradictory would appreciate the possibility of developing consistent ways to measure results.

private sector may decide that its only evaluation criteria are profits. The reality is usually more complex, with even for-profit firms requiring evaluation of their operations and targets.

For non-profit organizations, as are most social security agencies, administrative evaluation can never be based upon profits. In addition, the agency's vision is usually mandated by law, and even its administrative procedures are often defined in legislation or regulations promulgated by an authority outside the agency. Because most agencies must operate a program defined by law, they often do not want to concern themselves with criteria such as efficiency, customer service, reliability, and accountability. These agencies can be public or private, non-profit or for-profit, or national or regional, and may integrate financial and service provision features or specialize in and/or manage a single program or several.

For most countries, the final evaluation of social security agencies is not performed by the agency itself—often not even by the government—but by elected executive officers (presidents and prime ministers) and legislative bodies that make final decisions on the value of the programs and the ways they should be reformed. Nevertheless, just as pure

profit is rarely the single measure for evaluating a private company, a social security agency cannot wait every for political decision to be finalized before performing its day-to-day operations. Thus, while administrative approaches must be defined according to the vision of the agency in charge, they are also strongly defined by administrative practices and organization, which in turn are dependent upon the abilities and knowledge of managers as well as the technologies available in the market.

How can firms and agencies evaluate operations and results with the goal of better management? Is there a catalog of options, a menu-driven set of choices to select the best alternatives? The short answer to the latter question is no. Rather, the best way to manage is usually the result of a unique mix of managerial abilities, human capital, and the legacy left in place by past administrations regarding information systems and fixed assets. These and other factors that make each organization unique, such as its size and even its luck at a given point in time, provide it with the capacity to finance adjustments to the administration and manage the pressure of performing day-to-day operations.

The fact that there is a diversity of approaches that can be followed does not contradict the fact that there are management principles and techniques that have become generally accepted and tend to be used with more frequency than are others. Success can be achieved by agencies willing to discipline their administrations through an adequate mix of these techniques. Perhaps the principle that best summarizes these approaches is one that emphasizes the need to measure activities and results for successful management. "If you cannot measure, you cannot manage" is the relevant *cliché*. However, the act of measuring demands a definition of the objects to be measured, the metrics used to measure them, and a strategy for resolving problems.

One of the pioneers of administration evaluation was Edwards Deming, who provided the following insight on evaluation: "Measures of productivity are like statistics on accidents: they tell you all about the number of accidents in the home, on the road, and at the work place, but they do not tell you how to reduce the frequency of accidents" (1982, 15). Over the years, corporations and management schools have improved the methodologies that they use to define what to measure and how to measure it. In the following subsections, we discuss three approaches that have had a significant impact on the administration of social security agencies: the quality movement, balanced scorecard, and six-sigma approach. These terms do not refer to theories or hypotheses, and are not always viewed or only viewed as evaluation techniques. Rather, when an organization employs one of these techniques to define its operating, financial, and human resources processes, evaluation is a natural outcome.

Ideally, an organization will perform evaluations in a form fully integrated within its operating actions. For example, the data used for information will ideally be the same as those used for making day-to-day decisions. New technologies are making these approaches increasingly valuable. In the past, information for executive decision-making came in a statistical format with significant lags with respect to actual operations. Many modern organizations have been able to "eliminate the seams" so that all levels of management can access the same data online and on-demand.

II.3.1 Quality Movement Approach

Perhaps the most pervasive management approach that has been adopted by social security agencies is based upon the concept of *quality*, defined as the successful adoption of standards and compliance with them. While the term *quality* is quite general and can be employed within a number contexts, in this section we focus upon a specific approach based upon standardization and the measurement of compliance with a standard.

The quality movement (QM) approach began to be applied successfully in manufacturing. Measuring

quality in relation to the quantity of products produced with defects is a relatively straightforward concept. However, producers soon realized that to increase the quality of their final goods, it was necessary to imbue the whole production process with a quality approach that decreased errors in such areas as logistics, production lines, financial transactions, and responses to customer questions. Standards to promote quality control have been in use since the dawn of modern manufacturing, with the direct precedents of the modern approach originating from the arms production industry during World War II. During the 1980s, the QM approach began to gain the momentum that resulted in the prominence it now enjoys when it merged with management approaches that emphasize the measurement of operations and emphasize results as fundamental for the success of an organization.

