Economics Letters 122 (2014) 365–369

Contents lists available at ScienceDirect

Economics Letters

journal homepage: www.elsevier.com/locate/ecolet

An exercise to evaluate an anti-poverty program with multiple outcomes using program evaluation

T.M. Tonmoy Islam *

Department of Urban and Regional Studies/Department of Economics, University of Wisconsin–Green Bay, 2420 Nicolet Drive, MAC B326, Green Bay, WI 54311, United States

HIGHLIGHTS

- I introduce multidimensional poverty measures to the program evaluation literature.
- This shows a measure of program evaluation with multiple outcomes.
- I apply it to evaluate the Targeting the Ultra Poor (TUP) program in Bangladesh.
- TUP reduced multidimensional poverty of the treated group by 18 percentage points.

ARTICLE INFO

Article history:
Received 23 July 2013
Received in revised form
16 December 2013
Accepted 20 December 2013
Available online 30 December 2013

JEL classification:
I31
I32

Keywords:
Program evaluation
Multidimensional poverty index

ABSTRACT

To measure the effectiveness of programs with multiple outcomes, I combine multidimensional poverty measures with difference-in-difference matching estimators. I apply this technique on an anti-poverty program and show it is a more comprehensive measure of poverty reduction than a unidimensional measure.

1. Introduction

Measuring the effectiveness of anti-poverty programs has gained increased urgency as governments work to reduce poverty in order to meet the United Nation's Millennium Development Goals (MDGs). Historically the standard metric of program effectiveness has been income (per capita), and indeed one of the MDGs is halving the number of people living on less than a $1 a day (PPP) by 2015.[1] However, in recent years policymakers in both developed and developing nations have stressed the need to focus on how policy affects multiple dimensions of well-being. This is because in many cases, measuring poverty status of a person based on their earned income may yield some perverse results. For example, suppose there are three dimensions of well-being – income, consumption and health – and that a person earns income above the poverty line. Using a headcount ratio or the income-gap ratio or any other poverty measure that uses income as a dimension to measure poverty, the person is said to be non-poor. However, the person may be suffering from a serious illness that requires substantial out-of-pocket medical expenditures for treatment and so, after medical expenses are met, the person's remaining income may fall well below the poverty line. Therefore, this person's discretionary consumption level is the same as that of a person living in poverty, even though their income is above the poverty line. It would be logical to consider this person to be in poverty, but the typical income poverty measure would not count this person as poor.[2]

Thus, a central challenge is on how to quantify poverty across multiple dimensions and then evaluate the effectiveness of

* Tel.: +1 920 465 2396.
E-mail address: islamt@uwgb.edu.

[1] "MDG Monitor". http://www.mdgmonitor.org/browse_goal.cfm.

[2] Based on the recommendation of a 1995 National Academy of Sciences (NAS) panel, the Census Bureau is producing a Supplemental Poverty Measure whereby out-of-pocket medical expenses are subtracted from income when calculating income poverty. This has a large effect of nearly doubly poverty among the elderly http://www.census.gov/hhes/povmeas/methodology/supplemental/research/Short_ResearchSPM2010.pdf.

0165-1765/$ – see front matter
http://dx.doi.org/10.1016/j.econlet.2013.12.032

interventions across multiple outcomes. To address the first issue, economists have developed multidimensional poverty indices (MPI) (Alkire and Foster, 2011; Alkire and Santos, 2011; Bourguignon and Chakravarty, 2003). To answer the second issue, I combine a version of the MPI with the econometrics of program evaluation to assess the effectiveness of a large-scale experiment to combat chronic poverty in Bangladesh.

1.1. Multidimensional poverty index

The first step is to select an appropriate MPI, and there are many versions available to choose from (Alkire and Foster, 2011; Alkire and Santos, 2011; Bourguignon and Chakravarty, 2003). I use the following extension of the Foster et al. (1984) index (Bourguignon and Chakravarty, 2003) as an illustration:

$$P_\alpha (X; z) = \frac{1}{n} \sum_{j=1}^{m} \sum_{i=1}^{n} a_j I_{ij}(x_{ij} < z_j) \cdot \left(1 - \frac{x_{ij}}{z_j}\right)^{\alpha} \tag{1}$$

where, m is the number of dimensions being used to assess poverty of individuals/households, x_{ij} is the level of the dimension of well-being of individual/household i who is living in poverty, z_j is a predetermined level of the poverty line of dimension j, $I_{ij}()$ is an indicator function that is equal to 1 if $x_{ij} < z_j$, and 0 otherwise, a_j is the weight given to each dimension, and n is the total number of people/households in the community. α is a number which gives different measures of poverty.[3]

Among the many different versions of MPI available, I choose another version to illustrate the robustness of my results. I use another version of the multidimensional poverty measure that also in Bourguignon and Chakravarty (2003).[4]

$$P_\alpha{}^{BC} (X; z) = \frac{1}{n} \sum_{j=1}^{m} \sum_{i=1}^{n} a_j I_{ij}(x_{ij} < z_j) \cdot \left[\left(1 - \frac{x_{ij}}{z_j}\right)^{\theta}\right]^{\frac{\alpha}{\theta}}. \tag{2}$$

In Eq. (2), all parameters are the same as that in Eq. (1) except for the addition of θ, which represents a substitutability parameter. As θ increases, the substitutability between the dimensions decline.

