CONTENTS

INTRODUCTION AND OVERVIEW  1
MONITORING AND EVALUATION—AND IMPACT EVALUATION  2
OUTCOME INDICATORS  3
WAYS TO MEASURE WOMEN’S ECONOMIC EMPOWERMENT  4
THE EVALUATION DESIGN CHALLENGE  7

Written by Mayra Buvinic and Rebecca Furst-Nichols, of the United Nations Foundation.

The United Nations Foundation and the ExxonMobil Foundation are especially grateful to Oriana Bandiera, James Knowles, Agnes Quisumbing, Martin Valdivia and Chris Woodruff, for their thoughtful contributions. The authors wish to thank our ExxonMobil Foundation colleagues, Suzanne McCarron, Noa Gimelli and Jim Jones, for their insights on this work and their leadership and support of women’s economic empowerment programming and research. Emily Pryor, Alba Bautista and Katy Johnson of the United Nations Foundation contributed to the project and the report. Patrick McCarthy and Cynthia Langliands of ExxonMobil and Jennifer Eastes of ERM provided additional comments and editing. Design, copy editing and layout were by Bruce Ross-Larson and his team at Communications Development Incorporated, including Joe Caponio, Christopher Trott, Lawrence Whitely and Elaine Wilson.
INTRODUCTION AND OVERVIEW

When women are economically empowered, communities and nations benefit. Yet, there has been a crucial knowledge gap regarding the most effective interventions that directly advance women’s economic opportunities. In early 2012, the United Nations Foundation and the ExxonMobil Foundation joined forces, launching a project to address this gap and identify which development interventions best improve women’s productivity and earnings.

The two foundations, under the technical leadership of United Nations Foundation Senior Fellow Mayra Buvinic, convened a select group of more than 35 development economists and other experts from top universities, international agencies and non-profit organizations. The researchers worked on 17 review and empirical studies that investigated practical, implementable projects aimed at women’s economic advancement. Together, the findings, with supporting evidence from more than 135 additional studies, were compiled into a report, A Roadmap for Promoting Women’s Economic Empowerment, that outlines which interventions may work best to increase women’s productivity and earnings in developing economies. The Roadmap was released in September 2013.¹

In 2014, the United Nations Foundation and the ExxonMobil Foundation invited researchers who developed the Roadmap to help identify outcome measures or indicators for women’s economic empowerment programs, informed by the researchers’ first-hand experience with rigorous research and program evaluation.² The following researchers contributed to this effort:

- Oriana Bandiera, London School of Economics
- James C. Knowles, Independent Consultant
- Agnes R. Quisumbing, Cheryl Doss, Nancy Johnson and Ruth Meinzen-Dick, International Food Policy Research Institute
- Martin Valdivia, GRADE, Peru
- Christopher Woodruff, University of Warwick

Each researcher produced an independent ‘think piece,’ covering both distinct and overlapping measurement topics. The foundations then convened the researchers to discuss their recommendations and draft a common set of widely applicable measures across two categories: urban women entrepreneurs and business leaders, and rural women entrepreneurs and farmers. The recommended measures focused on the ExxonMobil Foundation’s three goals for its Women’s Economic Opportunity Initiative: increased productivity, income and well-being.

² A measure or an indicator (the terms are synonymous) is a quantitative or qualitative variable that provides a simple and reliable measure of an outcome.
MONITORING AND EVALUATION—AND IMPACT EVALUATION

Evaluating program effectiveness is an important step toward understanding the successes and failures of each intervention. Project-level monitoring and evaluation (M&E) and impact evaluation are different, but complementary methods of evaluating effectiveness.

DOING THINGS RIGHT?

Project-level M&E is usually done internally and is an important part of learning about a program. M&E examines ‘are we doing things right?’—that is, has the intervention been implemented and are people being served? Do people know how to use the technology? Did people learn skills as a result of the training? Without knowing this critical first measure, there is no point in looking for impact. However, project-level M&E can cause an incentive problem, since implementing agencies want to show they did well. It can also result in an optimistic view if it only looks at those who remained in the program and not those who dropped out.

DOING THE RIGHT THINGS?

Impact evaluation seeks to determine whether the desired outcomes are being achieved, that is, if ‘we are doing the right things,’ and is in reference to a valid counterfactual (comparison group). To measure program impact, we need to understand what would have happened without the program. This is especially true for those programs where other factors could have affected outcomes, such as macroeconomic factors, other initiatives, weather or political changes. If the evaluation is designed prior to program implementation, the intervention can usually be phased in randomly to get a valid control group. A large enough sample size is needed for this work, which is often the challenge for smaller programs.

