

Ten Steps
to a
Results-
Based
Monitoring
and
Evaluation
System

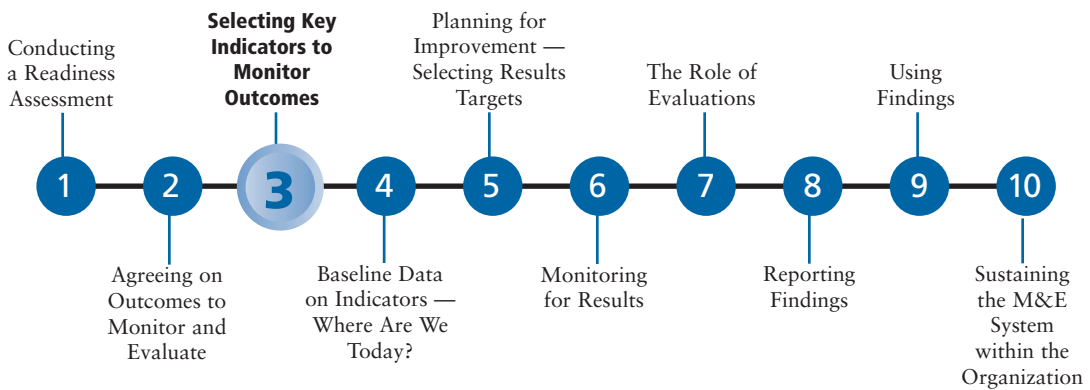
Jody Zall Kusek
Ray C. Rist



Chapter 3

Step 3: Selecting Key Performance Indicators to Monitor Outcomes

Figure 3.1



How will we know when we have achieved our desired outcomes? After examining the importance of setting achievable and well-defined outcomes, and the issues and process involved in agreeing upon those outcomes, we turn next to the selection of key indicators (figure 3.1). Outcome indicators are not the same as outcomes. Indicators are the quantitative or qualitative variables that provide a simple and reliable means to measure achievement, to reflect the changes connected to an intervention, or to help assess the performance of an organization against the stated outcome. Indicators should be developed for all levels of the results-based M&E system, meaning that indicators are needed to monitor progress with respect to inputs, activities, outputs, outcomes, and goals. Progress needs to be monitored at all levels of the system to provide feedback on areas of success and areas in which improvement may be required.

Outcome indicators help to answer two fundamental questions: “How will we know success or achievement when we see it? Are we moving toward achieving our desired outcomes?” These are the ques-

tions that are increasingly being asked of governments and organizations across the globe. Consequently, setting appropriate indicators to answer these questions becomes a critical part of our 10-step model.

Developing key indicators to monitor outcomes enables managers to assess the degree to which intended or promised outcomes are being achieved. Indicator development is a core activity in building a results-based M&E system. It drives all subsequent data collection, analysis, and reporting. There are also important political and methodological considerations involved in creating good, effective indicators.

This chapter specifically considers: (a) indicators required for all levels of the results-based M&E system; (b) translating outcomes into outcome indicators; (c) the “CREAM” of good performance indicators; (d) the use of proxy indicators; (e) the pros and cons of using predesigned indicators; (f) constructing indicators and tracking performance information; and (g) setting indicators using experience from developing countries.

Indicators Are Required for All Levels of Results-Based M&E Systems

Setting indicators to measure progress in inputs, activities, outputs, outcomes, and goals is important in providing necessary feedback to the management system. It will help managers identify those parts of an organization or government that may, or may not, be achieving results as planned. By measuring performance indicators on a regular, determined basis, managers and decisionmakers can find out whether projects, programs, and policies are on track, off track, or even doing better than expected against the targets set for performance. This provides an opportunity to make adjustments, correct course, and gain valuable institutional and project, program, or policy experience and knowledge. Ultimately, of course, it increases the likelihood of achieving the desired outcomes.

Translating Outcomes into Outcome Indicators

When we consider measuring “results,” we mean measuring outcomes, rather than only inputs and outputs. However, we must translate these outcomes into a set of measurable performance indicators. It is through the regular measurement of key performance indicators that we can determine if outcomes are being achieved.