From manufacturing, the QM approach began to expand into other fields, and today has specific applications in the service, government, healthcare, and finance industries. Among the more relevant applications for social security are those related to pension funds, including the investment of funds, operations of the pension system (*e.g.*, payment of benefits), determination of disability status, and registration of employers and workers; those related to healthcare at both the level of provision and risk management; and those related to safety and health at both the level of a fund that manages risks and the level of an employer. According to the Quality Management Institute (QMI), a provider of registration services for quality standards (www.qmi.com), a QM approach provides the following benefits: increased streamlining of processes, decreased scrap and reworked material, increased productivity, decreased costs, expanded production capacity, decreased cycle time, higher part-quality standards, extended process capability, increased process flexibility, greater preparation for new product production, and enhanced intercompany relationships.

The QM approach often employs standards developed by the International Organization for Standardization (ISO), a non-governmental organization within the framework of the United Nations, whose members are the national standards authorities. The ISO addresses many issues in addition to QM because standards have other important applications, such as safety and information. In addition, while the ISO is a global actor, it has no enforcement authority, and many industries have their own specialized standardization bodies. Regarding QM, the ISO issues the standards; the national accreditation boards regulate the certification bodies, which are composed of consultants certified by a national board; and the certification bodies certify the companies, government agencies, and other organizations willing to adopt ISO standards.

Who accredits? Who certifies? In large countries, where a significant demand for the QM approach has developed, a national body is often responsible for the accreditation of management systems. In the Americas, these bodies are the following: the ANSI-ASQ National Accreditation Board in the United States, the Standards Council of Canada (SCC) in Canada, the *Entidad Mexicana de Acreditación A.C.* (*EMA*) in Mexico, the International Automotive Oversight Bureau (IAOB), the *Organismo Argentino de Acreditación* (*OAA*) in Argentina, the Brazilian General Coordination for Accreditation (*CGCRE*) of the National Institute of Metrology, Standardization, and Industrial Quality (*INMETRO*) in Brazil, and the *Instituto Nacional de Normalización* (*INN*) in Chile. These national boards provide accreditation to certification bodies, which in turn support organizations interested in adopting and registering standards.

At the global level, the International Accreditation Forum (IAF) and IAF multilateral cooperative arrangements (MLAs) coordinate quality management system (QMS) and environmental management system (EMS) accreditation activities. While there can be several relevant ISO standards for social security agencies, depending upon the scope of their operations, the most commonly used is ISO-9001, which deals with QMS. Typically, an agency will

complete the following process to become ISO certified: 1) the agency makes a proposal to a certification body; 2) the certification body accepts the application and assigns an auditor; 3) the documentation is reviewed; 4) depending upon the standard, the agency passes through a pre-assessment stage; 5) a formal audit is performed; 6) the agency alters it polices or practices to conform to any necessary standards with which it does not yet conform; 7) the certification body registers and issues a certificate so that the organization can declare itself ISO certified; 8) the certification body performs surveillance audits; and 9) the certification body performs recertification audits.

The QM approach is widely recognized for its ability to provide a comprehensive and effective means of managing an agency's quality, directing an agency along a process of improvement, and helping an agency communicate to its clients and suppliers what to expect from the agency. For example, a certification body can help a social services agency improve the management of the pension process of benefit payment and can help a healthcare agency guarantee its beneficiaries that the procedures for refunding healthcare expenses or admitting patients for treatment are being applied equitably and consistently.

There are two main criticisms to the use of ISO standards: the cost of their implementation and the possibility of their misuse. The cost of implementation is not truly a deficiency in the QM approach. Certainly, there are costs of training and obtaining certification, but each organization must make decisions on the basis of clear criteria: Does it expect the decrease in operating costs and increase in the quality of services to compensate for the cost of adopting and certifying the standard? Criticism of the misuse of the process may be valid in a few cases, but the extent of misuse is unclear. The primary criticism is that an organization may obtain certification mainly for marketing purposes with no intention of making an effort to improve; that is, a flawed procedure may be certified and no attempt made to rectify the procedure. While it is likely that a few organizations may do this, this form of abuse may not be important in the mid-term. The reason is that the certification process involves costs that will not be justified if no improvements are achieved, and the external agents (clients and suppliers) will eventually realize that the organization is not providing any real benefit from the certification process. On the other hand, there is the legitimate consideration whether this type of abuse is more common among monopolies and public agencies that do not face the discipline of competition.