PRINCIPLES FOR MEASUREMENT

The research meeting emphasized the following principles for measuring women’s economic empowerment outcomes:

1. Given the interdependence of women’s economic and social roles, it is important to measure both economic and social (well-being) outcomes to understand women’s economic empowerment.
2. Similarly, it is important to measure effects at both the individual and household levels, considering the broader context of women’s well-being within the household.
3. The WHAT and the HOW of an evaluation matter equally: ‘what’ refers to the indicators selected, while the ‘how’ is the evaluation design.
4. No evaluation is better than a poorly designed evaluation.
5. Not every program can be rigorously evaluated, but something can be learned from every program.
6. Complementary qualitative work is important to understand the ‘why’ behind results.
OUTCOME INDICATORS

Outcome indicators document the causal chain between the program intervention and outcomes. Every step of the causal chain (input → direct outcome → intermediate outcome → final outcome) should be measured with clear indicators to better understand which effects materialize and why.

There are three types of outcomes: direct, intermediate and final. A direct outcome in women’s economic empowerment programs is the knowledge, skill or asset acquired via a specific intervention (e.g. a new process learned). An intermediate outcome is the practice or behavior change that results from the direct outcome (e.g. a change in business practices or a shift in the household division of labor). The final outcome is the intended principal effect of the program (Bandiera). Final outcomes include, for instance, increases in employment and profits as a result of changes in business practices brought about by business training. Other telling final outcome indicators are savings and investments in productive assets, such as farming machinery and business inventory, and measures of women’s empowerment and well-being. Figure 1 (adapted from Bandiera) shows the causal chain between different outcomes and gives examples of different outcome indicators.

Example: Innovative collateral offered through a credit program (input) enables the woman business owner to borrow capital (direct outcome), which she uses to hire additional workers (intermediate), to expand production and profits (final).

FIGURE 1. EXAMPLES OF DIRECT, INTERMEDIATE, AND FINAL OUTCOMES

<table>
<thead>
<tr>
<th>DIRECT OUTCOMES</th>
<th>INTERMEDIATE OUTCOMES</th>
<th>FINAL OUTCOMES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knowledge, skills and productive assets acquired, new technologies effectively used</td>
<td>Changes in business/farming practices, women’s decision-making roles in their businesses/farms, technology adoption and effective use, women’s self-confidence</td>
<td>Business income, employment, household income, asset ownership, subjective well-being, gender roles/norms, women’s self-confidence/self-esteem</td>
</tr>
</tbody>
</table>
Direct outcome indicators are most closely linked with the particular program intervention and can include the acquisition of information and concepts, training attendance and proper use of technology. These indicators pick up the most immediate results and should be constructed to fit the specific intervention. An important function of direct indicators is to provide information for monitoring, to gauge if the intervention was successfully delivered and adopted. For instance, the lack of growth in profits following a management training program could be due to the fact that the woman missed training lessons and did not learn critical concepts. Similarly, lack of results from access to high impact information or productive technologies may be because women have access to but do not know how to use the technology in question.

Intermediate outcome indicators, as the name implies, are measurable outcomes that are related to and occur before the final outcome. Intermediate measures should be causally related to final outcomes, as they can sometimes take the place of or become final outcome indicators, especially when the final outcome measure may take time to materialize.

Since final outcomes often depend on factors beyond the control of the woman or the intervention (such as fluctuations in economic cycles affecting business earnings), intermediate indicators should be selected that are more within her control. A risk is that intermediate outcomes may show impacts of the program intervention, but still fall short of providing information on final outcomes of interest. For instance, in Sri Lanka, de Mel et al. (2014) found that while business management training improved the management practices of women business owners, the changes in these practices did not lead to increases in profits or sales.³

Final outcome indicators are designed to determine whether the program had its desired outcome. It is ideal to measure final outcomes at varying intervals, since program impacts may be delayed or short-lived. Empowerment outcomes can be both intermediate outcomes, playing an instrumental role in helping to attain final outcomes (e.g., such as in the case of shifts in the allocation of household labor that enable women to devote more time to the enterprise or improved self-confidence that enables women to increase business risk-taking), as well as final program outcomes.

