



THE GENDER TRANSFORMATIVE POTENTIAL OF GRADUATION PROGRAMS

SONIA LASZLO
NOVEMBER 2019

This work was contracted by Fundación Capital under the project “Towards a Gender Transformative Graduation Approach”. Fundación Capital is carrying out this project with the aid of a grant from the International Development Research Centre, Ottawa, Canada. The views expressed herein do not necessarily represent those of IDRC or its Board of Governors.



IDRC | CRDI

International Development Research Centre
Centre de recherches pour le développement international

Canada

Acknowledgements

This paper was prepared for Fundación Capital and has benefitted from discussions with Chris Boyd, Wendy Chamberlain, Frank DeGiovanni, Syed Hashemi, Barbara Jackson, Tatiana Rincon, Carine Roenen, Jo Sanson, Dristy Shrestha, Jaya Tiwari, as well as participants at the “Workshop: towards a Gender Transformative Graduation Approach” held at IDRC in Ottawa on September 30th, 2019.

Abstract

By combining asset transfers with short-term consumption support, access to savings, training and regular coaching or mentoring, Graduation Programs aim to lift poor and ultra-poor populations out of poverty. Recent impact assessments have found strong economic effects on program beneficiaries. Although they typically target women, the evidence is less clear about the degree to which Graduation Programs have been gender transformative. That is, while we know that women's *economic* outcomes have generally improved, it is less clear whether the programs have been able to empower them in other domains. After providing a brief conceptual framework to explain the mechanisms through which such multi-faceted programs could be gender transformative, this report assesses the latest state-of-the-art research on Graduation Programs' effects on women's non-economic outcomes. The review finds that while the quantitative evidence suggests positive but weak effects on non-economic domains (autonomy, agency, political participation and mental health), the qualitative literature provides more nuanced evidence suggesting much stronger impacts. Specifically, three key elements seem to make a difference: regular and frequent life-skills coaching (especially when they challenge gendered roles in household decisions), self-help groups or other group based activities (allowing women a safe-space to build social capital and take on community roles), and involving and sensitizing men and boys from the beginning of the program. Meanwhile, the potential for these programs to be truly transformational is greatly hindered by the lack of childcare options as women have to juggle caring for children with added responsibilities from the program.

Executive Summary

Globally, Graduation Programs are multi-faceted programs aimed at reducing poverty by providing poor or ultra-poor households some combination of asset transfers, temporary consumption support, savings and loans access, training and coaching or mentoring services. Most programs target women either disproportionately or exclusively, recognizing the over-representation of women among the ultra-poor and the fact that supporting women yields a double dividend as they also disproportionately care for the health and wellbeing of their children. Recent evidence from different programs point to substantial improvements in living standards among beneficiaries (Banerjee et al., 2015). While Graduation Programs have been shown to increase women's *economic* outcomes, the evidence of their effects on women's empowerment more broadly is both scarcer and mixed.

This study aims to unpack the role that Graduation Programs can play in promoting gender transformative change and yield returns to the non-economic dimensions of women's empowerment. To do so, we begin with setting the conceptual framework by linking the concepts of gender transformative change to women's empowerment and the ways that Graduation Programs can target the different mechanisms that underlie these concepts. Beyond simply increasing beneficiaries' access to productive resources, empowering women often also means challenging deeply entrenched gendered norms and improving women's agency.

We then review the recent evidence of the impact that Graduation Programs have had on women's non-economic outcomes (e.g. agency and decision-making, mental health and wellbeing, political participation and the quality of spousal relationships). We review both quantitative and qualitative papers written and published in top peer-reviewed journals in several disciplines and supplement this evidence with high-quality research working papers and technical reports as well as information provided by Graduation Program practitioners around organizational policies and practices.

Lessons learned and best practices are summarized according to different stages of Graduation Programs. First, we explore the importance of pre-existing conditions and whether they are determining factors in predicting whether Graduation Programs succeed in improving women's empowerment. Second, we discuss matters that may arise in the targeting phase that facilitate or impede the potential for gender transformative change. Third, we examine program design features and elements that have the most potential for change: the coaching or mentoring component, the creation of women's groups (e.g. savings and loans groups or self-help groups) and the importance of involving men and boys (especially husbands) early on. We also identify a significant gap in program design: graduation programs typically do not include any early child care option, such that the economic elements of the program may in fact be making juggling child care with livelihoods even more difficult. Fourth,

we discuss matters around program implementation and identify the importance of gender-sensitizing local staff and establishing trust between beneficiaries and the implementing partner. Fifth, we touch on the limits of measurement, recognizing that measuring the non-economic dimensions of women's empowerment is both challenging and problematic. Finally, we consider some of the potentially negative unintended effects that have been known to occur in similar or single faceted programs.

Based on the evidence provided, this report provides the following messages and policy recommendations for multi-faceted anti-poverty programs if they aim to increase the probability their efforts will yield gender transformative change and empower women:

1. Preconditions matter. The review points to important roles that pre-existing local infrastructure, social norms, physical environment and socio-economic characteristics play in facilitating successful programs. In addition, a number of studies find that the pre-existing quality of the spousal relationship is documented in the qualitative literature as being an important ingredient for success. This literature has nevertheless suggested that a well-designed and carefully implemented coaching/mentoring program that is regular and frequent can make a large difference for women who start the program in particularly low-quality relationships.
2. Targeting efforts should be especially astute in the intersectionality of their target populations, as women face more than just their gender as a source of potential vulnerability: the literature has found that religious minorities and disabled women are especially vulnerable and may even self-exclude from participating. Special efforts should be undertaken to include women from especially marginalized populations.
3. Coaching programs have been shown to make significant gains to non-economic outcomes, particularly when they include regular and frequent visits by program staff and address challenging traditional gender roles and social norms, among other topics such as health, asset training, and financial training. In addition, regular visits help program staff monitor the effectiveness of the other program components. This component has the most potential for gender transformative change when it starts early in the program and includes male household members (especially husbands). While most programs do include some general life-coaching or soft-skills training, including a gender transformative curriculum is increasingly being integrated into Graduation Programs. The major concern with this component is its' expense, and more cost-effective ways should be explored.
4. Women's Groups. Both the quantitative and qualitative evidence point to the hugely transformative role that women's groups (e.g. Self-Help Groups,

savings and loans groups) play in empowering women in both economic and non-economic domains. In addition to providing a safe space for women to discuss personal and productive matters, they can play an important role in increasing social capital and decreasing hostilities between beneficiaries and non-beneficiaries.

5. Engaging with men (especially husbands) is of critical importance to maximize the change for positive gender transformative change. This review identifies coaching as one of the most natural component in which this can be integrated and early engagement is optimal.
6. Women still bear the lion's share in caring for children, the elderly and disabled, which will inhibit their ability to fully benefit from Graduation Programs. In the absence of child care options, programs should work with husbands to have them share the load child care and domestic chores. Helping improve access to affordable and safe formal child care would be especially beneficial for gender transformational potential. However, implementing formal child care within the programs may not always be feasible, especially considering the substantial costs involved. Seeking opportunities to engage and coordinate with local providers, policy makers and educational facilities may be especially fruitful.
7. Implementation considerations. Sensitizing local staff to be sensitive to gender and its intersectionalities should be prioritized early on. Some Graduation Programs have already implemented useful guidelines to this respect that can be adapted to different contexts.

The report also identifies two key areas for which additional research is desirable, along with a reconciliation of the quantitative and qualitative literatures:

1. Unpack the different components and quantify the multiplier effect of coaching and women's groups. The qualitative evidence of the importance of these components is striking. However, the quantitative evidence is hindered by identification issues in conducting impact assessments on bundled programs. Yet a better understanding of the contributions of the individual components versus the bundle and their magnitudes will be important as policymakers and donors are looking to scale up these programs, especially given the costs of certain components (e.g. coaching).
2. The qualitative literature and technical reports have identified intersectionality and pre-existing conditions as influential for program success. The quantitative evidence ought to push on investigating heterogeneous impacts of programs by considering the quality of the spousal relationship, the role of multiple sources of marginalization, local infrastructure and physical environment, and differential effects by socio-economic characteristics.

Table of Contents

Executive Summary	4
1. Introduction	8
2. Conceptual Framework	12
2.1. Gender Transformative Change and Women's Empowerment	12
2.2. Graduation Programs	14
3. A brief review of the evidence	18
3.1. Quantitative Studies	19
Autonomy/decision-making outcomes	19
Mental health and psychosocial outcomes	21
Political involvement	21
Summary	21
3.2. Qualitative Studies	22
4. Lessons learned/best practices	27
4.1 Pre-existing conditions	28
4.2 Targeting	29
4.3 Design	29
Coaching/mentoring	30
Self-Help Groups	31
Involving men and boys	32
Missing link: the care economy	32
Summary	33
4.4 Implementation	34
4.5 Measurement	35
4.6 Unintended effects	36
5. Conclusions	37
Appendix Table	39
References	40



1. Introduction

Since the establishment of the Millennium Development Goals in the 1990s, the international development community has increasingly relied on social protection programs to play an important role in reducing poverty and vulnerability around the world (Merrien, 2013). A clear indication of this heightened attention is the proliferation of anti-poverty policies and programs originating in the Global South such as Cash Transfer Programs (notably beginning in Mexico and Brazil in the 1990s) and Bangladesh's BRAC "Targeting the Ultra-Poor" programs.¹ Scholars and policy-makers might debate the optimal form an anti-poverty initiative should take, but there is general agreement on the need to consider such programs for redistributive (social justice) purposes and that the presence of market failures justifies an interventionist approach (Ravallion, 2013). A basic premise is that anti-poverty programs can break the vicious cycle between poverty and the inability to invest in physical and human capital, allowing for long term sustainable and inclusive prosperity.

Das et al. (2005) explain how cash transfers (one of many possible anti-poverty instruments) can help households reach their optimal level of investment in capital. Conceptually, some households underinvest in physical capital (e.g. assets) or human capital (e.g. health or education) – relative to the investment level that optimizes their well-being – a wedge that can be largely explained by market failures. For example, micro-finance programs have been promoted for their proposed solutions to market failures arising from asymmetric information (see Van Tassel (1999) for a theoretical view). Similarly, asset transfer programs provide the poor with much-needed seed

¹ BRAC (2013) defines the ultra-poverty lines at about 60 or 70 cents per day, considerably lower than the extreme poverty line set by the World Bank (1.25 \$/day in 2013, though they increased it to 1.90 \$/day in 2015)

capital which they are normally excluded from accessing due to their poverty status (Bardhan, 1996). Cash transfers, meanwhile, can help address another source of market failure: Intra-household bargaining dynamics lead to market failures and sub-optimal social outcomes if decision-makers do not fully internalize and take-into account the preferences of household members.

To see this, Cash transfer programs, especially those that target women and those with conditionalities attached to them, can resolve intra-household bargaining based market failures (Das et al., 2005). Conditioning on school enrolment, for example, leads to an increase in children's human capital beyond what a family might otherwise invest. And making women beneficiaries of the programs ought to afford them increased bargaining power and independence in household decision-making. Indeed, targeting women as beneficiaries is the mainstay of a vast array of social protection programs from cash transfers to micro-finance because of this double dividend. Anti-poverty programs targeting women not only improve women's outcomes but also improve children's outcomes since women are typically primary care-givers (UNICEF, 2007). Given this, it is not surprising that these programs feature prominently as a means to target – and indeed achieve – the Sustainable Development Goals (ILO, 2017).

While social protection programs do appear to yield encouraging results in the fight against poverty, the scholarly evidence is often mixed when it comes to meaningfully and sustainably improving women's lives. For example, while micro-finance programs have led to increased entrepreneurship, they often fail to reach the poorest and led to increased social pressures (and in the extreme have been linked to suicides – Ashta et al., 2015). Similarly, conditional cash transfer programs have been linked to increased marital dissolution (Bobonis, 2011) or have been found to reinforce traditional gender norms around childrearing and domestic work (Molyneux, 2007). And a large number of entrepreneurship or job training programs targeting women fail in the long term because of the lack of affordable and quality care of children, the elderly and disabled that otherwise disproportionately falls on women's shoulders. And this despite the recent focus on promoting Women's Economic Empowerment (WEE) as a central feature of the International Development debate and agenda.