For example, in the case of the outcome “to improve student learning,” an outcome indicator regarding students might be the change in student scores on school achievement tests. If students are continually improving scores on achievement tests, it is assumed that their overall learning outcomes have also improved. Another example is the outcome “reduce at-risk behavior of those at high risk of contracting HIV/AIDS.” Several direct indicators might be the measurement of different risky behaviors for those individuals most at risk.

As with agreeing on outcomes, the interests of multiple stakeholders should also be taken into account when selecting indicators. We previously pointed out that outcomes need to be translated into a set of measurable performance indicators. Yet how do we know which indicators to select? The selection process should be guided by the knowledge that the concerns of interested stakeholders must be considered and included. It is up to managers to distill stakeholder interests into good, usable performance indicators. Thus, outcomes should be disaggregated to make sure that indicators are relevant across the concerns of multiple stakeholder groups—and not just a single stakeholder group. Just as important, the indicators have to be relevant to the managers, because the focus of such a system is on performance and its improvement.

If the outcome is to improve student learning, then one direct stakeholder group is, of course, students. However, in setting up a results system to measure learning, education officials and governments might also be interested in measuring indicators relevant to the concerns of teachers and parents, as well as student access to schools and learning materials. Thus, additional indicators might be the number of qualified teachers, awareness by parents of the importance of enrolling girls in school, or access to appropriate curriculum materials.

This is not to suggest that there must be an indicator for every stakeholder group. Indicator selection is a complicated process in which the interests of several relevant stakeholders need to be considered and reconciled. At a minimum, there should be indicators that directly measure the outcome desired. In the case of improving student learning, there must be an indicator for students. Scores on achievement tests could be that particular indicator.

With the addition of outcome indicators (figure 3.2), we can expand on the performance framework for educational development outcomes introduced in the previous chapter.

What is the ideal number of indicators for any one outcome? The minimum number that answers the question: “Has the outcome been achieved?”

Figure 3.2
Developing a Set of Outcome Indicators for a Policy Area

Example: Education

Outcomes	Indicators	Baselines	Targets
1. Nation's children have better access to preschool programs	1. Percent of eligible urban children enrolled in preschool education 2. Percent of eligible rural children enrolled in preschool education		
2. Primary school learning outcomes for children are improved	1. Percent of Grade 6 students scoring 70% or better on standardized math and science tests		

The “CREAM” of Good Performance Indicators

The “CREAM” of selecting good performance indicators is essentially a set of criteria to aid in developing indicators for a specific project, program, or policy (Schiavo-Campo 1999, p. 85). Performance indicators should be clear, relevant, economic, adequate, and monitorable. CREAM amounts to an insurance policy, because the more precise and coherent the indicators, the better focused the measurement strategies will be.

Clear	Precise and unambiguous
Relevant	Appropriate to the subject at hand
Economic	Available at a reasonable cost
Adequate	Provide a sufficient basis to assess performance
Monitorable	Amenable to independent validation

If any one of these five criteria are not met, formal performance indicators will suffer and be less useful⁵.

Performance indicators should be as clear, direct, and unambigu-

ous as possible. Indicators may be qualitative or quantitative. In establishing results-based M&E systems, however, we advocate beginning with a simple and quantitatively measurable system rather than inserting qualitatively measured indicators upfront.

Quantitative indicators should be reported in terms of a specific number (number, mean, or median) or percentage. “Percents can also be expressed in a variety of ways, e.g., percent that fell into a particular outcome category . . . percent that fell above or below some targeted value . . . and percent that fell into particular outcome intervals . . .” (Hatry 1999, p. 63). “Outcome indicators are often expressed as the number or percent (proportion or rate) of something. Programs should consider including *both* forms. The number of successes (or failures) in itself does not indicate the rate of success (or failure)—what was not achieved. The percent by itself does not indicate the size of the success. Assessing the significance of an outcome typically requires data on both number and percent” (Hatry 1999, p. 60).

“Qualitative indicators/targets imply qualitative assessments . . . [that is], compliance with, quality of, extent of and level of . . . Qualitative indicators . . . provide insights into changes in institutional processes, attitudes, beliefs, motives and behaviors of individuals” (U.N. Population Fund 2000, p. 7). A qualitative indicator might measure perception, such as the level of empowerment that local government officials feel to adequately do their jobs. Qualitative indicators might also include a description of a behavior, such as the level of mastery of a newly learned skill. Although there is a role for qualitative data, it is more time consuming to collect, measure, and distill, especially in the early stages. Furthermore, qualitative indicators are harder to verify because they often involve subjective judgments about circumstances at a given time.