1.2. Program evaluation

The next step is to apply a version of the MPI to evaluate the effectiveness of a program. Programs that do not involve random placement of individuals in the treatment and the control groups confront the challenge of "re-creating" the experimental environment. One common way of evaluating non-experimental programs is by using the technique of matching. It involves comparing the outcome of a program participant (Y_1) with those of certain non-participants (Y_0) that have similar characteristics as the participants. Any difference in the outcomes between the participants and the non-participants can be said to be due to the impact of the program. Heckman et al. (1998) show that the estimated gain of a participant from the program, after controlling for program participation ($D = 1$) and characteristics X is:

$$E(Y_1 - Y_0 \mid D = 1, X). \tag{3}$$

If a particular domain of characteristics, or region of common support X, is used then the equation for the evaluation of a program becomes:

$$\sum_{i \in I_1} w_{N_0,N_1}(i) \left(Y_{1i} - \sum_{k \in I_{01}} W_{N_0,N_1}(i,k) \cdot Y_{0j}\right) \tag{4}$$

where Y_{1i} is the outcome of a person in the treated sample, Y_{0j} is the outcome for matched persons k. I_{01} is the set of people in the comparison group. $W_{N_0,N_1}(i,k)$ is a positive weight function so that for all i, $\sum W_{N_0,N_1}(i,k) = 1$, and N_0 and N_1 are the number of people in the treatment and the comparison group, respectively. $w_{N_0,N_1}(i)$ is the weight given to each participant, with I_1 being the set of participants, and therefore, $\sum w_{N_0,N_1}(i) = 1$. Under equal weights $w_{N_0,N_1}(i) = \frac{1}{N_0}$.

One popular matching method is to use propensity scores (Rosenbaum and Rubin, 1983; Dehejia and Wahba, 2002). The basic idea is to predict the probability of program participation as a function of observables Xs, and then match members of the treatment and comparison groups based only on a scalar probability value rather than matching across the potentially large dimension of Xs. Although this approach controls for 'selection on observables', a potentially more robust approach that also controls for time-invariant unobservable heterogeneity is to combine matching with a difference-in-differences estimator (Smith and Todd, 2005; Todd, 2006):

$$\frac{1}{N_0} \sum_{i \in I_1} \left(Y_{it1} - \sum_{k \in I_{01}} W_{N_0,N_1}(i,k) \cdot Y_{kt0} - \left(Y_{it'1} - \sum_{k \in I_{01}} W_{N_0,N_1}(i,k) \cdot Y_{kt'0}\right)\right), \tag{5}$$

where t' and t are the time periods before and after the implementation of the program respectively, and $W_{N_0,N_1}(i,k)$ is the weight given to the members of the control group that have been matched with a member of the treatment group using propensity score matching at time t'.

The typical evaluation focuses on a single outcome, Y, and when more than one outcome is addressed it is most common to evaluate each component sequentially. In the presence of multidimensional poverty measures such as in Eq. (1) the difference-in-difference matching estimator in Eq. (5) is modified slightly by replacing Y with P_α as follows:

$$\frac{1}{N_0} \sum_{i \in I_1} \left(P_\alpha(y; z)_{it1} - \sum_{k \in I_{01}} W_{N_0,N_1}(i,k) \cdot P_\alpha(y; z)_{kt0} - \left(P_\alpha(y; z)_{it'1} - \sum_{k \in I_{01}} W_{N_0,N_1}(i,k) \cdot P_\alpha(y; z)_{kt'0}\right)\right). \tag{6}$$

Since a decrease in the value of $P_\alpha(y; z)$ means that poverty level has decreased, a negative value of Eq. (6) would indicate that the program was successful.

2. Data

I implement the difference-in-difference matching estimator in Eq. (6) to evaluate the effectiveness of the Targeting the Ultra Poor

[3] When $\alpha = 0$, $P_\alpha (X; z)$ becomes the MPI headcount ratio, and it measures the average deprivation of dimensions per person. When $\alpha = 1$, $P_\alpha (X; z)$ is the poverty gap ratio, which measures the average normalized shortfall among the dimensions of all the individuals/households in the sample. When $\alpha = 2$, the measure is called the squared-poverty gap, and it measures the average normalized shortfall of the dimensions of well-being, like the $P_1 (X; z)$ measure, but in this case, it puts more weight on poorer people in the community. As $\alpha \to \infty$, increasing weight is placed on the worse-off individuals. So, this shows that, for a given α, the higher the value of $P_\alpha (X; z)$ the higher is the poverty in that community.

[4] A referee pointed out that other available versions of the multidimensional poverty measures can be used to illustrate the main findings and the robustness of the results. I do acknowledge that other versions can be used to show the importance of using multidimensional poverty measures in evaluating programs with multiple outcomes. However, I used the measures from Bourguignon and Chakravarty (2003) because they are very similar to the widely-used Foster et al. (1984) measure, relatively simple to implement and explain to policymakers and they have been shown to satisfy most of the axioms of a proper multidimensional poverty measure (Lasso de la Vega et al., 2009).