A (relatively) new class of programs known as "Graduation Programs (GP)" is especially promising in this area, precisely because it combines multiple facets of social protection (IPCIG, 2017). Graduation Programs have varying components depending on the context, though most include some asset transfer, cash transfer for consumption support, skills training, health/education information, savings programs, and life-skills coaching/mentoring. The history, design features, and virtues of GPs are well described and laid out in Hashemi and Umaira (2011) and de Montesquiou et al. (2014) and so won't be repeated here. By design, these programs intend to break the vicious cycle of poverty and an inability to invest in physical and human capital in a multi-faceted way recognizing that poverty is multi-

dimensional and resulting from multiple causes. Indeed, the influential multi-country evaluation of GP programs in 6 different countries by Banerjee et al. (2015) echo what has been found in a multitude of other programs: beneficiaries have greater asset holdings, incomes, and access to savings and are more food secure and healthier than comparable non-beneficiaries. What's more, GPs are designed to be short term in nature (normally between 1 to 3 years), enough to break the cycle and setting beneficiaries on a sustainable trajectory, whereby reducing program dependency and the disincentive effects that are of considerable concern in anti-poverty policy design (Besley and Kanbur, 1993; Ravallion 2013).

The study by Banerjee et al. (2015), amongst others with similar findings, generated considerable interest, especially among GP implementers and scholars interested in the potential these programs have to meaningfully transform the lives of women. While most GP-type programs indeed target women and since beneficiaries report greater wealth, income, food security, then one might be confident in drawing the conclusion that the programs have empowered women. Banerjee et al. (2015) however find weak to no impact of the GPs in their study on widely used measures of Women's empowerment. However influential their study, the broader and to an extent more recent impact evaluation literature on GP's causal impact on Women's Empowerment provides mixed and nuanced results.

This paper sets out to review the recent literature of GPs' effects on women's outcomes beyond the typical array of economic outcomes (e.g. incomes, savings, asset holdings etc.) and to instead investigate what has been found in terms of the potential of GPs to be gender transformative. That is, to what extent have GPs been found to deal with the market failures originating from bargaining (power) dynamics both within and outside the household? To understand this potential, we must first understand whether and how the GPs breakdown the barriers and limitations to meaningful empowerment by addressing social norms, the burden of care and household work as well as potential unintended effects (such as male backlash). To do so, we appeal to a review of high quality quantitative and qualitative impact assessment studies, as well as from technical reports and other materials from implementing organizations, supplemented with information and feedback from GP practitioners.

This paper is neither a meta-analysis of GPs nor one of gender transformative approaches to social protection. This exercise is also not a handbook for selecting gender transformative tools to include in GPs. Instead, it aims to provide readers with a broad overview of the recent and rigorous evidence of the gender transformative effectiveness of GPs and examples of what has worked in some contexts and why. Taken together, the evidence provided here enables the reader to critically assess the evidence and how it may apply to different contexts. The paper is structured as follows. Section 2 provides a brief conceptual framework establishing what we mean by gender transformative change and Women's Empowerment and how

these concepts might map into Graduation Programs. The purpose of this section is to understand *the mechanisms (how)* Graduation Programs can translate into gender transformative change. Section 3 presents a brief review of key influential and rigorous studies of the impact of GPs on Women's empowerment. This section provides us with a reading of *whether* Graduation Programs have yielded positive effects on women's empowerment (beyond simply looking at women's economic outcomes). In section 4, we deconstruct the lessons and best practices according to different stages of program impact from pre-conditions to measurement. This section is intended to help shed light on *why* Graduation Programs might have failed or succeeded in promoting women's empowerment and gender transformative change. Section 5 concludes.



2. Conceptual Framework

2.1. Gender Transformative Change and Women's Empowerment

To understand how Graduation Programs can transformatively improve women's lives, it is useful to establish a common understanding of the meaning of gender transformative change. There are three useful definitions employed in the international development dialogue. The Population Council defines a gender transformative approach to mean *"that promoting gender equality – the shared control of resources and decision-making – and women's empowerment are central to an intervention"* (Population Council, 2019). Similarly, according to CARE USA, gender transformative approaches *"aim to go beyond individual self-improvement among women toward transforming power dynamics and structures that act to reinforce gendered inequalities"* (Hillenbrand et al., 2015, p. 10), implying the need to consider gender inequality and inequity in the household and community. They build on the work by the CGIAR which takes the gender transformative approach to go *"beyond just considering the symptoms of gender inequality, and addresses the social norms, attitudes, behaviors and social systems that underlie them"* (CGIAR, 2012, p. 2). At the core of this approach is the recognition that meaningfully promoting and improving Women's Empowerment will require challenging both the power dynamics within the household and community as well as the existing social norms around gender.

What does this mean for anti-poverty programs, such as Graduation Programs, to be gender transformative? These need to move beyond simply targeting women and girls and providing them with resources and skills. To be sure, targeting women and girls and redressing inequities with respect to resources and skills are important ingredients for promoting their empowerment. At best, these can be viewed as necessary conditions for gender transformative change. However, as the review will show, these ingredients will not lead to meaningful and long-term empowerment and improved well-being for women and girls if the root causes of these inequities are not addressed. To illustrate, enhancing a woman's ability to save will not translate into her ability to start and lead a successful business if her savings are appropriated by her husband or if she is expected to limit paid activity in order to care for her young children or elderly relatives.

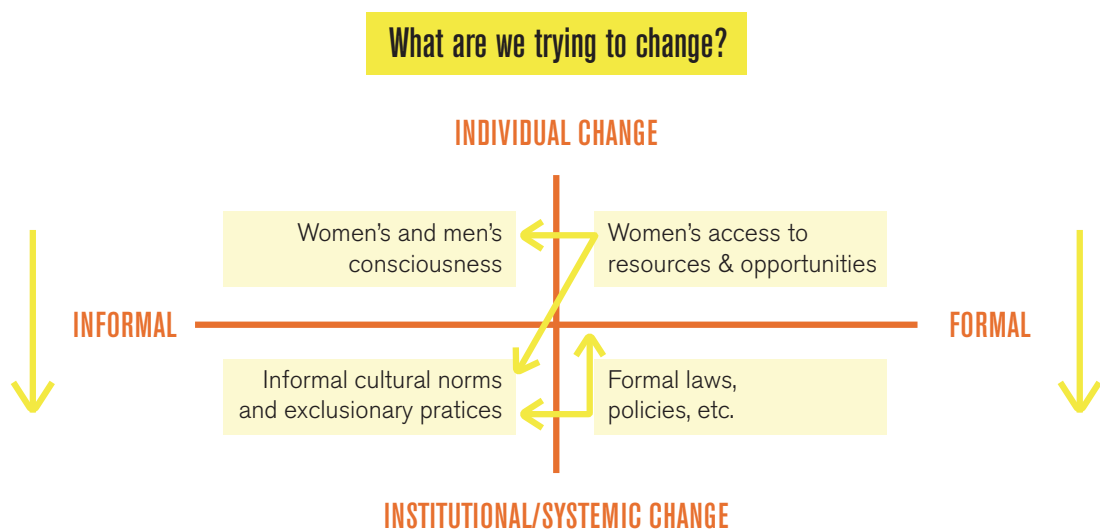
To be sure, the concept of gender transformative change is closely connected to the concept of women's empowerment. Kabeer's (1999) influential paper conceptualizes women's empowerment as connecting resources, agency and accomplishments. For a disempowered woman to become empowered, she needs to have access to resources, be able to act on and use them, and combined these must translate into accomplishments. There are two important implications of this conceptualization in the context of gender transformative change. The first is that we are dealing with a *process of change*, recognizing that there is both a sequence to empowerment or transformation and that it may take time to become fully realized. The second is that it is *multi-dimensional*: simply providing access to resources will not lead to meaningful change in a woman's life if she does not have the power to make decisions over the resource and so the potential for the enhanced access will not translate into improved well-being. Resources, agency and achievements must thus be taken as a tripartite in understanding and promoting empowerment and transformation.²

With these concepts, we can consider the ingredients of change as illustrated in Hillenbrand et al. (2015, p. 11), a slight modification of the influential framework in Rao and Kelleher (2005, p. 60), and reproduced in Figure 1. This figure is helpful to understand how social protection policies can be transformative, and also their possible limitations. Historically, most anti-poverty programs (e.g. cash transfers, microfinance and asset transfer and training) have targeted the top right quadrant by attempting to promote and improve women's access to opportunities and resources. To put it bluntly, this is the easy part. What is much harder to accomplish is changing the informal dimensions of change – namely “women's and men's consciousness” and the “informal cultural norms and exclusionary practices”. Recent social protection programs are increasingly incorporating attempts to make changes in these domains, recognizing that the failure to do so will severely limit the transformational

² Women's agency is itself a very multi-dimensional concept, and Kabeer's seminal (1999) piece discusses it in the context related to the “the ability to set goals and act on them”. Donald et al. (2017) provides an excellent discussion about how to measure agency in the context of women's empowerment in developing countries. Indicators that are often used to measure agency include autonomy in decision-making or bargaining (arguably one of the most common proxies in the current literature), psycho-social factors such as self-esteem, self-confidence, the locus of control, self-efficacy and aspirations, to name a few. See Donald et al. (2017) and Laszlo et al. (2017) for more on measurement.

potential of providing access to resources and opportunities. For example, in the absence of child care options, Roelen et al. (2019) describe how “evil eye” beliefs prevent women in Haiti from carrying their infant children on their backs rendering combining work and child care a physically impossible task. Similarly, providing an asset transfer to a woman will not be transformative if the asset is appropriated by her husband. On the bottom right hand quadrant of Figure 1 are “formal laws and policies”. The extent to which social protection programs can accomplish change in this domain will depend on the degree to which they are implemented by or in close collaboration with governments that have the political will to do so.

Figure 1



Source: Figure 1 in Hillenbrand et al. (2015, p. 11), a modification of Figure 1 in Rao and Kelleher (2005, p. 60).

2.2. Graduation Programs

The major innovation of Graduation Programs is that they combine multiple interventions, each of which in and of themselves have the potential to introduce change in one or more quadrant of Figure 1. As in Hashemi and Umaira (2011), de Montesquiou et al. (2014) and Banerjee et al. (2015), most Graduation Programs include some combination of the following interventions: asset transfer and training, cash transfer for consumption support, savings programs, health and education training, and regular coaching/mentoring. It is easy to see how some components deal directly with improving access to resources and opportunities (e.g. the asset or cash transfer, savings programs). Many training or coaching and mentoring programs target the informal domains in Figure 1, especially women's and men's consciousnesses. It is somewhat more challenging to see how these programs can change the domains in the lower two quadrants, unless they are brought to scale or are able to generate considerable spillovers and accompanied by government and policy cooperation and coordination.