Qualitative indicators should be used with caution. Public sector management is not just about documenting *perceptions* of progress. It is about obtaining objective information on *actual* progress that will aid managers in making more well-informed strategic decisions, aligning budgets, and managing resources. Actual progress matters because, ultimately, M&E systems will help to provide information back to politicians, ministers, and organizations on what they can realistically expect to promise and accomplish. Stakeholders, for their part, will be most interested in actual outcomes, and will press to hold managers accountable for progress toward achieving the outcomes.

Every indicator has cost and work implications. In essence, when we explore building M&E systems, we are considering a new M&E system for every single indicator. Therefore, indicators should be chosen carefully and judiciously.

Performance indicators should be relevant to the desired outcome, and not affected by other issues tangential to the outcome.

The economic cost of setting indicators should be considered. This means that indicators should be set with an understanding of the likely expense of collecting and analyzing the data.

For example, in the National Poverty Reduction Strategy Paper (PRSP) for the Kyrgyz Republic, there are about 100 national and subnational indicators spanning more than a dozen policy reform areas. Because every indicator involves data collection, reporting, and analysis, the Kyrgyz government will need to design and build 100 individual M&E systems just to assess progress toward its poverty reduction strategy. For a poor country with limited resources, this will take some doing. Likewise, in Bolivia the PRSP initially contained 157 national-level indicators. It soon became apparent that building an M&E system to track so many indicators could not be sustained. The present PRSP draft for Bolivia now has 17 national-level indicators.

Indicators ought to be adequate. They should not be too indirect, too much of a proxy, or so abstract that assessing performance becomes complicated and problematic.

Indicators should be monitorable, meaning that they can be independently validated or verified, which is another argument in favor of starting with quantitative indicators as opposed to qualitative ones. Indicators should be reliable and valid to ensure that what is being measured at one time is what is also measured at a later time—and that what is measured is actually what is intended.

Caution should also be exercised in setting indicators according to the ease with which data can be collected. “Too often, agencies base their selection of indicators on how readily available the data are, not how important the outcome indicator is in measuring the extent to which the outcomes sought are being achieved” (Hatry 1999, p. 55).

Figure 3.3 is an additional checklist for assessing proposed indicators.

The Use of Proxy Indicators

You may not always be precise with indicators, but you can strive to be approximately right. Sometimes it is difficult to measure the outcome indicator directly, so proxy indicators are needed. Indirect, or proxy, indicators should be used only when data for direct indicators are not available, when data collection will be too costly, or if it is

“Better to be approximately correct than precisely wrong.”

(Anon.)

Figure 3.3

Checklist for Assessing Proposed Indicators

Outcome to be measured: _____

Indicator selected: _____

Is the indicator . . .

1. As direct as possible a reflection of the outcome itself? _____
2. Sufficiently precise to ensure objective measurement? _____
3. Calling for the most practical, cost-effective collection of data? _____
4. Sensitive to change in the outcome, but relatively unaffected by other changes? _____
5. Disaggregated as needed when reporting on the outcome? _____

Source: United Way of America 1996.

not feasible to collect data at regular intervals. However, caution should be exercised in using proxy indicators, because there has to be a presumption that the proxy indicator is giving at least approximate evidence on performance (box 3.1).

For example, if it is difficult to conduct periodic household surveys in dangerous housing areas, one could use the number of tin roofs or television antennas as a proxy measure of increased household in-

Box 3.1

Indicator Dilemmas

The Chicago Museum of Science and Industry—a large, cavernous museum with many monumental-size exhibits, including an entire submarine and a coal mine—wanted to conduct a study to determine which exhibitions were of greatest interest to its visitors. They found that it was impossible to count how many visitors viewed every exhibit, so they decided to use a proxy indicator. They did this by determining where they needed to replace floor tiles most often. And where did they find the floor tiles most in need of replacement? In front of the exhibit of hatching baby chicks.

Source: Webb et al., 1966.

come. These proxy indicators might be correctly tracking the desired outcome, but there could be other contributing factors as well; for example, the increase in income could be attributable to drug money, or income generated from the hidden market, or recent electrification that now allows the purchase of televisions. These factors would make attribution to the policy or program of economic development more difficult to assert.