WAYS TO MEASURE WOMEN’S ECONOMIC EMPOWERMENT

The list below outlines some of the main ways to measure women’s economic empowerment, according to the type of outcome indicator. More specific indicators can be derived from these general categories. It tailors measures to urban versus rural environments, but does not recommend one indicator over another. The choice of indicator will depend on the nature of the program being evaluated, and not every program will be able to measure the final outcomes suggested.

DIRECT OUTCOME INDICATORS
- Specific information and knowledge acquired in training (measure has to be program specific)
- Skills developed through training programs (measure has to be program specific)
- Pre and post training scores and training attendance records
- Technology adoption and effective use (measured through frequency of use of high impact information and other productive technologies)

INTERMEDIATE OUTCOME INDICATORS
- Changes in business or farming management practices
- Perceived value of business and agricultural training and technology
- Intensity of mobile phone use for productive purposes
- Participation and involvement in formal and informal business-related associations (including indicators of participation, such as attendance rates)
- Changes in traditional gender roles in the household and changes in indicators of self-esteem (from independent reporting or self-reporting on questions regarding traditional household chores)

FINAL OUTCOME INDICATORS
Urban women (entrepreneurs and business leaders)
- Profits (variable and difficult to measure)
- Revenues (easier to obtain than profits, but still difficult to measure, especially for microfirms)
- Employment, including self-employment, regularity of employment, hours worked and income earned, as well as employment generated by the firm (potentially easier to measure in larger firms)

Rural women (entrepreneurs and farmers)
- Individual and household assets, including women’s individual ownership and control of assets
- Investment in productive assets such as land and livestock, including ownership and control questions
- Employment, including self-employment, regularity of employment, hours worked and income earned

All women
- Household income/expenditure and control over it
- Individual savings (independent from household and joint male and female savings)
- Women’s well-being, including indicators of self-esteem, satisfaction with work and life, and stress levels (many of these indicators are based on self-reporting so care needs to be taken to choose well-constructed measures)
- Women’s empowerment indicators of increased choice and decision making in family and public life (through independent reporting or self-reporting of behavior)
GOOD PRACTICES FOR DESIGNING OUTCOME INDICATORS

When constructing outcome indicators through surveys or questionnaires, it is desirable to:

1. **Keep it numerical:** Quantitative measures with comparable units allow interpersonal comparisons that are key to any program evaluation. For instance ‘how many hours do you work per week?’ yields comparable answers; ‘do you work full-time or part-time’ does not as the definition of full-time and part-time might differ for different people.

2. **Keep it easy:** Most respondents have low levels of education, and this raises challenges for the measurement of concepts requiring probabilities or percentages. Visual and manual instruments can often be effective. Two examples that are known to work well are ‘smiley scales’ to measure satisfaction and ‘bag of beans’ to elicit the allocation of scarce resources—e.g. time devoted to different activities, or earnings by month.\(^5\)

3. **Keep it short:** As fatigue sets in, it becomes difficult for participants to focus on the task. Survey modules should alternate fun/easy sections with long/tedious sections. Piloting will help to identify which sections respondents are most likely to answer.

4. **Keep it consistent:** Use the same unit of measure whenever possible (e.g. in a time use survey always ask about hours, or minutes), refer to the same time interval when asking recall questions, and make sure multiple choice options are mutually exclusive and the list is complete.

5. **Give a way out:** Always include ‘don’t know’ and ‘refuse to answer’ options.

---

\(^{4}\) See Bandiera, Oriana. 2014. Independent Research Note.

\(^{5}\) Respondents are given a set of cards representing different activities (e.g. in a time use survey these would be ‘work,’ ‘taking care of children,’ ‘cooking,’ etc.) and a bag of beans that they are asked to allocate to the different cards in proportion to the time they devote to each activity.
THE EVALUATION DESIGN CHALLENGE

FEATURES OF A GOOD EVALUATION METHODOLOGY

Outcome indicators are just one component of a good evaluation methodology, and they need to be embedded into a robust evaluation design. An effective program impact evaluation is characterized by several key components. First, a valid counterfactual or comparison group demonstrates whether and how the program changed outcomes for the participants compared to controls, that is, those with similar or identical characteristics but who did not participate.\(^6\) Second, a causal chain should be incorporated that shows how program effects materialize, which helps inform whether to further refine or scale-up programs (Bandiera). When measuring the income of individuals in households with more than one earner, a third consideration comes into play: since interventions may cause shifts in a household’s allocation of labor or other resources, the evaluation should be designed to uncover what has occurred both at the household and individual levels (Knowles).