In a recent review of social protection programs, Hanna and Karlan (2017) discuss Graduation Programs and stress the importance of understanding how each component separately addresses the problems they set out to change. In addition, they flag a number of areas for future research – such as deconstructing the impacts by component and they pay particular attention to both the potential and caveats of the life-coaching component:

“More work is needed to tease apart the different components: asset transfer (addresses capital market failures), savings accounts (lowers savings transactions fee), information (addresses information failures), life-coaching (addresses behavioural constraints, and perhaps changes expectations and beliefs about possible return on investment), health services and information (addresses health market failures), consumption support (addresses nutrition based poverty traps) etc. There will be no simple answer to the aforementioned queries, but further work can help isolate the conditions under which each of these components should be deemed necessary to address. And furthermore, for several of these questions, there are key open issues for how to address them; for example life-coaching can take on an infinite number of manifestations (...) Much remains to be learned, not just about the promise of such life-coaching components, but how to make them work (if they work at all)” (p. 540)

In addition to making explicit the problems to be solved by each component, this passage illustrates the role that life-coaching can play in addressing some of the informal dimensions of change from Figure 1 – namely around behavioural constraints. This is where the potential for changing attitudes, behaviours and norms can kick in. The caveats such as the infinite number of manifestations of the life-coaching component in the quote are salient and will be revisited in the review and discussion below. In contrast, while a cash transfer can vary in design by amount and by conditionalities, these are finite and easily comparable. There are programmatic and logistical reasons why coaching programs may vary far more. Depending on the implementing agency and the objectives of the donors/funders of the program, life-coaching could target a vast array of life-skills such as cognitive skills, business planning, life planning, redressing social and gender norms, to name but a few. Different coaching programs may include different elements. In addition, there are numerous ways in which these programs could deliver the training, and this may vary by context: in person coaching by mentors or peers, electronic mentoring through tablets, the frequency and duration of coaching can vary, the content may need to change depending on local culture, language and social norms, the identity of the mentor (male or female, local or not), and which household members (husbands, boys, other kin) are present during the coaching sessions. Each element can have

important effects on the effectiveness of coaching programs. Standardizing this component is problematic conceptually given the likely influence of local cultural and social norms, complicating the assessment of different coaching programs: the type of coaching that works in one population or sub-population might not work in another. Nevertheless, as the evidence reviewed and discussed below will show, this component plays an important role in ensuring the effectiveness of the other components, and thus may need to be tailored to the broader intervention for the specific context. In sum, it is thus be safe to say that the potential effectiveness of the bundled Graduation Program is greater than the sum of its parts. Put another way, the life-coaching component is well positioned to amplify impacts from the other components. In fact, this is why most graduation programs begin with a life-coaching component that follows households throughout the lifetime of the program.

Two papers are especially helpful in conceptualizing the reasons why GPs can be transformative. Jones et al. (2017) provide a discussion of how social protection programs in agriculture (including Graduation Programs) can be gender transformative. They stress the importance of incorporating a gender sensitive approach to programming, not only to target women but to design programs that meet women's needs, giving them voice and agency and ultimately to improve the synergies between program components (p. 94). In this light, taking one component on its own may fail to capture the multiple ways in which they influence women's wellbeing. For example, they illustrate this point by drawing attention to the fact that most consumption-smoothing interventions are designed with women as consumers and don't always consider women as producers – potentially understating the potential impact on household welfare of such interventions. Furthermore, in addition to their roles in both consumption and production, consumption-smoothing interventions can reduce anxiety, increase feelings of hope, happiness and life-satisfaction (p. 90). In this sense the intervention has an amplified effect as it insulates from negative consumption shocks, increases the productive potential and improves mental health. Conversely, they also provide examples of interventions that may not yield as strong an impact as desired: interventions aiming to increase women's control over household financial resources may in fact be limited in their transformative potential if consumption patterns remain gendered (p. 91)

Pritchard et al. (2015) discuss the factors enabling and constraining GPs. While their paper does not explicitly take a gender lens and focuses on a single GP (the Chars Livelihood Program in Bangladesh), it is easy to map this framework into one that is useful to understand and conceptualize how GPs can have a gender transformative potential. In terms of factors that constrain graduation, the authors point to policy and design constraints and to household-level constraints. In the former case, they suggest that for maximum graduation potential, the program needs to be designed (and log-framed) as a graduation-focused program and that indicators and thresholds must be set from the start (p. 42). This suggests that to maximize the gender transformative potential of Graduation Programs, gender

issues must be integrated from the start, including setting indicators and thresholds early in the implementation phase. In the latter case, they noted that a number of beneficiary households saw their asset values fall, which they attribute to (1) poor reinvestment of profits, (2) payment of dowries, (3) investment in low quality lands and (4) payment of loans (p. 43). Here, there is an explicit mention of gender-based norms inhibiting the potential of asset holdings – dowries – even though such payments are prohibited by law. Note also how this connects to the lower quadrants of Figure 1 in which *de jure* laws do not always translate into *de facto* practices, and how programs need to consider both formal and informal rules in order to understand the gender transformational potential and limitations of social protection programs.

For factors enabling graduation, Pritchard et al. (2015) consider environmental factors in addition to policy/design factors and household-level factors. For policy and design factors, the authors warn against being overly ambitious in setting outcome thresholds against which to gauge whether households have graduated out of poverty and against expecting large spillover effects of the broader impacts of graduation programs (p. 44). This will be especially important on the gender domain, considering the complex nature of the concepts involved (women's empowerment) and even greater concerns around its measurement (Laszlo et al., 2017). Environmental factors in Pritchard et al. (2015) are understood as the ability households have to reduce their vulnerability to physical environments concerns (floodplains in their case). In addition to the gendered dimension of physical environmental issues (such as climate change, see Oxfam (2002)), we may add the social environment in which beneficiaries live, namely as they relate to gender norms. Marcus (2018) discusses the need for a supportive policy and institutional environment to enable programs to meaningfully empower women – the availability of child care being one of many institutional factors that relax significant gendered constraints. Regarding household level factors, Pritchard et al. (2015) explicitly single out agency – a participant's ability to influence household decisions (regarding investments) – as an enabler, which has a clear relevance for women's empowerment.

To summarize the conceptual framework, *a priori* there is considerable potential for GPs to be gender transformative. While programs may target women explicitly and even exclusively, the degree to which the components (or bundle of components) will improve women's lives in a sustainable manner will largely depend on the degree to which these have been accompanied with changes in power relations and in the formal and informal structural constraints to women's empowerment. This section has reviewed some key conceptual issues highlighting the importance of relaxing and reversing unfavorable formal and informal norms and practices to maximize this potential. The next section reviews the state of the evidence on whether Graduation Programs have led to transformational change for women. In the penultimate section, we deconstruct where the weak points might be by looking at the lessons and best practices from the various stages of GPs: pre-conditions (socio-cultural and economic environment predating the introduction of the program), targeting, design, implementation and measurement.



3. A brief review of the evidence

We now turn our attention to the existing literature evaluating GPs' effects on women's empowerment, with an eye to evaluating the gender transformational impact of these programs. This review does not survey all impact assessments of GPs – most programs target women and most impact assessments evaluate the impacts on incomes, wealth, asset holdings, food security and a number of other economically relevant outcomes.³ And the evidence is overwhelmingly positive for these outcomes (see Banerjee et al. 2015 and related literature). This review recognizes that gender transformative change must look beyond such economic outcomes. Thus, the first part of this section will review impact assessments that look at women's (non-economic) empowerment outcomes, and this literature is mostly quantitative. The second part reviews a number of key qualitative studies that are able to uncover more nuanced impacts on women's empowerment outcomes. For both the quantitative and the qualitative literature, we restrict attention to impact assessments employing rigorous impact assessment techniques, published in peer reviewed journals or appearing in working paper series known for their scholarly rigour (e.g. the National Bureau of Economic Research or the World Bank). Where appropriate, we supplement the discussion with technical papers and information about these programs provided by implementers of GP practitioners.

³ Sulaiman (2016), for example, provides a meta-analysis of the effects of GPs in comparison to simple asset transfer or cash transfer programs.

3.1. Quantitative Studies

This research has identified 12 papers that have attempted to quantitatively assess the impact of Graduation Programs on non-economic outcomes among women.⁴ For the study to be included in this section, it needed to provide sufficient methodological detail to ascertain that it meets a minimum level of econometric/statistical rigour (e.g. Randomized Controlled Trial, Difference-in-difference estimation, Propensity Score matching). In fact, most these studies in this quantitative literature use a Randomized Controlled Trial design. Of the 12 papers, 8 are in peer review journals in scientific journals, the remaining 4 are either part of a working paper series, reports, or unpublished university working paper (World Bank, Centre for the Study of African Economies, National Bureau of Economic Research, IDS Research Report). The peer-reviewed papers are published in top general interest (e.g. Science) and top discipline journals (e.g. Quarterly Journal of Economics, Social Science and Medicine), suggesting the scholarly interest in these programs and the potential scale of readership. These 12 papers evaluate 14 different projects across Asia (Afghanistan, Bangladesh, India, Pakistan), Africa (Burundi, Burkina Faso, Ethiopia, Ghana, Kenya, Uganda) and Latin America and the Caribbean (Honduras, Paraguay, Peru). While the literature evaluating GPs is relatively recent (most papers have been published since the mid-2000s), it is worth noting that the literature evaluating impacts on women beyond their immediate economic outcomes is considerably more recent. The Appendix Table gathers information from these 12 studies, summarizing the different facets of the graduation program being evaluated, indicating (if reported in the paper) the key component of the life coaching/monitoring, the non-economic outcome measure and the estimated impact of these measures.⁵ The first main observation is that the measures used to capture non-economic outcomes varies drastically across studies.

Autonomy/decision-making outcomes

The 6 projects evaluated by Banerjee et al. (2015) and the BRAC/TUP evaluation by Bandiera et al. (2017) use the same set of outcome measures, commonly employed in the Women's Empowerment literature. The most direct one is an index of women's decision-making, which typically measures to what extent women share important household-level decisions such as major household purchases, investment in children's education and health, and autonomy over how to spend her own income or over visits to friends and family. These measures are often used as proxies for autonomy, agency and bargaining power. None of these 7 projects show any strongly statistically significant long-term impact on these index measures.⁶ However, the Banerjee et al. (2018) follow up study for the Ghana site studied in the Banerjee et

⁴ One study (Ismayilova et al., 2018a) considers child outcomes (boys as well as girls), not women's outcomes explicitly. It is however included in this study because it is connected to their other paper (Ismayilova et al., 2018b) and the project under evaluation includes life-coaching aimed at changing norms and behaviours around violence and teenage pregnancy).

⁵ This table does not report the estimated impact on economic outcomes.

⁶ Evaluating an earlier phase of the BRAC's Challenging the Frontiers of Poverty Reduction – Targeting the Ultra Poor – Emran et al. (2014) use the ratio of saris (female clothing) to lungis (male clothing) as an indicator of the balance of power in expenditure decisions, although recognize the limitations in using such indicators to capture either women's welfare or empowerment. They do not observe any statistically significant impact of the program on this ratio.

al. (2015) does find a weakly positive impact on their women's decision-making index after 3 years, for a treatment arm in which beneficiaries receive the full graduation package. For the other treatment arms in their study, in which beneficiaries receive either the full package minus the savings component, savings only or asset transfer only, the results on female empowerment are statistically insignificant.

Meanwhile, several other papers also include similar decision-making or autonomy measures and the results are more nuanced. While Blattman et al. (2016) find no statistically significant impact on women's autonomy over purchases in the case of the WINGS project in Uganda, Ismayilova et al. (2018b) find a positive impact of Trickle-Up's Burkina Faso program on women's financial autonomy and Bedoya et al. (2019)'s evaluation of Afghanistan's TUP finds a positive effect on women's empowerment but a statistically insignificant effect on her role in household expenditure decisions.⁷ The Mahecha et al. (2018) evaluation of the Paraguayan *Sembrando Oportunidades Familia por Familia (SOF)*, a graduation program implemented by Fundación Capital and the Paraguayan Government, considers three indices to capture women's empowerment: an index of empowerment within the community, an index of autonomy and independence in household decision-making, and an index of perceived gender roles within the household. Their results relative to impacts on the first two indices are sensitive to the estimation method, though their results pertaining to perceived gender roles are statistically significant across methods.⁸ It is worth noting that *SOF* explicitly addressed empowerment and self-esteem in their life-skills coaching – though the bundled nature of the program does not allow an assessment of whether the measured impacts on gender roles is causally driven by the coaching. Mahecha et al. (2018).

The study by Devereux et al. (2015) of Concern Worldwide's program in Burundi and the Roy et al. (2015) study of BRAC's Specially Targeted Ultra-Poor program in Bangladesh also find nuanced effects of graduation programs on the different domains of decision-making and autonomy. Both studies found that decision-making or control over resources shifted towards women in some dimensions but that in other dimensions, women actually lost control or had to now share control with her husband.⁹ Roy et al. (2015) in fact find that women's ownership, decision-making and control over the transferred asset improved, yet they document a loss of decision-making power over other assets to their spouse. They also document that the nature of the transferred asset – livestock – requires the beneficiary to stay close to the homestead consequently reducing her mobility out of the home (including to seek income generating activities) and increasing her workload. While this appears in stark contrast to many notions of empowerment, their qualitative work in Roy et al. (2015) and Das et al. (2013) suggests that the beneficiaries actually preferred to stay close to their homestead as it was viewed as preferable to the outside low-quality, low-pay and high-stigma employment option.¹⁰

7 In fact, these reported results are largely based on indices – the effects of individual components of the empowerment indices show more nuanced results, and I defer the reader to consult their paper for more details as they are too numerous to list here.