The Pros and Cons of Using Predesigned Indicators

Predesigned indicators are those indicators established independently of an individual country, organization, program, or sector context. For example, a number of development institutions have created indicators to track development goals, including the following:

- MDGs
- The United Nations Development Programme's (UNDP's) Sustainable Human Development goals
- The World Bank's Rural Development Handbook
- The International Monetary Fund's (IMF's) Financial Soundness Indicators.

The MDGs contain eight goals, with attendant targets and indicators assigned to each. For example, Goal 4 is to reduce child mortality, while the target is to reduce by two-thirds the under-five mortality rate between the years 1990 and 2015. Indicators include (a) under-five mortality rate; (b) infant mortality rate; and (c) proportion of one-year-old children immunized against measles. (For a complete list of MDG indicators, see annex 3.)

The UNDP created the Human Development Index (HDI) in 1990 as a way of measuring human progress and the quality of life in all countries of the world. "The HDI constitutes the first comprehensive attempt to measure achievements in development from a human perspective, expressed in terms of numerical indicators that permit inter-country and inter-temporal comparisons . . . The index also provides an initial working tool that could be further developed and refined, and that could guide country efforts to establish relevant databases" (UNDP 2001).

More specifically, "[t]he UNDP's Human Development Index measures a country's achievements in three aspects of human devel-

opment: longevity, knowledge, and a decent standard of living. Longevity is measured by life expectancy at birth; knowledge is measured by a combination of the adult literacy rate and the combined gross primary, secondary, and tertiary enrollment ratio; and standard of living, as measured by GDP per capita” (UNDP 2001).

The World Bank’s Rural Development Indicators Handbook, based on the World Development Indicators, defines and disseminates international statistics on a broad set of rural indicators for rural well-being, improvement in the rural economy, development of rural markets, improvement of accessibility and communication, sustainable management of the resource base, and policy and institutional framework. Specific indicators include, for example, rural population below the poverty line, agricultural gross domestic product, agricultural exports, paved roads, potential arable land, and local tax revenue.

Thus, the Rural Development Indicators Handbook helps to develop a common approach to monitoring and evaluating progress both within and across countries using a common, clearly defined set of indicators. The Handbook also contains a Rural Score Card—a composite indicator that can be used, for example, to assess a country’s overall progress (or lack thereof) toward achievement of rural poverty reduction (World Bank 2000).

In light of regional financial crises in various parts of the world, the IMF is in the process of devising a set of Financial Soundness Indicators. These are indicators of the current financial health and soundness of a given country’s financial institutions, corporations, and households. They include indicators of capital adequacy, asset quality, earnings and profitability, liquidity, and sensitivity to market risk (IMF 2003).

On a more general level, the IMF also monitors and publishes a series of macroeconomic indicators that may be useful to governments and organizations. These include output indicators, fiscal and monetary indicators, balance of payments, external debt indicators, and the like.

There are a number of pros and cons associated with using pre-designed indicators:

Pros:

- They can be aggregated across similar projects, programs, and policies.

- They reduce costs of building multiple unique measurement systems.
- They make possible greater harmonization of donor requirements.

Cons:

- They often do not address country specific goals.
- They are often viewed as imposed, as coming from the top down.
- They do not promote key stakeholder participation and ownership.
- They can lead to the adoption of multiple competing indicators.

There are difficulties in deciding on what criteria to employ when one chooses one set of predesigned indicators over another.

Predesigned indicators may not be relevant to a given country or organizational context. There may be pressure from external stakeholders to adopt predesigned indicators, but it is our view that indicators should be internally driven and tailored to the needs of the organization and to the information requirements of the managers, to the extent possible. For example, many countries will have to use some predesigned indicators to address the MDGs, but each country should then disaggregate those goals to be appropriate to their own particular strategic objectives and the information needs of the relevant sectors.

Ideally, it is best to develop indicators to meet specific needs while involving stakeholders in a participatory process. Using predesigned indicators can easily work against this important participatory element.

Constructing Indicators

Constructing indicators takes work. It is especially important that competent technical, substantive, and policy experts participate in the process of indicator construction. All perspectives need to be taken into account—substantive, technical, and policy—when considering indicators. Are the indicators substantively feasible, technically doable, and policy relevant? Going back to the example of an outcome that aims to improve student learning, it is very important to make sure that education professionals, technical people who can construct learning indicators, and policy experts who can vouch for the policy relevance of the indicators, are all included in the discussion about which indicators should be selected.