Fourth, some interventions are particularly challenging to measure due to the inherent difficulty in developing a robust evaluation design, and results from these evaluations should be examined for possible alternative explanations. Identifying a valid comparison group, for instance, for interventions targeted at small and medium enterprises (SMEs) is problematic since there are fewer women-owned SMEs than women-owned microenterprises, making it difficult and costly to obtain a large enough sample for a randomized trial. Without randomization, it is difficult to make unambiguous attributions of causality (Woodruff). Isolating the impact of mobile technology is also difficult due to the nature of cell phone adoption and use, which often happens rapidly and universally in a given population. Information is nearly impossible to ‘control,’ meaning that it is hard to know which groups have been exposed to the intervention or technology and whether a valid comparison exists (Vaitla).\(^7\) In both of these cases, evaluators need to examine and be clear about possible other causes, in addition to the intervention, of the outcome measured.

A final issue that is often not sufficiently considered in program evaluations is the length or sustainability of the intervention’s impact. Some interventions take time to show effects. A recent study by Valdivia (2013) in Peru showed a positive effect of business management training on businesswomen’s profits two years after the program was completed, although it had shown no effect when business outcomes were first measured. The effects of training built up over time and emerged with a two-year delay.\(^8\) However, the opposite can also be true: other effects diminish or disappear over time, as was the case with growth in business profits from the combined effects of business training and a capital transfer to poor women entrepreneurs in Sri Lanka (Woodruff).

---

\(^6\) See Bandiera for a list of commonly used comparison groups, with notes on their validity.


Cases like these justify the importance of considering the long-term effects of program interventions and building in robust intermediate and final outcome indicators (Woodruff). Some indicators are better than others at being able to capture effects that vary over time. For instance, asset measurements have proven to be a valid measure of longer-term household income, but a less robust measure of effects that vary over time (Knowles; Quisumbing and colleagues).

**DESIRED OUTCOMES**

Across the different types of economic empowerment programs and women reached, researchers agreed that desired overall outcomes for measurement include growth in both the objective dimension of income or assets and the more subjective dimension of empowerment. Increases in income should translate into increased well-being for women – considering both monetary and in-kind income, as well as the allocation of women’s time to work and leisure (Quisumbing and colleagues; Woodruff). The subjective dimension of empowerment can be gauged using proxy indicators that can be measured objectively, such as investments in children’s schooling and health (Knowles), or by women’s self-reporting on attributes linked to empowerment, including labor market roles and the allocation of household tasks (Bandiera); ownership and control over assets (Quisumbing and colleagues); life satisfaction, social capital and household decision-making (Valdivia); and emotional states related to satisfaction and well-being, including women’s stress levels (Woodruff).

**THE CHALLENGE OF MEASURING WOMEN’S ECONOMIC OUTCOMES**

In addition to design challenges common to most program evaluations, there are particular challenges related to measuring women’s empowerment: first, women are embedded in households, which makes it difficult to separate and measure the effects of programs in one domain when they ‘spill over’ into another domain, as happens often in households. For instance, and especially in rural settings, measuring a change in the income of a single household member is not only difficult in itself, but such measurement may neglect possible positive or negative ‘spillover’ effects of the individual’s activities on the income-earning activities of other household members (Knowles). While this challenge is applicable in theory to all family members, it is particularly an issue for women because of the strong interdependence between their economic and family roles.

Second, this intertwining of economic and social roles in women’s lives means that barriers to either dimension can prevent progress on the other (Bandiera; Woodruff). Women’s family roles may influence business choices and returns to those businesses. Microenterprises owned by women are often interlinked with household activities, for instance, and decisions on sector, time spent and growth objectives are made with both roles in mind. Where interventions cause the business to scale-up, they may also result in more stress in juggling household responsibilities, offsetting any potential gains in profits or revenue (Woodruff). Potential gains in profit or business growth may also be compromised by family or community pressures to share windfall cash, pressure to which women are exposed more than men in poor households in developing countries.
Third, measuring the unobservable, subjective elements of economic empowerment is difficult but important, since it both impacts final outcomes and may itself be a valued outcome. Valdivia identifies three dimensions of ‘empowerment’ that can be measured: life satisfaction, social capital and household decision-making. However, Quisumbing and colleagues argue that women’s control of income or their degree of participation in economic decision-making are the most robust known measures of empowerment and economic well-being. In a departure from these more subjective measures, Knowles suggests a frequently used set of proxy indicators for women’s empowerment: household expenditure on women’s and children’s medical care, and school enrollment and expenditure on children’s education, by gender of the child.