8 Though initially designed as an RCT, the implementation could not follow the strict protocol to ensure internal validity.

9 These ambiguous effects may be one reason for which indices fare poorly in measuring program impacts: if the index components show effects of opposite signs, they could potentially cancel themselves out in an index.

10 This preference might be partly explained by local religious or social norms that stigmatize poor rural women from participating in the local economy, imposing important reputational costs that can be avoided by staying in the homestead (Roy et al., 2015, p. 14). This result might be different in different contexts where those reputational costs are lower or non-existent.

Mental health and psychosocial outcomes

Most of the quantitative studies listed here also include some measure of mental or psychosocial wellbeing, which have been linked to women's empowerment through its relationship with agency (e.g. Donald et al., 2017). Only 2 of the 6 programs evaluated in the Banerjee et al. (2015) study show positive and significant effects on an index of mental health (which comprises of indicators of stress, happiness and life satisfaction): Ariwara (Peru) and Proyecto MIRE (Honduras). These results are strong enough to remain statistically significant and positive when the authors pool the 6 study sites, though they acknowledge that the results on mental health dampen between their first and second endline surveys. The Bandiera et al. (2017) paper similarly finds strongly statistically significant effects in Bangladesh for their mental health index (a combination of self-reported happiness and mental anxiety measures), four years after the transfer. In post-conflict setting Afghanistan, Bedoya et al. (2019) show very strong effects on psychological wellbeing – measured as an index of indicators on self-reported happiness, stress, depression, self-esteem and cortisol, adapted to the local social norms –considerably more so for women than for men. The results hold even when analyzing each component of the index separately. Ismayilova et al. (2018a) similarly show important reductions in stress and depression and an increase in self-esteem among children in beneficiary households in Trickle-Up's programme in Burkina Faso. Mahecha et al. (2018) find positive effects of the Paraguay SOF program on both aspirations and expectations. As most of these papers acknowledge, it is impossible to disentangle what mechanism drives these results – part of the problem, as above, is the inability to isolate the effects of the separate components of the graduation program.

Political involvement

Women's empowerment should also reflect their ability to take part in and influence community decisions (Kabeer, 1999). Linking back to Figure 1, increased political involvement is one way to manifest gender transformative change. If a program can empower women to better advocate for themselves in the political domain, the more likely formal pro-gender equality laws and policies can be brought in. Again, the Banerjee et al. (2015), Bandiera et al. (2017) and Bedoya (2019) studies systematically evaluate the impact of graduation programs on an index of political involvement. In all but the Peru and Honduras study sites evaluated in these three papers, the effects on political involvement are statistically significant and positive.

Summary

The recent quantitative literature evaluating the impact of graduation programs on the non-economic outcomes of women's empowerment and gender transformative change generally show either statistically significantly positive or insignificant results. The lack of consensus in these studies could be due to measurement issues – non-economic factors are very difficult to quantify and obtaining comparable indicators in different settings is problematic given different socio-cultural contexts. Understanding this lack of consensus is also complicated by our inability to isolate

which component of the evaluated programs are transformative or whether it is the bundle as a whole that is transformative.¹¹ Still lacking in this area is a clean test of the mechanisms laid out in section 2. Finally, given that some components (such as the life-coaching/mentoring component) vary considerably across programs in their application and curriculum, it is difficult to know to what extent one model works in one context and not another.

Blattman et al. (2016) attempt to unpack the apparent weak results on non-economic empowerment related outcomes, despite strong economic impacts of the Uganda WINGS program. In fact, they identify somewhat contradictory results: while husbands seemingly encouraged participation in the program, they were also likely to appropriate their wives' resulting increased income. Similarly, while spouses were more likely to exert control over their wives, the women reported somewhat higher quality in their relationship with their partner. Unlike most other graduation programs, the WINGS program in Uganda features only three components: a cash transfer, five days of business training, and on-going supervision.¹² Their experiment randomized the amount of supervision (0 visits, 2 visits, 5 visits plus advice), allowing them to tease out the effect of the supervision component from the others. They find that supervision has little impact on business income or consumption in the long term, though they do find that businesses which received more supervision were more likely to survive at follow-up. The training sessions included group dynamics training to encourage the endogenous formation of self-help groups, which appear to influence beneficiaries' relationships with their neighbours: hostility increased among beneficiaries that did not participate in a self-help group. Social capital appears to matter – highlighting the importance of community level factors in the empowerment process. Similarly, Devereux et al. (2015) evaluate a randomized

control trial in Burundi's program in which they vary the intensity of home visits. One treatment group receives 1 visit a month and the other treatment group receives 3 visits per month. They did not however find any significant differences between the two groups in either economic or non-economic outcomes

3.2. Qualitative Studies

Qualitative studies on the impacts of GPs tend to delve into mechanisms more than their quantitative counterparts. This subsection reviews 7 studies covering (1) Bangladesh's Challenging the Frontiers of Poverty Reduction (CFPR), BRAC's early Graduation Program (Ahmed et al., 2009; Roy et al., 2015), (2) India's Swayam Krishi Sangam (SKS) Program (Goodwin et al., 2018), (3) Trickle-Up's programs (targeting disabled persons) in Guatemala, Nicaragua and Mexico (Sanson et al., 2018), (4) Kabeer's (1999) study of the Programs in Pakistan and in India, (5) Haiti's Fonkoze's program in Roelen et al. (2019) and (6) Concern Worldwide's program in Burundi (Devereux et al., 2015).¹³

¹¹ Although some papers do attempt to isolate effects of certain components, such as Banerjee et al. (2018) and Sedlmayr et al. (2018).

¹² They note the considerable disproportionate expense of supervision, echoing Pritchett's (2018) concern that graduation programs may not be cost-effective. Blattman et al. (2016) do raise the importance of finding more cost-effective means to deliver the supervision component.

¹³ Both the Roy et al. (2015) and Devereux et al. (2015) include both quantitative and qualitative analysis and the qualitative results were discussed in section 3.1. in interpreting the quantitative results. Das et al. (2013) IFPRI working paper is not counted as a separate study in this paper, but it provides a lot of the details of the qualitative analysis presented in Roy et al. (2015).

Ahmed et al. (2009) conduct one of the first quantitative and qualitative studies of BRAC's CFPR program. While they find considerable improvements on economic outcomes (asset holdings, food consumption, savings behaviours and dwelling improvements), their quantitative study does not address women's non-economic outcomes. Instead, they appeal to qualitative evidence drawn from 5 case-studies to delve into why the program worked or not for beneficiaries. Pointing out that quantitative results speak to *average* effects, their qualitative results speak to differentiated effects and they propose reasons for success or failure. For example, they interview one woman who despite reporting an increase in the number of hours spent doing chores, reported a greater sense of security and stability and made her more confident. This case shows a nuanced result on empowerment – more hours spent on chores may reflect a reinforcement of gender roles within the household, yet the increased sense of security and confidence suggest an improvement in mental health and the potential for increased agency. They also document two examples of beneficiaries who were less successful, and the reasons provided shed considerable light on the process and mechanisms. In the first case, failure was in part attributed to a poor relationship with the spouse and with the program officer. In the second case, the authors attribute business failure to the beneficiary's inability to overcome social constraints within her community. Recalling the conceptual framework from section 2 and Figure 1 (from Hildebrand et al. (2015) and Rao and Kelleher (2005)), failure of the program in these cases was despite the increased access to resources – it had more to do with social considerations than economic ones.

Goodwin et al. (2018) revisit the Indian SKS program that was originally evaluated by RCT in Bauchet et al. (2015). The Bauchet et al. (2015) study failed to find any statistically significant effect of the graduation program on economic outcomes and provided 4 broad explanations for failure to estimate impact (data issues, design and implementation issues, low program take-up or high drop-out, and shifting from wage to self-employment). They argue, furthermore, that these programs work best in thin labour markets, in contrast to Andhra Pradesh's strong market in their setting. While Goodwin et al. (2018) critique the Bauchet et al. (2015) study on the basis of "narrow economic indicators" used in this and other RCT studies, their qualitative *social inclusion* study yields different conclusions. Based on semi-structured interviews with 15 graduates of SKS, they document narratives suggesting that asset ownership was in fact transformative and that the program lead to increased feelings of agency and control over decision-making; they find that the asset allowed women to play an active economic role without male support (p. 129). They point out that a major epistemological difference between RCT-type quantitative studies and qualitative studies lies in *who* determines success of the project – the evaluators in the RCT case or the beneficiary in the social inclusion study case (p. 134). It is worthwhile pointing out that, unlike the papers reviewed in section 3.1, the Bauchet et al. (2015) study did not include non-economic outcome indicators.

Similarly, Kabeer (2019) revisits two settings in which programs that were evaluated using RCT – one in Pakistan (the Pakistan Poverty Alleviation Fund) and one in West Bengal, India (by Bandhan) that were reported in Banerjee et al. (2015) – and found to have statistically significant and positive impacts on economic outcomes but moderate if any significant impact on measures of women's empowerment.¹⁴ Her qualitative study fills a number of gaps in our understanding of the causal processes underlying graduation programs in three critical ways, and suggests that the RCT studies were not conducted in a way to be able to flesh out some of the mechanisms. The first relates to the quality of the relationship with the spouse. Women who were already in a co-operative relationship with their husbands tended to do quite well, echoing the qualitative evidence described above by Ahmed et al. (2009).¹⁵ While this suggests that preconditions matter, Kabeer (2019) also documents cases of women whose relationships with their husbands improved over the course of the program.

The second is the importance of Self-Help Groups (SHG). Initial objectives of SHG were to encourage savings and loans. In many GPs these SHG also act as savings commitment devices, allowing women to protect their savings from their husbands, providing them with some form of financial security (p. 211). But beyond their immediate role in promoting financial inclusion, these SHG were seen to be a safe space for women to discuss and find support on a number of other issues from the personal to the productive: "SHGs had come to represent a valued new set of relationships for some of the women with impacts on their consciousness, agency and engagement in collective action (p. 210)." This quote provides another hint of the transformative potential of graduation programs, when they include ways to enhance social capital. In addition, Kabeer (2019) documents instances of hostility between project participants and non-participants in the West Bengal case. A similar experience of hostility is also found in Haiti (Roelen et al., 2019), Burundi (Devereux et al., 2015) and in Uganda (Blattman et al., 2016). In the latter case, Blattman et al. (2016) find that the SHG play an important role in minimizing hostilities between groups.

The third relates to differences in entrenched gender norms, which Kabeer (2019) explains can account for some of the differences in the magnitude of the impact between Pakistan and India. Specifically, women in the India study were far more easily able to make use of their additional resources without the support of an adult male, echoing the earlier finding in India's SKS program (Goodwin et al., 2015). Part of the differences in this respect are driven by the intersectionality with social identity (p. 212). One lesson to be drawn from these heterogeneous impacts is the need to carefully consider the role of intersectionality at design or implementation phase – in addition to the degree of pre-existing patriarchal social norms. Indeed

¹⁴ Though conducted in the same setting and evaluating the same programs as those in Banerjee et al. (2015), Kabeer's (2019) study is "stand-alone", neither contemporaneous with nor integrated within the RCT evaluations. Furthermore, and the work was implemented by different organizations (Orangi Charitable Trust for Pakistan and Trickle-Up for West Bengal).

¹⁵ The pivotal role played by household harmony and cooperation within the relationship was corroborated by discussions with GP practitioners.

women are not a homogenous group – many women may face other sources of marginalization (such as religious or ethnic affiliation, race and sexual identity) that may act as barriers to participating or, if they do participate, may limit the degree to which they are able to benefit from their participation.