II.3.2 Six-Sigma Approach

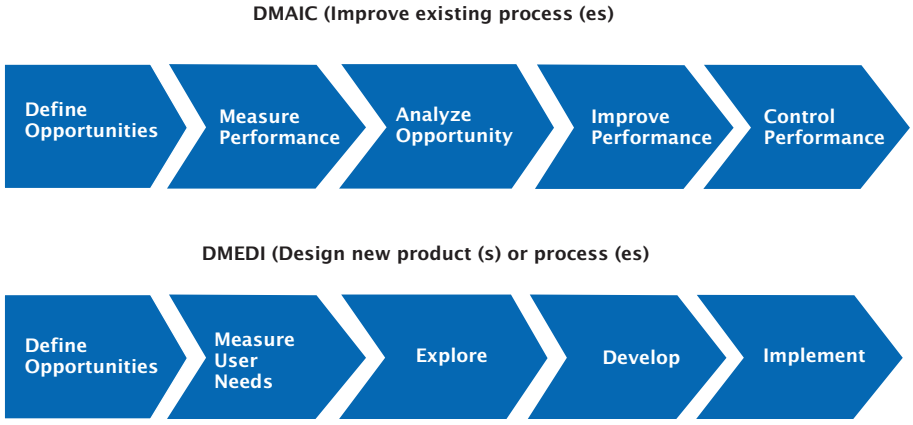
The six-sigma approach, which has been considered very successful in helping organizations achieve very high quality in their operations and supply of goods and services, allows the definition of administrative processes and operationalization of their measurement in a rich way, integrating statistics, management and strategy. Its application requires more than simply the commitment of management; it demands that the entire administration be redesigned to follow its management principles.

The term applied to the technique contains the Greek letter *sigma* (σ), which is employed by statisticians to denote a measure of deviation of a variable from its mean or expected value. Thus, an administration capable of adequately defining a process or result can measure it and calculate how far it deviates from an expected value through a measure denoted by the symbol. The *six* in six-sigma means that very low levels of error are allowed to measure errors; other numbers can be allowed, but the one used in the original applications denotes a high commitment to quality. For example, an agency in charge of a pension plan may state that it must pay pensions on the first day of the month, and will measure the average duration of its delay. Thus, errors will be measured statistically, and reduced to a certain level.

The six-sigma approach follows a strategy that has become well known within the business world. The six-sigma approach uses data and performs

statistical analyses to measure and improve a company's operational performance, typically by describing applications for product development and improving existing processes. Figures II.1 and II.2 summarize the two approaches generally used.

Figure II.1
DMAIC and DMEDI Six-Sigma Models



Source: Adapted from Islam 2006, 24.

Figure II.2
DMAIC Define Phase

Objective	Activities	Tools	Deliverables
Understand the problem	Develop a business case	Stakeholder analysis form	Project charter
Develop a problem charter	Develop opportunity statement	Charter template	Goal statement
Identify critical-to-quality issues (CTQs)	Develop goal statement	Affinity diagram	Scope definition
Develop a plan to manage the project	Identify project scope	Process maps	Process maps
	Develop project plan		Team
	Process mapping		Critical to quality
			Define tollgate

Source: Adapted from Islam 2006, 25.

DMAIC is an acronym used to encapsulate the five phases of the six-sigma improvement methodology—define, measure, analyze, improve, and control—as part of the process for product or process improvement. *DMEDI* is an acronym used to describe the five phases of product development—define, measure, explore, develop, and implement. The DMEDI process was designed to ensure that the desired business and financial results are achieved. While the limited scope of this report does not allow us to explain the six-sigma approach in detail (refer to Islam 2006 for a thorough introduction to the approach), we will demonstrate that it has been one of the more successful strategies operationalized within several industries in recent times.

As discussed in the section above on the QM approach, it is clear that the standardization of operations can be useful in measuring errors, and thus it can be useful to apply a six-sigma strategy. However, it should also be clear that the use of ISO standards is in no way necessary for applying this approach; what is necessary to apply a six-sigma strategy is a commitment by the entire organization to its application because it requires substantial levels of training and consistency in its application to be effective.