In sum, increases in women’s economic empowerment span economic and social spheres, have ‘spillover’ effects to other domains in women’s lives and to other family members, enhance subjective well-being, and include ownership and control over income and independence (from men) in individual decision-making regarding economic and social matters. Outcome indicators should be designed to capture all of these different elements of empowerment.

**ILLUSTRATIVE INDICATORS FOR URBAN WOMEN ENTREPRENEURS AND BUSINESS LEADERS**

For women entrepreneurs and business leaders in urban areas, ways to measure business growth and income (profits and revenues, and employment obtained and generated) and management practices (resulting from training interventions) are key considerations in communicating overall outcomes. This section explores both areas, and provides some examples of possible indicators.

Measuring profits is difficult, especially for the smallest enterprises (Valdivia; Woodruff). Measurement requires the respondent to recall figures on sales and cost figures, and assumes she has this information readily at hand. This poses general challenges for accuracy, which are intensified if the intervention itself changes the quality of such calculation through new knowledge and skills. A follow-up survey, then, could detect an impact on profits attributed to the intervention when the true impact is the woman’s improved registry of sales or the inclusion of fixed costs. An additional issue, as with any survey that seeks income information, is that trust often impacts a respondent’s willingness to provide accurate information; respondents are likely to underreport income to unfamiliar individuals. If trust in the surveyor or program has improved by the follow-up survey, it is possible to attribute a program effect that is not there, but rather a result of improved trust or allegiance to the program (Valdivia).

Since profits correlate most directly with income, they remain an important indicator, but revenues are also important to collect (Woodruff). Profits may vary such that they do not reflect the true value of the firm, for instance, if the measurement period (e.g. monthly, quarterly) is below average for the industry surveyed, or if investments have been made too recently for increased profitability to materialize (Valdivia). A well-designed randomized experiment can overcome the measurement period issue if the treatment and control groups are aligned.
Measurement of intermediate outcomes that lead to the profit stream is appropriate when investments are too recent to show results.

Profits may not be an appropriate indicator for some women for whom enterprise growth is not the goal, and indeed could be counterproductive. Women who juggle household responsibilities with their businesses may experience more stress when their business expands (Woodruff). These women may look instead for stability of income or increased productivity as a result of the intervention, rather than business growth. Thus, women’s well-being is an important outcome that should be measured. In these cases, it will be left to the judgment of evaluators and researchers how well these outcomes convey real increases in women’s economic empowerment. For instance, increases in business productivity that maintain business profits may be more easily interpreted as increasing economic empowerment than increases in business productivity with reduced business profits. Qualitative data about the types of women participating in the intervention, and their goals for running a business, should be useful to collect alongside profit and revenue information to address this issue.

Changes in management practices are one commonly used intermediate indicator for women entrepreneurs, but others may be identified by specifying the proposed theory of change, which links inputs to direct, intermediate and final outcomes. For instance, innovative collateral offered through a credit program (input) enables the woman business owner to borrow capital (direct outcome), which she uses to hire additional workers (intermediate), to expand production and profits (final).

Appropriate intermediate outcomes can vary significantly by context (Valdivia), and are more directly in the woman’s control at any given time than final outcomes (Woodruff). A risk is that intermediate outcomes may show impacts of the program intervention, but still fall short of providing information on final outcomes of interest. For instance, in Sri Lanka, de Mel et al. (2014) found that while business management training improved the management practices of women business owners, the changes in management practices induced by training did not lead to increases in profits and sales.9

A related measurement issue that evaluators should be aware of is a potential statistical bias when a program evaluation includes many independent indicators of similar outcomes. This artificially increases the probability of having one or more false positive outcomes that are statistically significant. One way to address this potential statistical bias is to collapse many indicators into a single summary measure known as the ‘mean standardized treatment effect’ (Valdivia).10

---


10 This single summary measure, the ‘mean standardized treatment effect,’ is devised by taking the unweighted average of the standardized outcomes of a family of outcomes (see Valdivia).
ILLUSTRATIVE INDICATORS FOR RURAL WOMEN ENTREPRENEURS AND FARMERS

In this section, we focus on three main areas of measurement that are particularly applicable to rural women entrepreneurs and farmers. The first, measuring assets, is especially applicable to rural populations due to the relative ease of measurement and accuracy in predicting empowerment. The second area, household income and household consumption expenditures, can be applied in both urban and rural settings. The third area covers measurement challenges and opportunities related to mobile technology, given the growing importance of this technology in addressing constraints women face to increase productivity and earnings in rural economies.