Roelen et al. (2019) turn their attention to children's outcomes, but the study is included here because it offers key insights on non-economic outcomes and the gender transformative approaches used in the program they evaluate – Chemen Lavi Miyò (CLM) implemented by FONKOZE in Haiti. This study is especially rich in documenting the significant gendered barriers to empowerment and how the implementing agency builds in a gender approach to the program. First, among the vast array of topics covered, the home visits (coaching component) explicitly cover sexual and reproductive health, family planning and childbearing at a young age. In addition, beneficiaries and their husbands are invited to participate in regular three-day training sessions aimed at increasing social capital as well as forming sustainable savings groups. Part of this training and coaching involved confidence building exercises and changing husbands' attitudes and behaviours. According to their study, the key to behaviour change was the frequency and regularity with which supervision (through coaching and training) took place. Simanowitz and Greely (2017) provide additional detail about FONKOZE's advocacy strategy, which includes changes in attitudes, discourse, process and content and behaviour.

Second, like the qualitative work by Ahmed et al. (2009) and Kabeer (2019), Roelen et al. (2019) document that the quality of the relationship between the beneficiary and her husband is a contributing factor in determining GP success. While their work suggests that the quality of the relationship at baseline is important, they do document a few cases in which the "(project) managers were able to improve relationships and foster spousal cooperation (p. 49)" and "helped some members extricate themselves from disruptive and potentially harmful relationships, such as by allowing them to build a house for themselves to live in (p.50)". This evidence points to the transformative potential of the project manager: beyond improving access to productive resources (such as the asset or cash transfer or savings programs), the project manager in these cases appear to play a role akin to a social worker to ensure that the remaining components of the program can be successful. This is echoed in the qualitative evidence in Devereux et al. (2015) which relates the experience of a particular case worker who was able to reduce conflict within the couple (p. 100).

Third, and perhaps most importantly, Roelen et al. (2019) identify the lack of child care as a significant barrier to women's ability to undertake the productive activities associated with the CLM programme. In addition to the scarcity of formal childcare options, women report lack of trust in leaving their children with others and, since social norms around the "evil eye" prevent women from carrying young children on their back, it is difficult to conceive how women will be able to fully engage

economically and benefit from the program. They further report that the program did not build in a mechanism to change gender norms around childcare and instead might have reinforced traditional gender norms in this domain (p. 44 and 48).

Devereux et al. (2015) report three interesting findings of qualitative analysis of the Concern Worldwide's Burundi program. Alongside the RCT in which they randomized the number of home visits, the qualitative analysis produced a number of interesting insights, echoing results from other settings. First, they provide evidence that participation in the program enabled improved spousal relationship citing one case in which the case manager was instrumental in advising the couple to reduce conflicts and tension. Another respondent stated that her increased income and contributions to the household lead to better communications with her spouse. Second, they provide qualitative evidence of increased self-esteem, confidence and social capital among program beneficiaries. Third, they also document resentment and jealousy by non-beneficiaries,

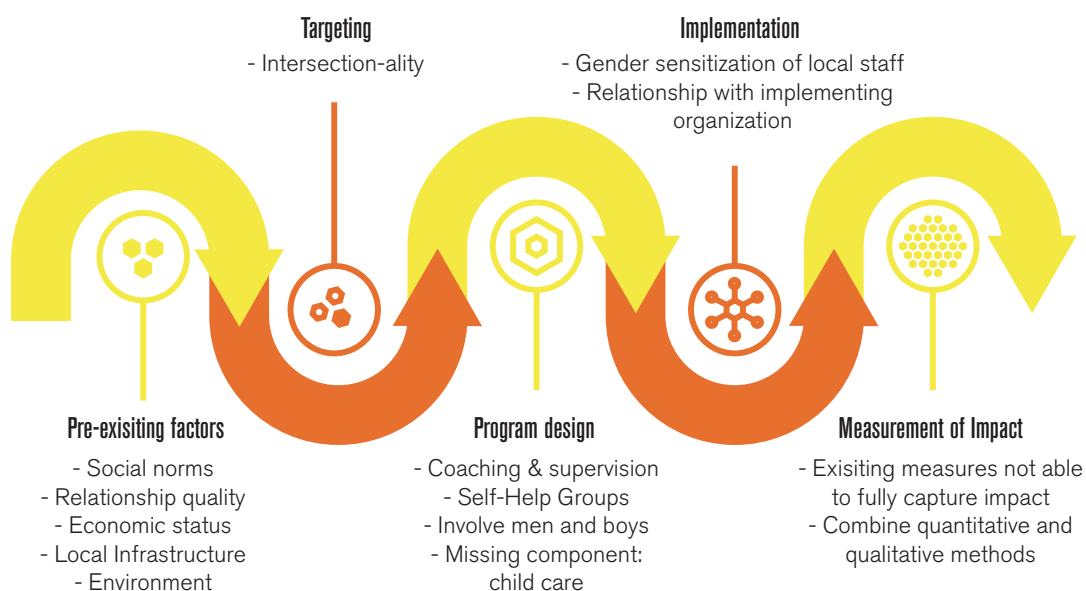
Finally, the Sanson et al. (2018) paper conducts a mixed-methods analysis of the impact on persons with disabilities of Trickle-Up's graduation programs (in Guatemala, Mexico and Nicaragua). While the population studied here is not explicitly women, we include this paper here for two reasons. First, persons with disabilities are similarly vulnerable as other marginalized persons and so lessons learned from this study can provide relevant insights. Second, the intersectionality between gender and disability status should be especially interesting and important to recognize, as women with disabilities may be an especially vulnerable population requiring special needs pertaining to their program participation. Sanson et al. (2018) identify difficulties involving persons with disability in GP programs, owing to "low self- confidence, suspicion of outsiders, risk-averseness, negative experiences with previous programmes, and fear of community stigma (p. 56)", considerations not unique to this population and that must be addressed at the targeting stage.



4. Lessons learned/best practices

Based on the review of the literature in section 3, and informed by the conceptual framework in section 2, we can collate some key lessons learned and best practices into 5 categories corresponding to 5 different stages of programming. This is depicted in Figure 2, as a process beginning with pre-conditions, followed by targeting, design issues, implementation issues and finally measurement of impact. We discuss each in turn.

Figure 2



4.1. Pre-existing conditions

The qualitative literature reviewed above suggested a few instances in which programs seemed to fare better or worse according to pre-existing conditions, while the quantitative literature is generally fairly silent on this issue. Three baseline factors emerged in this context in terms of predicting success of GPs: baseline economic status, psycho-social factors and quality of the spousal relationship. Though GPs typically target the ultra-poor by design, Kabeer (2019) and Sanson et al. (2018) do point to pre-intervention economic status as being a key ingredient for success. While Sanson et al. (2018) find that the poorest of the poor are likely to not even agree to participate in the program, Kabeer (2019) study presents two opposing predictions. In the case of West Bengal India, she finds that a group beneficiaries which she classifies as “fast climbers” – success stories so to speak – were those who were poor but already engaged productively before the intervention (the Adivasi), suggesting strong returns to prior experience and labour market skills for graduation. Worse-off beneficiaries in this case chowed greater progress. Meanwhile in the Pakistan program, she finds the reverse: better off beneficiaries did better.

This relationship with economic status may partly be driven by associated psychosocial factors. Indeed, the psychological consequences of poverty are well-documented where poverty has been linked to negative affect and stress (Haushofer and Fehr, 2014) as well as low self-esteem and low-confidence (Kakwani and Silber, 2005). Both Sanson et al. (2018) and Premchander et al. (2018) discuss these as being factors that inhibit participation in Graduation Programs in Haiti and India, respectively. Premchander et al. (2018) discuss how the Bandhan program in fact explicitly built confidence building into their coaching program, but “stayed shy of directly addressing the psychological wellbeing of the beneficiaries and capturing the status of physical and mental health in respect of illness, happiness, stress, anxiety, etc. (p. 17)”

Finally, the discussion in section 3.2 also identified the quality of the relationship with the husband as being an important determinant for success in graduation programs (Ahmed et al. 2009; Kabeer, 2019; Roelen et al., 2019).¹⁶ This ties directly to the role of intra-household dynamics, agency and social norms around traditional gender roles and the degree of autonomy and control. The qualitative evidence did point to instances in which program officers were able to help beneficiaries by either triggering a positive change in the relationship or in assisting them to exit particularly toxic relationships.¹⁷

The bottom line is that some baseline conditions do matter but that some of these can be addressed as part of the Graduation Program. Two recommendations come out of this. The first is that to the extent possible, programs should map out existing

¹⁶ This was corroborated in discussions with several of GP practitioners.

¹⁷ To our knowledge this is not something that GP implementers actively seek out to do systematically.

social norms and power dynamics within household and community to help guide curriculum for coaching. This may not always be feasible to do *ex-ante*, especially in cases where gender transformative change isn't a stated objective of the program or the program has already begun. The second is to work with a trusted local implementing partner who knows and is known by the community: they already have a sense of what preconditions may matter in that particular context.

There is also considerable space for more research in this area. Most of the evidence on the importance of preconditions in predicting graduation success comes from small sample qualitative studies or anecdotal evidence from GP practitioners.¹⁸ The larger sample quantitative evidence is fairly silent on this but could easily consider estimating heterogeneous treatment effects.

4.2. Targeting

Most GPs target women, either disproportionately or exclusively. Even recognizing that simply targeting women is not sufficient to bring about gender transformative change, a number of papers have suggested GPs could do better to improve their targeting efforts to reach out to especially vulnerable groups and women. Special attention should be placed on intersectionality, as women of particular ethnic or religious groups (e.g. in Kabeer's (2019) study) and with disabilities (Sanson et al., 2018) may resist participating and thus be excluded from the benefits of GPs because of social stigma, political or other psychosocial reasons. As discussed previously, Premchander et al. (2018) and Sanson et al. (2018) flagged the lack of self-esteem and self-confidence as significant barriers to program take-up. Furthermore, Roelen (2014) points out that many GPs by construction exclude women who are not able to work, implying that the most vulnerable women will not benefit from the programs. The main recommendation coming from this dimension is the need for GPs to heed integrate a more intersectional approach to targeting women, recognizing the special vulnerabilities of marginalized sub-populations. This may be accomplished by actively engaging stakeholders by giving them voice and agency as is discussed in the case of the CLM programme by FONKOZE in Haiti.¹⁹

4.3. Design

The literature reviewed above identified three especially promising areas in which GPs can be gender transformative – the coaching/mentoring component, the establishment or encouragement of Self-Help Groups (SHG), and the need to involve men and boys. It also identified one considerable gap in existing programs: the lack of reliable child care options. We discuss each in turn here.

¹⁸ Local public infrastructure and the physical environment are also important pre-existing conditions that can be influential in predicting success. For instance, better quality of local roads and larger local markets may theoretically improve the sustainability of entrepreneurship activities generated from a Graduation Program. While these sorts of pre-existing conditions would affect both men and women, there may be reasons to believe that they have gendered dimensions as well.

¹⁹ That said, there exists a legitimate argument that a single program cannot solve all poverty problems. Since graduation programs typically aim for productive self-sufficiency of beneficiaries, it might not necessarily be the best program to alleviate all poverty, especially for those who are physically unable to engage in productive activity. For these individuals, other forms of social protection (such as consumption support) may be preferred and more effective.

Coaching/mentoring

The transformative potential of regular coaching and mentoring around life skills – broadly defined – cannot be understated. De Montesquiou and Sheldon (2014) describe it as “perhaps the most critically successful factor for the Graduation Approach” and Devereux et al. (2015) and Devereux (2017) call it “the X-factor”. There are many varieties of coaching, however, as discussed in Hanna and Karlan (2017). They vary by content with different GPs implementing different curricula according to program priorities and local setting. They vary by the frequency of visits across programs and in the case of Devereux et al. (2015) and Blattman et al. (2016) within program. They vary by the degree to which they target the beneficiary exclusively or the household more widely – that is, do husbands participate in their wives’ coaching/mentoring sessions. The qualitative evidence is clear on the transformational potential of this component. The quantitative evidence is more difficult to obtain since the impact of this component cannot be isolated from the impact of the bundle as a whole (with the exception of the Blattman et al. (2016) in which they experimentally manipulated the inclusion of this component). Bauchet et al. (2015) propose that one reason why some of the coaching/supervision programs have failed is the low level of substantive engagement with the field staff.