II.3.3 Balanced Scorecard Approach

The balanced scorecard (BSC) approach is a methodology that has become popular for organizing the large flows of information available to contemporary organizations. As with the more theoretical approaches (economic, actuarial, and OR), the explosion in database sizes and computing capacity has made it feasible to adopt models of a sophistication that had been unattainable just a few years ago. It is expected that the improvement in IT will be very large and that the capacity to develop these models will increase at a high rate for at least several decades.

Before explaining the BSC approach, we would like to point to a few examples that illustrate the volume of information that may be available to a social security organization:

- A social security agency can access online data pertaining to applications for disability benefits to determine who is asking for benefits and for what reasons at the exact moment that the applications are presented.
- A health insurance fund can determine the level of occupancy of each type of hospital bed among the thousands it finances. It may also be able to determine which patients are in the beds and their diagnoses.
- A childcare system can provide parents with immediate access to information on childcare centers that have spaces available, either by accessing the Internet or calling a toll-free telephone number. Parents may also be able to access information on their eligibility for subsidies, the services provided by a center, and the status of their child at a given point in time.

As these limited examples demonstrate, it is very challenging for any manager in charge of receiving all this information to organize it in a meaningful way. The BSC approach, one attempt to do so, is applied by most large software packages to allow all employees in an organization to participate in strategic management and access updated data on everyday activities so that they can contribute to organizational success.

One advantage of the BSC approach is that it allows performance reviews to be performed more easily and consistently across all levels and departments in an organization. The BSC strategy should not be seen solely as a way to manage and organize data but also, because of its adoption of a "causality view," a way of organizing information in ways that signal why certain operations are successful or unsuccessful. The BSC approach borrows from psychological approaches to achieve the best possible interface between statistical models and human understanding. Thus, one of its aims is to achieve a structure of information that allows all

individuals in the organization to easily access and evaluate the information relevant to their particular situation. The BSC approach is sometimes illustrated through a graphic model placed within four "walls," each one dedicated to a strategic feature: human resources, financials, operations, and customer service. This description changes across applications, depending upon the user organization. For example, the walls of an alternative model may be customer service, financials, internal processes, and learning and growth.

From the preceding description, it can be inferred that a BSC application may make intensive use of all the other techniques mentioned in this chapter for evaluation and monitoring. A good BSC application contains in-built econometric and actuarial models to identify cause-and-effect relationships, is able to measure whether and how a quality strategy is working, and employs operations researchers in an intensive manner. Box II.2 illustrates data mining, a technique highly complementary to the BSC approach that has been increasing in popularity.

II.3.4 Financial Approach

In practice, there appear to be two "syndromes" that affect the evaluation of social security programs. One syndrome arises from the fact that evaluation is biased towards financials because of the political weight given to fiscal control. The other syndrome arises from a focus upon simple and partial information due to a lack of appropriate financial information; some public programs and agencies do not develop financial statements properly, making it very difficult to measure costs and apply basic management techniques. Even though financials are only tools to help provide final services, they are necessary for good management.

In this section, we do not explain the manner in which financial evaluations are performed by social security agencies but rather point out why and how the two syndromes must be addressed. A program evaluation approach based predominantly upon financials must evolve into a more balanced approach by the incorporation of other factors. This dependence upon financials appears more common among pension programs and agencies established as state-owned corporations. Part of the reason for this syndrome can be attributed to the increasing focus upon internal control approaches. Across Latin America and the Caribbean (LAC), transparency and budget control legislation has created large regulatory bodies that focus upon mainly financial issues in their surveillance of social security programs and of public programs. For example, the evaluation of procurement processes basically follows compliance processes that pertain to laws on purchasing. Certainly, compliance with a law on purchasing by a public agency should always be very high in the agenda, but it should not obscure that the goal of the agency is to provide health services.

The other syndrome, a lack of financial information, is somewhat common in public programs that have operated in a centralized fashion (that is, as part of a government department) and thus have not developed proper financial accounting procedures. This is the case in some public health systems and in childcare and other programs financed directly by central budgets and managed centrally. As information, financials are basically signaling devices that help agencies organize their operations, human resources, and services. Thus, in the absence of basic developments, programs and agencies find themselves unable to develop similarly basic evaluation and monitoring strategies.