Asset measures are preferable to measures of employment and income, especially for women farmers and rural producers. Physical and financial assets (such as livestock and bank accounts) are easier to measure in rural environments and more accurate predictors of empowerment than income-based measures (Knowles; Quisumbing and colleagues). Asset measures are less sensitive, however, to detecting short-term variations, so they are better medium- and long-term indicators of wealth (Knowles; Quisumbing and colleagues). To detect short-term variations, they need to be replaced or complemented by estimates of household consumption (or a proxy based on income-sensitive consumption items).

Household income, the measurement of income-earning activities of all household members, is applicable as an outcome indicator to a wide range of interventions (Knowles). If measured well (to account for spillovers to other household members, as described previously), it captures actual changes in income as a result of an intervention. Unfortunately, household income is difficult and costly to measure directly in practice, particularly in rural households. Short-term fluctuations due to weather variation or holiday expenditures that are unrelated to the intervention are likely to introduce unwanted variability to household consumption expenditures (Knowles).

Knowles recommends asset indices as a promising alternative measure of long-term household income which is closely correlated with household consumption, and gives asset index examples in his note. Asset indices are constructed using information on housing characteristics (e.g. materials of walls, floor and roof) and ownership of consumer durables, relying entirely on information about physical assets. Asset indices have been used frequently in large-scale surveys such as the Demographic and Health Surveys (DHS) and the Multiple Indicator Cluster Surveys (MICS). Corrections need to be made when comparing asset indices across regions, since types of household durables owned in different areas could vary substantially.

Quisumbing and colleagues propose measuring tangible (physical), financial and intangible (social networks and skills) assets by looking at men and women's

11 Work by the Gender, Agriculture, and Assets Project has shown that gendered asset ownership measures are responsive to program interventions of moderate (three to five years) duration (Quisumbing and colleagues).
12 Knowles also demonstrates the properties of household asset measures when compared to household income and consumption measures (for the case of Cambodia).
exclusive and joint use, control and ownership. Like Knowles, they recommend measuring at the household level by obtaining an inventory of all the relevant tangible assets, asking each individual who owns and has rights over each, as well as its value. Use can be discerned by asking who uses the asset most of the time, control by asking who decides whether to make the asset purchase, and ownership through questions around rights to sell, mortgage, rent out, or give away the asset.

To measure financial assets, each individual in the household should be asked about all forms of savings and accounts, including whether there is a positive balance, and if they are individually or jointly held. Availability of cash or in-kind credit, in addition to remittances and public sector transfers, rounds out a complete set of questions on financial assets. Questions about social capital (e.g. group membership, social networks, perceived status) and human capital (skills) are used to measure intangible assets (Quisumbing and colleagues).

Savings, which may be in the form of physical or financial assets, can be defined as the difference between household income and consumption during a given period (one year). Savings is an important outcome indicator because increased savings can: (1) increase household investment in productive assets; (2) smooth consumption in event of emergency expenditures; and (3) increase women’s empowerment if women have control over some savings. Like other asset measures, Knowles recommends measuring savings at the household level to avoid the risk of capturing a shift away from other household assets rather than a true increase. However, if measured at the individual level, savings can be a good proxy measure for women’s empowerment (Knowles).

Quisumbing and colleagues recommend collecting both asset and income information because their research shows that results from asset and income measures are not highly correlated with each other; each reveals different information about how a program is affecting women’s economic empowerment. When these measures appear to move in different (and unexpected) directions, it is important to draw on project logic models or theory of change, as well as qualitative information from beneficiaries and project staff, to understand why differential impacts are appearing.

Assessing the impact of mobile phone technologies on women’s economic empowerment presents particular challenges and opportunities. Two types of mobile interventions are especially relevant: provision of market price information and mobile money applications that facilitate cash transfers. Women’s net income (profit), the stability of these income flows, and changes in women’s asset stocks are the key final outcomes resulting from increased access to market price information and mobile funds. Important intermediate outcomes include volume of money transferred from/to women through mobile technologies and market-level changes in prices, as cell phones improve information flows. Lastly, direct outcome indicators of interest include actual usage of market prices and money transfer services (Vaitla).