Conceptually, how can coaching be so transformational? Coaching components can directly target social norms and gender issues in ways that the other components cannot. Ismayilova et al. (2018b) provides an excellent example of this from Trickle-Up’s Burkina Faso program in which they included gender-sensitive family coaching. Coaching was provided to all household members, including modules designed to raise awareness around child protection issues (e.g. the child marriage) and challenging gendered social norms and beliefs around domestic violence and women’s’ role in household decision-making. Similarly, Bandiera et al. (forthcoming) evaluate a multi-faceted program by BRAC in Uganda. While it is not a GP per se, it targets girls specifically by providing them with vocational training (hard skills) and life-skills training (soft skills). The life-skills training focused on a large number of sexual and reproductive health and rights topics, as well as softer skills around management, conflict resolution and leadership. Evaluating the program using a Randomized Controlled Trial, they provide evidence that vocational skills alone are not enough to lead to sustainable improvements in empowerment and suggesting that that programs should be bundled with life-skills training.²⁰ One intention of the life-skills training was to loosen internal constraints such as low self-confidence and aspirations. This paper directly posed the question about whether gendered social norms and low aspirations can be changed. Their results show that even 4 years after the intervention, beneficiary girls had more control over their body. They also find positive effects on their aspirations, though those results tended to be short-lived with the exception of their views on ideal age at marriage and to start childbearing. The BOMA project in Kenya similarly includes gender-focused life-skills and human rights training and coaching by a local mentor (BOMA, 2018).

20

Formally, the life-skills training isn’t labeled as a coaching program, rather these take the form of “development clubs” led by a female mentor selected from the community and trained.

Discussions with GP practitioners reinforced the importance of life-skills coaching and stressing the need to adapt to local context. Different modules are included in different, tailoring the content to specific thematic issues that take priority in the local. Including men from the start in programs was seen to be critical to mitigate potential conflicts that arise from increased bargaining power in household decisions afforded by increased asset holdings or cash. The timing matters as well. According to many GP practitioners, the earlier beneficiary households can be followed and coached, the more successful the initiative. This helps to establish trust between the beneficiary, her household and the project. The timing and sequence of the content of the training also matters: starting too early challenging traditional gender roles in especially sensitive setting, for example, may be counterproductive. The sequence of the modules thus plays a role in the extent to which coaching can change social norms.

Indirectly, the coaching and supervision by a program officer or mentor can help change some of the gendered social norms simply by their presence. We saw in the case of Roelen et al. (2017), for instance, that the program officers were able to intervene and improve some respondents' relationship issues. Regular mentoring and supervision can also play an important role in reinforcing messages around social norms and empowerment. The challenge facing GPs vis à vis this component, especially as they eye scaling up, is tailoring it to the relevant cultural and social context while keeping costs low.

Self-Help Groups

A number of GPs have introduced in some form or another self-help groups (SHG) into their programming. The most common form in which they manifest themselves are through the Village Savings and Loans Associations (VLSAs) or Rotating Savings and Credit Associations (ROSCAs), which were brought initially to encourage individual savings (de Montesquiou et al., 2014). Group savings and loans programs have been shown to be successful in raising savings because the group dynamics provide a mechanism for monitoring and enforcing loan repayment (Diamond, 1984) or because they can act as commitment device (Ashraf et al., 2006), for instance. Beyond solving credit market imperfections, groups savings and loans programs can also change the power dynamics between a woman and her spouse. Anderson and Baland (2003), for instance, demonstrate that these group savings and loans programs are sought after by women as they allow them to exert more control over their financial affairs and minimize the chances their earnings or loans get appropriated by their spouses. Groups have increasingly been integrated within GPs with a broader mission than encouraging savings or loans and often labeled as Self-Help Groups, as in the Program evaluated with Blattman et al. (2016) in Uganda. They have been shown to provide a safe-space in which women can obtain peer support on personal, social and economic issues and thus empower them at individual, household and community levels (Trickle Up, 2016; Bandiera et al., forthcoming; Kabeer, 2019). Some have argued that women's groups are in fact key to scaling up GPs (Premchander et al., 2018)

Involving men and boys

GP practitioners widely recognize the importance of involving men and especially husbands from the early stages of the program to maximize the changes of generating meaningful and lasting change to gender norms.²¹ Concern Worldwide is currently fielding a study in Malawi in which they are testing more explicitly the role of gender by targeting their program to women and men separately, and in a third treatment arm targeting women but adding a couples' empowerment component 'Transforming Gender and Power Relations' (Concern Worldwide, 2018). At the time of writing, to our knowledge, the results are not yet available.

Addressing gender dynamics and involving men must be done early, especially for programs that target women exclusively, so as to minimize any backlash or resentment by husbands and men. As we've seen above husbands are often invited to participate in regular coaching and mentoring, especially when these involve topics around gender roles within the household and attitudes (Ismayilova et al., 2018a and 2018b; Roelen et al., 2019). This can in principle be especially helpful, if done carefully, for situations in which the baseline relationship between the beneficiary and her husband is of poor quality. The more deeply entrenched gender norms, however, the harder it may be to trigger change, especially in the relatively short span of most GPs.

Missing link: the care economy

Being designed in large part to improve long-term livelihood strategies (de Montesquiou et al., 2014), many women in GPs will inevitably run into difficulties in juggling caring for family members (especially young children) and engaging in economically productive activities. The challenges of managing care, largely still disproportionately borne by women around the world, are well documented in Folbre (2018) and the literature cited therein. Roelen et al. (2019) painstakingly documents the challenges mothers face in the Haitian context balancing work and care and assert that GPs do not typically integrate childcare options into their programming, thus limiting the potential program effectiveness. There is an increasing body of evidence from the Global South that providing access to subsidized and quality care leads to increased maternal economic engagement (e.g. Clark et al., forthcoming). While small scale projects may reasonably incorporate a child care option, this may likely not be realistic for most programs and the challenge will of course be bringing these to scale.

²¹ Marcus and Harper (2015) of the ODI provide an excellent discussion of and toolkit for how to bring about changes in gender norms and identify engaging with men and husbands as key to prevent backlash. Engagement should also be undertaken within the community over concerns of stigma, which can be a contributor to backlash.

Summary

Graduation Programs can be designed to maximize their potential for gender transformative change. In the literature reviewed, and corroborated by discussions with several GP practitioners, three interconnected design features have been shown to increase the success potential for beneficiary women: regular and continued coaching and mentoring, self-help groups, and involving men and boys (especially husbands). At the heart of this potential is the role these play in affecting men's and women's consciousnesses and in changing social norms and gender roles within the household and community. Reverting back to Figure 1, these GP design features conceptually target the left half of the diagram. The qualitative and anecdotal evidence is nearly unanimous in making this case. With the exception of a few rare RCT that manipulate one component or another, the quantitative evidence has so far been unable to tease apart convincingly the impact of these components on economic or non-economic empowerment outcomes. While the quantitative literature has been able to show convincing evidence of select interventions (e.g. group savings and loans programs) outcomes ranging from savings (Ashraf et al., 2006) to intra-household decision-making (Anderson and Baland, 2003), understanding and quantifying the potential multiplier effect these components can have is far from obvious. And none of the programs evaluated and reviewed here include any child care component. Numerous authors have stressed the importance of unpacking GPs and investigating the contribution of each component separately and combined, though these have largely been left as suggestions for future research (Hanna and Karlan, 2017; Bandiera et al., forthcoming; Banerjee et al., 2018). Yet quantitatively estimating these impacts will be important in taking these programs to scale, especially with respect to coaching and mentoring as these are known to be quite expensive (Blattman et al., 2017).

Two studies have nevertheless attempted to unpack the effects of certain components within graduation programs. Banerjee et al. (2018) return to the same study sites in Ghana as those evaluated in Banerjee et al. (2015) to investigate complementarities between some components. They conduct an RCT with multiple arms: Graduation from Ultra Poverty (GUP), GUP without savings, savings only and asset only, and control. They find that the individual programs (savings only or asset only) generate similar effects on economic outcomes as the GUP. For women's empowerment, they find some positive effects from only the full GUP treatment. The savings only or assets only treatments failed to improve either the health, mental health, political participation or women's empowerment, however, consistent with the results in the Banerjee et al. (2015) study. Perhaps these components are not the ones in which we would expect gender transformative change to happen the most. Meanwhile, Sedlmayr et al. (2018) evaluate the role of a "light-touch" behavioural component in Village Enterprise's Uganda program in which beneficiaries received information about goal-setting, plan-making and other psychological services along with a cash transfer (cash+) and evaluated against a simple cash transfer. While they find that

the cash+ treatment did better than the cash only treatment in terms of increased asset holdings, it did not statistically improve the non-economic outcomes (trust, women's empowerment or protection against intimate partner violence)

4.4. Implementation

A number of qualitative studies and technical reports (and corroborated by GP practitioners) have identified two especially important ingredients for success at the implementation level. The first is the importance to ensure gender sensitization of local staff, especially if gender norms are deeply entrenched within the local community. While most programs include some training for local staff, especially around their role in supervising the household/beneficiary during the regular coaching or mentoring sessions in facilitating the self-help groups, it is not immediately clear that this training always includes some gender sensitization training. Trickle-Up (2016) provides an excellent example of the types of tools that their organization provides self-help group facilitators with a constructive gender justice manual. Similarly, Sanson et al. (2018) reflect on the importance of sensitizing local staff with respect to the intersectionalities at play.²²

Haiti's FONKOZE's program is an interesting case study in this regard. Roelen et al. (2019) document issues around the predominance of men among local staff, especially the staff interacting directly with the beneficiaries, and most egregiously that "beliefs held by programme staff reinforce existing beliefs about gender norms, particularly in terms of roles in caregiving (p. 45)." Compounding this, Simanowitz and Greely's (2017) research found that some beneficiaries of their CLM program found their facilitator "paternalistic" and "patronizing" (p. 15). Meanwhile, the organization recognizes this issue and is working with a prominent Haitian feminist to deepen gender sensitization at all levels of its operation, including with this local staff (Jean-Gilles, 2019). A very useful lesson from this case is that in contexts where gender norms are deeply entrenched, organizations may need to provide extra support and training to sensitize local staff around gender and its intersectional issues.

The second implementation ingredient predictive of success lies in the relationship between the implementing partner and the stakeholders. A key component here, not surprisingly, is the importance of trust. Heredia's (2016) qualitative study on Peru's Haku Wiñay graduation-type program reveals what many have anecdotally said or heard: potential beneficiaries' distrust, disbelief and are generally skeptical of organizations that purport to provide valuable services as well as in kind and cash transfers. Roelen et al. (2019) documents the numerous ways in which lack of trust (of kin, of the community and of strangers) is a barrier to smooth implementation of GP programs in the case of Haiti. Lack of trust is especially likely to be a concern for implementation in remote and isolated locations, according to one GP practitioner.

Engaging early with local authorities and early and frequent visits to beneficiaries are critical procedural steps.

²² Sanson et al. (2018) reported how Trickle-Up needed to deal with stigma and how the other (able-bodied) participants would deal with this ("marginalization and stigma also need to be addressed in graduation programs" (p. 61)).

4.5. Measurement

In completing the sequence in Figure 2, it might be useful to highlight concerns about measurement. Measuring women's empowerment, as discussed above, is notoriously difficult. While there are a number of widely accepted indicators of women's *economic* empowerment – namely those associated with outcomes and achievements (e.g., labour market outcomes, access to loans/savings, entrepreneurship rates, etc.) – measuring the *non-economic* dimensions is especially challenging. Cultural and social factors require obtaining instruments that are locally appropriate and locally understood given the low levels of literacy among many ultra-poor populations. Given the complexity of the concept itself, especially around agency, it is difficult to imagine reducing these measures to an index comparable across projects, countries and time. Ismayilova et al. (2018b) state it well in the context of the Burkina Faso GP: “These cultural nuances of household power dynamics are not captured by existing belief scales measuring gender norms and decision-making power which could potentially explain the lack of significant change for these outcomes (p. 456).” Bauchet et al. (2015) similarly flag the significant problems of measurement error even in the economic outcomes in the SKS (India) evaluation.