Indicators should be constructed to meet specific needs. They also need to be a direct reflection of the outcome itself. And over time,

It will take more than one try to develop good indicators. Arriving at a final set of appropriate indicators will take time.

new indicators will probably be adopted and others dropped. This is to be expected. However, caution should be used in dropping or modifying indicators until at least three measurements have been taken.

Taking at least three measurements helps establish a baseline and a trend over time. Two important questions should be answered before changing or dropping an indicator: Have we tested this indicator thoroughly enough to know whether it is providing information to effectively measure against the desired outcome? Is this indicator providing information that makes it useful as a management tool?

It should also be noted that in changing indicators, baselines against which to measure progress are also changing. Each new indicator needs to have its own baseline established the first time data are collected for it. (The topic of setting baselines is covered in further detail in chapter 4.)

In summary, indicators should be well thought through. They should not be changed or switched often (and never on a whim), as this can lead to chaos in the overall data collection system. There should be clarity and agreement in the M&E system on the logic and rationale for each indicator from top level decisionmakers on to those responsible for collecting data in the field.

Performance indicators can and should be used to monitor outcomes and provide *continuous feedback and streams of data* throughout the project, program, or policy cycle. In addition to using indicators to monitor inputs, activities, outputs, and outcomes, indicators can yield a wealth of performance information about the process of and progress toward achieving these outcomes. Information from indicators can help to alert managers to performance discrepancies, shortfalls in reaching targets, and other variabilities or deviations from the desired outcome.

Thus, indicators provide organizations and governments with the opportunity to make midcourse corrections, as appropriate, to manage toward the desired outcomes. Using indicators to track process and progress is yet another demonstration of the ways that a results-based M&E system can be a powerful public management tool.

Setting Indicators: Experience in Developing Countries

More and more developing countries—and even regions—are beginning to set indicators to track progress toward their development goals. Boxes 3.2 through 3.4 review experiences in the Africa region, Sri Lanka, and Albania.

“The central function of any performance measurement process is to provide regular, valid data on indicators of performance outcomes.”

(Hatry 1999, p. 17)

Box 3.2**The Africa Region's Core Welfare Indicators**

Efforts are underway throughout the Africa region to create the basic statistical and technical building blocks of M&E systems. Among these building blocks are the core indicators surveys that have been conducted in a number of African countries, including Ghana, Malawi, Mozambique, Nigeria, and Lesotho. The Core Welfare Indicators Questionnaire (CWIQ) was created jointly by the World Bank, the UNDP, and UNICEF to monitor development objectives through the use of leading indicators in general, and social indicators in particular. “Leading indicators are indicators which give advance warning of a future impact, whose emergence may be delayed or difficult to measure” (<http://www4.worldbank.org/afr/stats/pdf/cwiq.pdf>).

Specifically, the CWIQ helps governments collect indicators related to household well-being, and indicators of access to, usage of, and satisfaction with basic services on an annual basis.

CWIQ features include the following:

- A fixed set of core questions with flexible modules
- Quick data entry and validation
- Simple reporting
- Large sample
- Short questionnaire
- Easy data collection.

“The CWIQ is not a complicated survey. It incorporates a package of features, which, when taken together, ensure wide coverage and a rapid turnaround time” (www.worldbank.org/afr/stats/pdf/ghcoreinds.pdf).

The CWIQ also “ . . . provides key social indicators for different population subgroups—within and across countries; [acts as] . . . an instrument for monitoring changes in key social indicators over time; and provides countries with a simple tool that produces rapid results” (World Bank p. 1).

At the same time, using the CWIQ does not prohibit in any way participant countries from also developing their own specific socioeconomic indicators.

For an example of a completed CWIQ, go to <http://www.worldbank.org/afr/stats/pdf/ghcoreinds.pdf>, which contains the Core Welfare Indicators for Ghana (<http://www4.worldbank.org/afr/stats/pdf/cwiqloop.pdf>).

Source: World Bank.

Box 3.3**Sri Lanka's National Evaluation Policy**

The government of Sri Lanka's National Evaluation Policy seeks to: (a) create an evaluation culture and to use evaluations to manage for results; (b) promote evaluation through capacity building with respect to staff, institutions, tools, and methodologies; (c) enable learning of lessons from past experiences; (d) improve the design of development policies and programs through integration of evaluating findings; and (e) establish accountability, transparency, and good governance.

