Evaluation and Program Planning 68 (2018) 47–56

Contents lists available at ScienceDirect

Evaluation and Program Planning

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

Evaluation of cash transfer programs in sub-Saharan Africa: A methodological review

Ebenezer Owusu-Addo[a,c,*], Andre M.N. Renzaho[a,b], Ben J. Smith[a]

[a] School of Public Hea lth and Preventive Medicine, Monash University, Australia
[b] School of Social Sciences and Psychology, Western Sydney University, Australia
[c] Bureau of Integrated Rural Development, Kwame Nkrumah University of Science and Technology, Ghana

ARTICLE INFO

Keywords:
Cash transfer programs
Complex interventions
Methodological review
Realist evaluation
Sub-Saharan Africa

ABSTRACT

Cash transfer programs (CTs) have been rigorously evaluated since their inception in the 1990s. However, to date, there has been no study critically examining the utility of the methodological approaches used to evaluate CTs. This article reviews the approaches used to evaluate CTs in sub-Saharan Africa (SSA) to provide recommendations for improving future CTs evaluations.

We conducted searches for CTs evaluation studies in SSA in the peer-reviewed and grey literature using electronic databases, hand searching of selected journals, organisational websites, Google Scholar and Scirus Internet search engines.

The review included 53 evaluation studies which were largely outcome-focused evaluations (95%; n = 50). The studies were undertaken within 24 CT programs comprising 11 unconditional CTs (UCTs), eight conditional CTs (CCTs) and five combined UCTs and CCTs. The review finds that while there is evidence of CTs impacts on a broad range of outcomes, the current evaluation approaches have primarily been experimental designs and have largely failed to provide explanations for mechanisms of change. To improve CTs policy and practice, there is the need to consider theory-based evaluation approaches such as realist evaluation that provide insights about the contexts and mechanisms through which programs generate outcomes in different circumstances.

1. Introduction

The Sustainable Development Goals (SDGs) identify poverty and inequality as critical barriers to improving health among socially disadvantaged groups across the world (UN, 2016). To this end, the first priority policy action, SDG 1.3 is to 'implement nationally appropriate social protection systems and measures for all and by 2030 achieve substantial coverage of the poor and vulnerable'. A particularly promising social protection scheme that could help in the fight against poverty and inequality is cash transfer programs (CTs). A recent study by the World Bank (2015) revealed that globally, as at 2014, there were about 194 cash transfer programs (CTs) with exponential growth found in sub-Saharan Africa (SSA). CTs are non-contributory safety net programs that give cash grants to poor households and vulnerable groups to satisfy basic consumption needs. They are classified into two: conditional cash transfers (CCTs) and unconditional cash transfers (UCTs). CCTs transfer money to households and/or individuals on conditions that beneficiaries adopt certain 'healthy' behaviours including school enrolment and attendance, child growth monitoring, utilisation of health services or sexually transmitted infection (STI) disease prevention. UCTs, similarly, provide money transfer, but do not have any explicit conditions (Baird, Ferreira, Özler, & Woolcock, 2013; Fiszbein & Schady, 2009). Generally, CTs have three core objectives: (a) reduce short-term poverty; (b) reduce long-term poverty by improving the accumulation of human capital; and (c) reduce food insecurity and other vulnerabilities (e.g. HIV/AIDs, orphans and vulnerable children crisis). The benefits of CTs as anti-poverty interventions have been widely reported in the literature (see Bastagli, Hagen-Zanker, Harman, Barca, & Sturge, 2016; DFID, 2011; ICAI, 2017; ODI, 2015, 2016). It has been estimated that in 2014, 718 million people were reached by CT programs globally (World Bank, 2015).

In SSA, CTs have become a mainstay social policy instrument for poverty reduction and for tackling a wide range of vulnerabilities. A key feature of CT programs in the region is their on-going rigorous evaluation (Davis, Gaarder, Handa, & Yablonski, 2012; Davis, Handa, Hypher, Rossi, & Winters, 2016). The benefits of program evaluation are enormous including, but not limited to, finding out what works for whom and under what circumstances, improving program delivery, accountability, evidence-based practice and increasing the likelihood of future funding, and hence sustainability (Pawson, 2013; Smith, 2011).

* Corresponding author.
E-mail addresses: ebenezer.owusu-addo@monash.edu, eowusu-addo.canr@knust.edu (E. Owusu-Addo).

https://doi.org/10.1016/j.evalprogplan.2018.02.010
Received 1 August 2017; Received in revised form 19 December 2017; Accepted 11 February 2018
Available online 14 February 2018

To realise these benefits, program evaluators need to pay particular attention to the decisions they make in relation to evaluation designs and methods, and the implications of these for gathering evidence that is meaningful and beneficial for both policy and practice.

Various approaches have been used in evaluating CTs, but to date, these methods and their robustness are poorly documented and assessed, especially their ability to address the complexities surrounding CTs and provide insights into how these bring about various impacts. This article therefore, reviews the methodological approaches used to evaluate the impact of CTs in SSA, and provides recommendations for improving future evaluations of CT programs. To do this, the article first discusses the complexities surrounding CTs and the implications of this for their evaluation. This review aims to offer insights for CTs policymakers, program managers, and evaluators with regard to the appropriateness of the current approaches being used in gathering evidence, and their ability to answer essential questions concerning what works, for whom and in what circumstances.

2. Literature review

2.1. Evaluating CTs as complex programs

Three categories of interventions have been identified in the literature namely: simple, complicated and complex interventions (see Campbell, Murray, Darbyshire, Emery, & Farmer, 2007; Ling, 2012; Rogers, 2008). Simple interventions are discrete and standardised in nature, and can be manipulated under ideal conditions to produce the desired outcomes (Ling, 2012) while complicated interventions are linear in nature with interrelated but non-interacting components that are expected to function in a predictable way with outputs leading to desired and anticipated outcomes (Glouberman & Zimmerman, 2002; Rogers, 2011).