Unfortunately, the matter of measurement is unlikely to be resolved soon, especially for outcome variables that would be most appropriate to identify gender transformative change. That said, there are a number of widely used instruments to measure agency (e.g., household decision-making, individual psychosocial measures, see Donald et al. (2017)) that provide promising avenues for impact assessment, but users should beware of their advantages and disadvantages (see Laszlo et al. 2017 for a discussion).

The fact that the quantitative and the qualitative results on the gender transformative results of GPs yield at times conflicting conclusions does suggest an issue around measurability of the impacts. Increasingly, quantitative studies are fine-tuning and designing more nuanced methods to quantify the impacts on non-economic outcomes (e.g. Ismayilova et al, 2018a and 2018b; Bandiera et al., forthcoming) and there is growing interest in combining qualitative impact assessments with rigorous experimental methods. Both methods prove essential for understanding the mechanisms through which GPs can produce meaningful and lasting change and could sharpen their instruments to enable fleshing out the mechanisms outlined in Figure 1.

4.6. Unintended effects

Finally, without going into the degree of detail to do justice to the existing literature on the topic, it is important to mention that like any social protection policy, GPs may have a number of unintended negative effects even though they might show overwhelmingly large positive effects on the lives of the poor. While my review of the literature on GPs have found few such instances – namely, around the hostility of non-beneficiaries towards beneficiaries (Blattman et al., 2016; Devereux et al. 2015; Kabeer, 2019; Roelen et al., 2019) – it is useful to remind the reader of some of the negative spillovers that some social protection programs that make up the GP bundle have been shown or suspected to have.

Similarly, targeting women as beneficiaries could trigger backlash from spouses, backlash that could lead to a worsening of the relationship and in the extreme case, an increase in intimate partner violence (IPV). The evidence from other social protection programs, such as cash transfers or microfinance, shows an ambiguous effect on IPV, and despite some tragic cases the evidence seems to largely point towards a reduction.²³ The review above in section 4.3 stressed the importance of involving men from the get-go in GPs, which should by design be helpful to insure against male backlash.

As GP practitioners are increasingly looking to scale-up their programs, it is important to recognize that any program that introduces wealth, say through asset or cash transfers, may have general equilibrium effects. Of considerable concern, especially for more remote areas cut off from markets where the ultra-poor often reside, is the possibility of price effects (Cunha et al., 2019). Introducing assets of a particular type on a wide scale could affect their local prices, and widespread cash transfers could be inflationary. One way to minimize the inflationary potential of cash transfers is by limiting the amount and duration of the transfers, as many GPs do in fact do.

It is also important to note that several authors have pointed to the fact that social protection programs often reinforce traditional gender norms: e.g. Molyneux (2007) in the case of cash transfer programs and Roelen et al. (2019) in the case of GPs. Introducing gender sensitive training in coaching and mentoring can help in this respect, but unless fathers become more involved in providing childcare or childcare options are integrated within the bundle, it is unclear that GPs will be able to be immune from this consideration.

Finally, Figure 1 also shines the spotlight on the need to change formal laws to generate gender transformative change. A recent working paper by Bahrami-Rad (2019) suggests that even well-intentioned law reforms, aimed specifically at empowering women by giving them inheritance rights can backfire if informal gender norms are sufficiently entrenched. In Indonesia, they find that female inheritance rights not only increase the probability of arranged (and cousin) marriages – to keep land in the family, so to speak – and decreased women's engagement in economically productive activities.

²³ Schuler et al. (1998) provides compelling arguments as to the ambiguity surrounding whether program aimed at increasing women's income (in this case, microcredit programs) lead to an increase in IPV. On the one hand men will resist their perceived loss of bargaining power within the household. On the other hand, financial stability can give women in bad relationships the security to exit the marriage (Bobonis, 2011). Sedlmayr et al. (2018) do find some measurable reduction of IPV for beneficiaries of the GP in Uganda.



5. Conclusions

This report has reviewed the latest in the state-of-the-art evidence on the gender transformative potential of Graduation Programs. The evidence points to considerable improvements on women's economic outcomes, but for these programs to sustainably and meaningfully empower women, they must challenge the underlying structures that constrain women from fully reaching their potential. Improving access to resources alone is a necessary but insufficient condition to maximize the potential impacts of these programs. Helping women achieve greater agency and breaking down entrenched gender norms are both theoretically grounded and empirically supported to promote empowerment in its broadest sense.

By appealing to a review of high quality quantitative and qualitative evidence, supplemented with technical reports and feedback on policies and practices from Graduation Program practitioners, this review has identified several key ingredients that maximize the transformational potential of these programs. First, early, regular and frequent life-skills coaching or mentoring by program staff is well positioned to multiply the effects of the other program components. Coaching that includes gender sensitization modules are especially likely to lead to gender transformative change. Second, programs that include groups (e.g. self-help groups or savings and loans groups) and provide women with a safe-space to engage with their peers are associated with greater agency, community engagement and lower degree of hostility between beneficiaries and non-beneficiaries. Third, involving men (especially husbands) early on is important not only to increase the effect of the intervention

but also to decrease the degree of backlash that may be associated with women's increased bargaining power in household decisions. Encouraging husbands to play a greater role in child care may be especially productive given the additional demands that the economic elements of the programs will have on women's time. Fifth, some Graduation Programs have been very innovative in promoting gender sensitization of local staff, but more could be done to address the intersectionality with other forms of marginalization.

Most graduation programs address one or several of these key ingredients. While there is a general recognition that the soft-skills or coaching components are potentially key (the "X-Factor"), there is less consensus on how to implement these components at scale. Even at a small scale, they are expensive and subject to local conditions. Scaling up this component is sometimes seen as problematic, though innovative technology solutions such as the use of tablets delivering (gender sensitising) soft-skills trainings shows promise.

It is worth also emphasizing that Graduation Programs were not originally intended to be gender transformative even if they disproportionately target women. While this review has been able to provide some evidence of what might or might not work to improve the gender transformational potential of these programs, there is also a sense that there is no "magic bullet". Indeed, it may be too much to ask one class of anti-poverty programs to solve all poverty issues globally. Instead, the ability to scale up Graduation Programs will inevitably require linkages with larger macro-economic policy initiatives, institutions and infrastructure.

Finally, the report identifies several areas for future research. The first is the need to rigorously disentangle the impact of each component relative to the whole and to better quantify the potential multiplier effects that the coaching or groups components have. The second is to better understand how intersectionality factors into program design and program impact.

Appendix Table

Table 1 - Review of Quantitative Studies

Paper (*: peer review)	Country	Program	Asset Transfer	Consumption support	Savings	Health/ Ed Training	Coaching/ mentoring	WEE outcome measure	WEE Result	Method
Emran et al. 2014 (EDCC)*	Bangladesh	TUP/BRAC	in kind	cash	?	Yes	Asset/business training	Health status Health improvement Ratio saris to lunghis Female children working Female children literate Years of ed of female children	0 + 0 0 0 0	DID / DIDM
Banerjee et al. 2015 (Science)*	Ethiopia	Relief Society of Tigray	in kind	in kind	indiv	No	Yes	Mental Health Index Political Involvement Index Women's Decision-making Index	0 + 0	RCT
	Ghana	Graduation from Ultra Poverty	in kind	cash	indiv	Yes	Yes	Mental Health Index Political Involvement Index Women's Decision-making Index	0 + 0	RCT
	Honduras	Proyecto MIRE	in kind	in kind	indiv	Yes	Yes	Mental Health Index Political Involvement Index Women's Decision-making Index	+ 0 0	RCT
	India	Bandhan	in kind	cash	indiv	Yes	Yes	Mental Health Index Political Involvement Index	0 +	RCT
	Pakistan	Pakistan Poverty Alleviation Fund	in kind	cash	indiv/ group	Yes	Yes	Mental Health Index Political Involvement Index Women's Decision-making Index	0 + 0	RCT
	Peru	Arariwa	in kind	cash	indiv/ group	Yes	Yes	Mental Health Index Political Involvement Index Women's Decision-making Index	+ 0 0	RCT
Roy et al. 2015 (JDE)*	Bangladesh	TUP/BRAC	Yes	Yes	No	Yes	Yes (Community Support, awareness)	Asset ownership & control Decision-making	mixed	RCT
Devereux et al. 2015. (IDS report)	Burundi	Concern (Terintambwe)	Yes	Yes	Yes	Yes	yes	Decision-making Social Capital	mixed	Quasi experimental
Blattman et al. 2016 (AEJ- Applied)*	Uganda	WINGS	No	Cash	No	No	Business mentoring	Autonomy in purchases Emotional/physical abuse Degree of partner control partner relationship quality Woman lives w partner at endline	0 0 - + 0	RCT
Bandiera et al. 2017 (QJE)*	Bangladesh	TUP/BRAC	in kind	cash	?	Yes	Asset training	Mental Health Index Political Involvement Index Women's Decision-making Index	+ + 0	RCT
Ismayilova et al. 2018 (SSM)*	Burkina Faso	Trickle-Up	Seed capital	no	VSLA	No	yes (targeting norms and behaviours around child violence & early marriage)	child self esteem child Depression child Violence at home child Violence at work	+ + + +	RCT
Ismayilova et al. 2018 (Psych of Violence)*	Burkina Faso	Trickle-Up	Seed capital	no	VSLA	No	yes (targeting norms and behaviours around child violence & early marriage)	Financial autonomy Quality of marital relationship emotional abuse physical abuse	+ + + 0	RCT
Mahecha et al. 2018 (Working paper)	Paraguay	Sembrando Oportunidades Familia por Familia	Cash	cash	indiv	?	Yes (incl life-plan)	Aspirations Empowerment	+/0 +/0	RCT but not internally valid
Bedoya et al. 2019. (WB Working Paper)	Afghanistan	WB supported program	in kind	cash	yes	yes	Yes (life coaching, including on WEE dimensions)	Labour choices for women Psychological well-being WEE index (DM and other including aspirations for daughter, political involvement, social capital)	+ + +/0	RCT
Sedlmayr et al. 2018 (CSAE Working paper)	Uganda	Village Enterprise	No	cash	yes	no	yes (business/ savings)**	Women's empowerment, protection from IPV	0 0	RCT
Banerjee et al. 2018 (NBER Working Paper)	Ghana	Graduation from Ultra Poverty	Yes	yes	yes	yes	yes	Female empowerment Mental health	+/0 +/0	RCT