II.3.5 Legal and Technological Challenges

Evaluation strategies are bound to face challenges arising from privacy issues and other legal constraints, as well as the rapid development of IT, software capabilities, and applications for evaluation, which will only become more complex in the years to come. While transparency is valued in a democratic society, it is clear that not all information can or should become public. For example, citizens have interests in keeping private personal information regarding medical treatments, marital status, or pension income.

Box II.2
Data Mining

Managers are often faced with an array of administrative fads with peculiar and sometimes mysterious or motivating names. However, it is not always easy for the non-expert to judge when a claim of innovation is legitimate. *Data mining* is a phrase that has been gaining popularity, and we explain in this box how it fits into a general evaluation strategy for a social security administration.

Data mining is defined as the process of exploration and analysis by automatic and semiautomatic means of large amounts of data with the purpose of discovering rules and patterns that have meaning. Data mining has certain synonyms, including learning by machine, statistical learning, knowledge discovery, and artificial intelligence.

Data mining is usually subdivided into two large sets of applications. In supervised learning, the goal is to predict the value of a result on the basis of the number of metrics of inputs. In non-supervised learning, there is no measurement of the product because the goal is to describe association among patterns in measures of inputs.

Some prototypical examples of data mining include predicting the order of answers in a database of transactions, predicting whether a customer will default based upon creditor consumer data, detecting fraud when invoicing registries, predicting activity from a database of transactions, detecting attacks to a network from traffic data, and identifying spam from patterns of words in headers and texts of e-mails.

The more practical approach used for most businesses applications is supervised learning. When using it, it is indispensable to consider the very large databases that allow the use of the typical paradigm of machine learning. Usually, there is a need for large samples for learning and one or more additional samples for validation.

It is not difficult to identify useful applications for social security agencies. For example, an agency may be concerned about how to recapture workers who have left the system and moved into the informal economy. While analysis is often quantitative, it is also possible to develop more sophisticated qualitative models to analyze such factors as the role of wages and behavior of firms in the regional economy. By nature, social security is a program with very large databases that contain much information on individuals and employers. While this large volume of information was difficult to manage in the past, the contemporary capacity to process data has allowed data mining to be very useful for a modern social security agency.

With respect to the other approaches studied in this Report, it should be mentioned that data mining employs econometrics and statistics intensively. In this sense, it is part of the approaches explained in previous sections. Good data mining applications can be used in QM, six-sigma, and BSC approaches.

For social security agency officials, the goal should not be to learn the technique of data mining. Rather, they should have an understanding of the subject at an operative level, including understanding the limits of data mining, the tools available, the quality of the work carried out by specialists, and the use menu-driven tools.

The success of an application is highly dependent upon the consolidation, cleaning, and organization of the data. The results will be applicable to BSC and other approaches. The software for data mining is sold commercially and usually compatible with the architecture of organizational systems.

Even information on the activities of public agencies can be subject to improper use if an inadequate regulation on revelation is adopted; for example, a requirement to provide information on public bids on contracts may unduly increase the costs of providing services. Thus, it is necessary to identify and evaluate the main legal challenges faced by evaluation strategies as technology provides for more substantial but also more complex strategies.

II.4 Comparative Advantages of the Approaches

This section presents a comparison of the relative strengths and weaknesses of each of the approaches

discussed above. Table II.2 summarizes the discussion. Because administrators are typically not experts in any of the academic approaches, their goal should be gaining understanding of the power of these approaches to be able to discuss them with experts and ultimately apply them to their agencies.

II.5 Approaches of the Main International Organizations

The final section of this chapter briefly describes the approaches used by some of the main international organizations. Each organization has a specific vision according to which it develops its goals. Naturally, its evaluation approach tends to be consistent with its goals. Some specialized agencies that manage a subset of social risks use approaches that are weighted towards financial, fiscal, social, or other criteria. Some agencies are regional and some global, some public and some private, and some non-governmental and some hybrids. It is important to note that some areas of evaluation seem to have received more attention while, with respect to certain social risks, other areas show a large gap in the measurement and understanding of realities.