As part of the evaluation policy, the government is mandating the use of performance indicators for all policy, program, and project preparation initiatives. For this purpose, the government is encouraging partnerships with civil society organizations (for example, the Sri Lanka Evaluation Association) and NGOs to introduce participatory evaluations in the public sector. The government is also encouraging universities and public sector training institutions to include evaluation modules to share knowledge on evaluation techniques and methodologies.

Also see annex 4: The Sri Lanka National Development Plan for Monitoring and Evaluation.

Source: Sri Lanka Evaluation Association and Ministry of Public Development and Implementation 2003.

To the extent possible, indicators should be developed based on the particular needs of a given country or organization. “. . . [T]he appropriate choice of performance indicators differ for different countries, times, and sectors. The only valid general rule is, therefore, when performance measurement is appropriate and cost-effective, performance should be assessed according to that combination of output, outcome and process indicators that are realistic and suitable for the specific activity, sector, country, and time” (Schiavo-Campo 1999, pp. 80–81).

Again, developing good indicators inevitably takes more than one try, and arriving at the final set of indicators will take time.

What we are ultimately building is a performance framework to provide countries and organizations with the means to develop strategies, set outcomes, build indicators, establish baselines, and set targets. This process will help guide the best use of budgets, resources, and personnel to achieve the desired outcomes.

Box 3.4

Albania's Three-Year Action Plan

Monitoring and evaluation systems—both implementation and performance-based—will be developed and used by the government of Albania to provide feedback on major programs constituting the Three-Year Action Plan (including all major strategic initiatives currently underway in Albania's public sector: the National Strategy for Social and Economic Development [NSSSED]; the Medium-Term Expenditure Framework; the Stabilization and Association Agreement; the Anti-Corruption Action Plan; and the Strategy for Decentralization and Local Autonomy). The government has assigned the Coordination Department within the Council of Ministers to oversee and coordinate implementation monitoring of the Three-Year Action Plan. Similar responsibilities for NSSSED performance monitoring will be assigned to the NSSSED Department within the Ministry of Finance.

The Ministry of Finance is expected to oversee the overall performance and implementation management by the 12 line ministries covered by the NSSSED. Responsibilities include: (a) procedures for setting indicators that will be tracked and reported on; (b) instructions to the line ministries on how to select indicators; (c) processes for selecting indicators to ensure they measure results that key stakeholders care about; and (d) procedures clarifying how information is to be collected against the indicators to ensure verification and reporting consistency.

Progress is also being made in the Education Ministry, which recently developed a draft NSSSED progress monitoring matrix. A new M&E unit has also been established within the Education Ministry, including six representatives of different departments. A variety of education indicators will be developed in connection with the government's Growth and Poverty Reduction Strategy, Poverty Reduction Support Credit, Education Project, and Education for All, initiatives. Education indicators include, among others, school attendance by educational level, teacher salaries, share of GDP spent on education, pupil-teacher ratio, percentage of the teaching force that meets ministry standards for qualified teachers, average class size, education completion rates overall, and education rates disaggregated for rural and poor families.

More generally, the Albanian government has basic statistical capacity (although there is room for improvement), and recently established a policy analysis unit. The government also has the indicators in place with respect to the MDGs.

Source: World Bank 2002a.

The following are examples of indicators at various levels:

Box 3.5 provides some useful examples of program and project level indicators.

Box 3.6 provides an example of an outcome and some possible indicators.

Box 3.5

Program and Project Level Results Indicators: An Example from the Irrigation Sector

Project name	Strengthening irrigation in a specific country area
Project goals	Improve agricultural productivity Raise farm income.
Indicators	
Outcome indicators	New area under irrigation Higher yield Increased production Increased farm income.
Output indicators	Construction of 10 new irrigation schemes Reconstruction of five old irrigation schemes Twenty-five farmer training sessions.

Source: Adapted from IFAD 2002, p.19.

Box 3.6

Outcome: Increased Participation of Farmers in Local Markets

Possible outcome indicators

- Percent change in annual revenue
- Percent change in amount of spoiled crops
- Percent change in crop pricing due to competition
- Percent change in agricultural employment.