References

- Ahmed, A., M. Rabbani, M. Sulaiman and N. C. Das. 2009. "The Impact of Asset Transfer on Livelihoods of the Ultra Poor in Bangladesh" BRAC Research Monograph No. 39
- Anderson, S. and J.-M. Baland. 2002. "The economics of ROSCAs and intrahousehold resource allocation" *Quarterly Journal of Economics* Vol. 117: 963-995
- Ashraf, N., D. Karlan, and J. Lin. 2006. "Tying Odysseus to the mast: evidence from a commitment savings product in the Philippines" *Quarterly Journal of Economics* Vol. 121: 635-672
- Ashta, A., S. Khan, P. Otto. 2015. "Does Microfinance Cause or Reduce Suicides? Policy Recommendations for Reducing Borrower Stress" *Strategic Change* Vol. 24: 165-190.
- Bahrami-Rad, D. 2019. "Keeping It in the Family: Female Inheritance, Inmarriage, and the Status of Women" Unpublished manuscript. or <http://dx.doi.org/10.2139/ssrn.3250395>
- Bandiera, O., R. Burgess, N. Das, S. Gulesci, I. Rasul, and M. Sulaiman. 2017. "Labor Markets and Poverty in Village Economies" *Quarterly Journal of Economics*: 811-870
- Bandiera, O. N. Buehren, R. Burgess, M. Goldstein, S. Gulesci, I. Rasul and M. Sulaiman. Forthcoming. "Women's Empowerment in Action: Evidence from a Randomized Control Trial in Africa" *American Economic Journals: Applied Economics*
- Banerjee, A., E. Duflo, N. Goldberg, D. Karlan, R. Osei, W. Parienté, J. Shapiro, B. Thuysbaert, C. Udry. 2015. "A multifaceted program causes lasting progress for the very poor: Evidence from six countries" *Science* Vol. 348, Issue 6236.
- Banerjee, A., Karlan, D., R. Osei, H. Trachman, C. Udry. 2018. "Unpacking a Multi-Faceted Program to Build Sustainable Income for the Very Poor" NBER Working Paper 24271
- Bardhan, P. 1996. "Efficiency, Equity and Poverty Alleviation: Policy Issues in Less Developed Countries" *Economic Journal*, Vol 106 (438): 1344-1356
- Bauchet, J., J. Morduch and S. Ravi. 2015 "Failure versus Displacement: Why and Innovative Anti-poverty Program Showed No Net Impact on South India" *Journal of Development Economics* Vol. 116: 1-16
- Bedoya, G., A. Coville, J. Haushofer, M. Isaqzadeh, J. Shapiro. 2019. "No Household Left Behind: Afghanistan Targeting the Ultra Poor Impact Evaluation" World Bank Policy Research Working Paper 8877.
- Besley, T. and R. Kanbur. 1993. "The Principles of Targeting" in Lipton, M. and J. van derGaag (eds.) *Including the Poor*, Washington, D.C. The World Bank
- Blattman, C., E. Green, J. Jamison, M.C. Lehmann and J. Annan. 2016. "The Returns to Microenterprise Support among the Ultra Poor: A Field Experiment in Postwar Uganda" *AEJ-Applied Economics* Vol. 8 (2): 35-64
- Bobonis, G. J. 2011. "The Impact of Conditional Cash Transfers on Marriage and Divorce" *Economic Development and Cultural Change* 59(2), 281-312.
- BOMA. 2018. *Increasing Voice, Choice and Agency: An Analysis of BOMA Participants from the Bill & Melinda Gates Foundation Funded Cohort*. The BOMA Project.
- BRAC. 2013. *An end in sight for ultra-poverty: Scaling up BRAC's graduation model for the poorest*. BRAC Briefing Note #1.
- CGIAR Research Program on Aquatic Agricultural Systems (CGIAR). 2012. "Building Coalitions,

Creating Change: An Agenda for Gender Transformative Research in Development Workshop Report, 3-5 October 2012, Penang, Malaysia." CGIAR Research Program on Aquatic Agricultural Systems. Penang, Malaysia. Workshop Report: AAS-2012-31.

Clark, S., C. Kabiru, S. Laszlo, S. Muthuri. Forthcoming. "The Impact of Childcare on Poor Urban Women's Economic Empowerment in Africa" Demography

Concern Worldwide. 2018. Enabling Sustainable Graduation out of Poverty for the Extreme Poor in Malawi. Graduation Programme Brief No. 3.

Cunha, J., G. De Giorgi, and S. Jayachandran. 2019. "The Price Effects of Cash Versus In-Kind Transfers" Review of Economic Studies Vol. 86(1): 241-281

Das, J., Q-T. Do, and B. Ozler. 2005. "Reassessing Conditional Cash Transfer Programs" World Bank Research Observer Vol. 20(1): 57-80.

Das, N., R. Yasmin, J. Ara, M. Kamruzzaman, P. Davis, J. Berhman, S. Roy, S., and A. Quisumbing. 2015. "How Do Intrahousehold Dynamics Change When Assets Are Transferred to Women? Evidence from BRAC's Challenging the Frontiers of Poverty Reduction—Targeting the Ultra Poor Program in Bangladesh" IFPRI Discussion Paper #01317

de Montesquiou, A., T. Sheldon, F. DeGiovani, and S. Hashemi. 2014. From Extreme Poverty to Sustainable Livelihoods A Technical Guide to the Graduation Approach. CGAP and Ford Foundation.

Devereux, S. 2017. "Do Graduation Programs Work for Africa's Poorest?" Chapter 11 in Lawson, D., Ado-Kofie, L., and Hulme, D.(eds) What Works for Africa's Poorest: Programmes and policies for the extreme poor Rugby, UK: Practical Action Publishing.

Devereux, S., K. Roelen, R. Sabates, D. Stoelinga, and A. Dyevre. 2015. "Final Evaluation Report: Concern's Graduation Model Program in Burundi" IDS Report December 2015

Diamond, D. 1984. "Financial Intermediation and Delegated Monitoring" Review of Economic Studies Vol. 51: 393-414

Donald, A., G. Koolwal, J. Annan, K. Falb and M. Goldstein. 2017. "Measuring Women's Agency" World Bank Policy Research Working paper #8148.

Emran, M.S., V. Robano, S. Smith. 2014. "Assessing the frontiers of Ultrapoverty Reduction: Evidence from Challenging the Frontiers of Poverty Reduction/Targeting the Poor, an Innovative Program in Bangladesh" Economic Development and Cultural Change Vol. 62 (2): 339-380.

Folbre, N. 2018. Developing Care: Recent Research on the Care Economy and Economic Development. Ottawa: IDRC

Goodwin, S., A.P. Voola, R. Voola. 2018. "What is Program Success in Ultra-Poverty? An exploration of Randomized Control Trials (RCT) and Social Inclusion Evaluations" Third Sector Review Vol. 24 (1)

Hanna, R. and D. Karlan. 2017. "Designing Social Protection Programs: Using Theory and Experimentation to Understand how to Help Combat Poverty" Handbook of Economic Field Experiments Vol. 2

Hashemi, S. and W. Umaira. 2011. New pathways for the poorest: the graduation model from BRAC. CSP Research Report 10.

Haushofer, J. and E. Fehr. 2014. "On the Psychology of Poverty" Science Vol. 344 (6186): 862-867.

Heredia, J.A. 2016. "Una mirada a los procesos locales" in Eds. J. Escobal and C. Ponce Combinando protección social con generación de oportunidades económicas: una evaluación de los avances del programa Haku Wiñay. Lima: GRADE

Hillenbrand, E., N. Karim, P. Mohanraj and D. Wu. 2015. "Measuring gender-transformative change A review of literature and promising practices" CARE USA.

ILO. 2017. World Social Protection Report 2017-2019: Universal social protection to achieve the Sustainable Development Goals. Geneva: ILO.

International Policy Center for Inclusive Growth (IPCIG). 2017. Debating Graduation. Policy in Focus Vol. 14 (2)

Ismayilova, L., L. Karimli, J. Sanson, E. Gaveras, R. Nanema, A. Tô-Camier, J. Chaffin. 2018a. "Improving mental health among ultra-poor children: Two-year outcomes of a cluster-randomized trial in Burkina Faso" *Social Science of Medicine* 208: 180-189.

Ismayilova, L., L. Karimli, J. Sanson, E. Gaveras, R. Nanema, A. Tô-Camier, J. Chaffin. 2018b. "An Integrated Approach to Increasing Women's Empowerment Status and Reducing Domestic Violence: Results of a Cluster-Randomized Controlled Trial in a West African Country" *Psychology of Violence* Vol. 8(4): 448-459.

Jean-Gilles, S. 2019. Rapport sur l'Etat des Lieux de l'empowerment des Femmes Membres du Programme CLM et CLM+. Unpublished Report. FONKOZE: Port au Prince.

Jones, N., R. Holmes, E. Presler-Marshall, M. Stavropoulou 2017. "Transforming Gender Constraints in the agricultural sector: The Potential of Social Protection Programmes" *Global Food Security* Vol. 12: 89-95.

Kabeer, N. 1999. "Resources, Agency, Achievements: Reflections on the Measurement of Women's Empowerment." *Development and Change* Vol. 30: 435-464.

Kabeer, N. 2019 "Randomized Controlled Trials and Qualitative Evaluations of a Multifaceted Programme for Women in Extreme Poverty: Empirical Findings and Methodological Reflection" *Journal of Human Development and Capabilities* Vol. 20(2): 197-217

Kakwani, N. and J. Silber. 2007. *The Many Dimensions of Poverty*. Palgrave MacMillan: London.

Laszlo, S., K. Grantham, E. Oskay and T. Zhang. 2017. "Grappling with the Challenges of Measuring Women's Economic Empowerment" *GrOW Working Paper GWP-2017-12*.

Laszlo, S, F. Majid and L. Renee. 2019. "Conditional Cash Transfers, Women's Empowerment and Reproductive Choices" *GrOW Working Paper GWP-2019-21*

Mahecha, J.G., J. Maldonado, V. Léon-Jurado, L. Villa. 2018. "Evaluación de efectos del programa Sembrando Oportunidades Familia por Familia." *Universidad Los Andes Working Paper*.

Marcus, R. 2018. *The norms factor: Recent Research on gender, social norms and women's Economic Empowerment*. Ottawa: IDRC.

Marcus, R. and C. Harper. 2015. *How do Gender Norms Change?* ODI: Knowledge to Action Resource Series 2015

Merrien, F-X., 2013. "Social Protection as Development Policy: A New International Agenda for Action" *International Development Policy* Vol. 4.2: 89-106.

Molyneux, M. 2007. *Change and Continuity in Social Protection in Latin America Mothers at the Service of the State?* Gender and Development Programme Paper Number 1. UNRISD

Oxfam, 2002. *Gender, Development and Climate Change*. Oxford: Oxfam.

The Population Council. 2019. "The Evidence Project: Strengthening Family Planning and Reproductive Health Programs Through Implementation Science" *Webpage*. Last accessed June 20 2019.

Premchander, S., A. Mokkapat and S. Dutta. 2018. "Innovations for success and scale-up: an analysis of Bandhan's Targeting the Hard Core Poor programme in India" *Enterprise Development and Microfinance* Vol. 29 (1): 13-31

- Pritchard, M., S. Kenward, M. Hannan. 2015. "The Chars Livelihoods Programme in Bangladesh: Factors that Enable, Constrain and Sustain Graduation." *IDS Bulletin* Vol. 46 (2): 35-47.
- Pritchett, L. 2018. "Alleviating Global Poverty: Labour Mobility, Direct Assistance, and Economic Growth" CGDev Working Paper 479.
- Rao, A. and D. Kelleher. 2005. "Is there life after gender mainstreaming?" *Gender and Development* Vol 13 (2): 57-69.
- Ravallion, M. 2015. "The Idea of Anti-Poverty Policy." In Eds. A. Atkinson and F. Bourguignon. *Handbook of Income Distribution* Vol. 2: 1967-2061.
- Roelen, K., S.K. Kim, I. Barnett, D. Chancheni. 2019. "Pathways to Stronger Futures in Haiti: the Role of Graduation Programming in promoting Early Childhood Development" IDS report
- Roy, S., J. Ara, N. Das and A. Quisumbing. 2015. "'Flypaper effects' in transfers targeted to women: Evidence from BRAC's 'Targeting the Ultra Poor' program in Bangladesh" *Journal of Development Economics* Vol. 117: 1-9
- Sanson, J., F. Bielsa, and S. Kumar. 2018. "Economic Empowerment for People with Disabilities Through the Graduation Approach: Lessons from Guatemala, Nicaragua and Mexico" *Enterprise Development and Microfinance* Vol. 29 (1): 49-63
- Schuler, S., S. Hashemi and S. Badal. 1998. "Men's violence against women in rural Bangladesh: undermined or exacerbated by microcredit programmes?" *Development in Practice*, Vol. 8 (2): 148-157.
- Sedlmayr, R., A. Shah, M. Sulaiman. 2018. "Cash-Plus: Poverty Impacts of Transfer-Based Intervention Alternatives" CSAE – Working paper WPS/2017-15-2
- Simanowitz, A. and Greeley, M. 2017. *Poverty, voice and advocacy: a Haitian study*, Making All Voices Count Research Report, Brighton: IDS
- Sulaiman, M. 2016. "Making Sustainable Reductions in Extreme Poverty: A Comparative Meta-Analysis of Livelihood, Cash Transfer and Graduation Approaches" unpublished manuscript, Findevgateway.org (last accessed June 30).
- Trickle-Up. 2016. *Amra Aksathe Choli: We Walk Together: Self-Help Groups as a Safe Space for Gender Justice*. Trickle-Up: NY
- UNICEF. 2007. *The State of the World's Children 2007: Women and Children – the Double Dividend of Gender Equality*. UNICEF NY.
- Van Tassel, E. 1999. "Group lending under asymmetric information" *Journal of Development Economics* Vol. 60 (1): 3-25.