Evaluation by international agencies sometimes refers to evaluation of their own operations and sometimes to the programs and agencies in their member countries. The aim of this section is to provide understanding of the evaluation approaches used by international agencies and support the systematic use of evaluation approaches by social security agencies.

The World Bank

The World Bank (WB) performs evaluation of both its own operations as well as social data with the aim of supporting its operations. Whereas the former type of evaluation is performed by the Independent Evaluation Group (IEG), which the WB created specifically for that purpose, the latter is more commonly linked to its credit operations.

The members of the World Bank agreed in 2002 to focus upon supporting countries' abilities to manage for results. Additionally, they agreed to work towards harmonizing the results-based approaches of all the development agencies. The 2006 Independent Evaluation Group (IEG) Report recognizes that only limited advances have been made in the strategy, to which it attributes the lack of a systematic application of methodologies; specifically, a "lack of capacity and the additional cost of data collection" at the country level, insufficient incentives for collaboration across teams, and the placing of too much attention on easy targets to obtain and measure results (World Bank 2006).

With respect to development activities, the WB prints a manual that recommends the following list of tools: performance indicators, the logical framework approach, theory-based evaluation, formal surveys, rapid appraisal methods, participatory methods, public expenditure tracking surveys, cost-benefit and cost-effectiveness analyses, and impact evaluation (World Bank 2004).

The Inter-American Development Bank

In 2002, the IADB members also agreed to stress the role of results-focused evaluation strategies. The Office of Evaluation and Oversight (OVE) of the IADB lists the following priorities in its most recent annual report: Country Program Evaluations, Policy and Instrument Evaluations, Sector and Thematic Evaluations, Ex-post Project Evaluations, Oversight of Bank Systems and Processes, Evaluation Capacity Development, and Participation in the Development Evaluation Community of Practice. The last priority listed by the IADB refers to the effort, primarily that of the Evaluation Cooperation Group (ECG), to link the major international financial agencies (IADB 2006).

The Evaluation Cooperation Group

The Evaluation and Cooperation Group (ECG) has as members the development banks of Asia, Africa, the Americas, and Europe, as well as the WB and the International Monetary Fund (IMF). It was founded in 1996 to support a learning environment across these agencies, which, being very large and well funded, play a large role in determining how countries ultimately evaluate themselves. The ECG focuses mainly upon the evaluation of operations by producing documents regarding such matters as evaluating the independence of an evaluation body and establishing good practices for the evaluation of loans and public sector operations. Perhaps its main innovation has been its development of a "peer review" process that aims to strengthen good practices in the evaluation of the operations of the large development agencies, as well as establish cross-reviews among them.

The UN Evaluation Group

The UN Evaluation Group (UNEG) faces a special challenge, as it comprises 43 agencies in quite diverse fields. This agency was redefined in 2000 after having been established in 1984 as the Inter-Agency Working Group on Evaluation. In 2005, it issued its Norms and Standards for Evaluation (see http://www.uneval.org for a complete listing). Each UN agency has its own evaluation policy. In 2006, UNEG established a working group on the issue of "delivering as one" wherein "one UN pilots at the country level, with one leader, one program, one budget, and where appropriate, one office" (UNEG 2007). This ambitious goal has been advanced in the Americas with the development of a pilot case in Uruguay (UNEG 2008). It is certainly of great interest to the world community that the United Nations aims to make advances in the modernization and streamlining in the evaluation of and by its agencies. Among the 18 UN agencies involved in the Uruguay pilot are the International Labour Organization (ILO) and the Pan American Health Organization (PAHO), two organizations that are greatly involved in social protection. Regardless of their areas of specialization, it is in the best interest of all organizations to support this UN effort.

OECD Development Evaluation

The Organization for Economic Co-operation and Development (OECD) focuses its efforts upon the evaluation of development assistance via its evaluation network, which has close ties to the Development Assistance Committee (DAC). OECD committees and the network, which have as members the delegates of the member countries, are supported by the OECD Secretariat.

To gain understanding of this effort, it is important to note that OECD members are the relatively wealthy countries in the international community that often maintain international aid agencies. Thus, the OECD network works as a knowledge exchange center to promote evaluation, harmonize the practices of different countries, and facilitate the coordination of studies across countries (see http://www.oecd.org for more information).