While complex interventions share some of the traits of complicated interventions such as multiple components, these are characterised by interdependency, the role of human agency, a non-linear interaction between intervention components, and adaptation to changing conditions (Bamberger, Vaessen, & Raimondo, 2016; Campbell et al., 2007). Complexities in interventions are determined by a broad range of factors including: the context within which the program is designed and implemented; the specific components of the intervention; the interplay of program beneficiaries and institutions, structure and agency; the possibility of determining anticipated outcomes; and prior knowledge of factors affecting success or failure (Bamberger et al., 2016; Pawson & Tilley, 1997).

As complex interventions, CTs are heavily context dependent because their impacts are influenced by factors such as the policy environment, socio-economic conditions, organisational readiness, availability of complementary and supply-side services, and the behaviour of the target beneficiaries. Furthermore, CTs tend to include a broad mix of components, and may therefore achieve varied outcomes both intended and unintended, in different contexts. These complexities have implications for the methodologies and methods that may be used in their evaluation. The implications of complexity for program evaluation have been well discussed in the literature including the use of appropriate evaluation designs and methods, outcome patterns assessment, and the role of contexts in shaping outcomes (see e.g., Bamberger et al., 2016; Barnes, Matka, & Sullivan, 2003; Byrne, 2013; Callaghan, 2008; Ling, 2012; Westhorp, 2012). Because of the inherent complexities within CT programs, it is important to identify methodological approaches that can elucidate the causal process by which changes and impacts are achieved, and to address questions relating to program workings which are of relevance to policy and practice.

2.2. Approaches to evaluating complex programs

A recent study commissioned by the Department for International Development (DFID) identified six approaches to impact evaluation: experimental, statistical, theory-based, case-based, participatory, and synthesis studies (Stern, Stame, Mayne, Forss, & Davies, 2012, p. 24). Within this broad range of approaches, randomised controlled trial (RCT) designs have been deemed as the 'gold standard' for gaining an unbiased estimate of program effects (Rychetnik, Frommer, Hawe, & Shiell, 2002). While some have proposed that RCTs can be used to evaluate complex programs if the right measures are instituted (see Medical Research Council, 2008; White, 2013), it has been widely argued that RCTs are largely suitable for evaluating simple or complicated (but not complex) programs (Forss, Marra, & Schwartz, 2011; Ling, 2012; Pawson, 2013; Zimmerman, Dubois, Houle, Lloyd, & Mercier, 2012). A major limitation is that RCTs have difficulties accounting for how intervention components interact with each other and with complex open systems in unpredictable ways. Quasi-experimental designs, which are often adopted instead of RCTs for impact evaluation (Khandker, Koolwal, & Samad, 2010; White, 2013) are also greatly limited in their ability to answer questions about *how* and *why* programs work or fail to work within particular contexts (Marchal, Westhorp, Wong, Van Belle, & Greenhalgh, 2013; Pawson, 2013; Pawson & Manzano-Santaella, 2012; Pawson & Tilley, 1997). That is, experimental designs are limited in unpacking the 'black box' of complex interventions to explain how programs' outcomes are produced. It has been argued that it is necessary to not only determine what outcomes a program achieves, but to also understand the mechanisms that link causes and their effects, in order to inform program design and scalability (Chen, 2005; Pawson, 2013; Pawson & Manzano-Santaella, 2012).

Among the array of impact evaluation approaches, many evaluation theorists and writers (e.g., Bickman, 1987; Chen, 1990, 2005; Chen & Rossi, 1980; Donaldson, 2007; Pawson & Tilley, 1997; Rogers, 2008; Suchman, 1967; Weiss, 1995) have called for the use of theory-based evaluation approaches to evaluate complex programs. Theory-based evaluation is built upon intervention theory, and aims to identify and articulate this theory, to test it and/or improve it. As Birckmayer and Weiss (2000, p. 407) succinctly put it, "theory-based evaluation explores the how and why of program success or failure". Leeuw (2012) identifies two critical components of a theory-based evaluation namely: conceptual and empirical. The conceptual aspect focuses on identifying program theories while the empirical focuses on testing these theories to elucidate the process and mechanisms by which programs produce their outcomes. Among the array of theory-based evaluation approaches are realist evaluation (Pawson & Tilley, 1997), theory of change (Chen, 1990; Weiss, 1995), logic models (Bickman, 1987; Funnel & Rogers, 2011; Rogers, Petroscino, Huebner, & Hacsi, 2000), logical frameworks (USAID, 1980), and outcome hierarchies (Owen & Rogers, 1989; Owen 2006). Of these approaches, realist evaluation has been argued as providing a holistic approach to evaluating complex programs (Bamberger et al., 2016) as it focuses on program context, mechanism and outcome pattern configurations. A basic tenet of realist evaluation is that programs are complex interventions operating within open social systems, and therefore, require an understanding of their nature and their mechanisms of change (Pawson, 2013; Pawson & Tilley, 1997).

3. Methods

3.1. Selection criteria for inclusion of evaluations

For the purpose of this review, *evaluation* is defined as "the use of social research methods to systematically investigate the effectiveness of social intervention programs in ways that are adapted to their political and organisational environments and are designed to inform social action to improve social conditions" (Rossi, Lipsey, & Freeman, 2004, p. 16). This review includes evaluation studies that have examined or explored the impact of CTs in SSA. Studies were included if they