

Version: January 2025, conditionally accepted at
Nature Human Behavior as of April 2025, according
to the authors

A systematic review and meta-analysis of social safety nets, women's economic achievements and agency

Amber Peterman^{*1}, Jingying Wang², Kevin Kamto Sonke³, Janina Steinert⁴

Abstract: There are increasing calls for social safety nets (SSNs) to be designed and implemented to promote women's economic inclusion and agency, contributing to closing gender disparities globally. We investigate the extent to which SSNs achieve these goals through a systematic review and meta-analysis. Our sample includes 1,307 effect sizes representing 218,828 women across 45 low- and middle-income countries. We show significant pooled effects across all outcomes (hedges' $g = 0.107$, $p < 0.001$), driven by increases in economic achievements (productive work, savings, assets and expenditures) and agency (voice, autonomy and decision-making). Effects are robust for asset transfers, unconditional cash transfers, social care services, and public work programs. Impacts for conditional cash and in-kind transfers are smaller in magnitude and for the latter, insignificant. We conclude that SSNs can economically and socially empower women, yet evidence gaps remain.

Affiliations:

¹Department of Public Policy, University of North Carolina at Chapel Hill; Chapel Hill, NC, USA.

²Department of Sociology, University of Zurich; Zurich, Switzerland.

³Sanford School of Public Policy, Duke University; Durham, NC, USA.

⁴TUM School of Social Sciences and Technology, Technical University of Munich; Munich, Germany.

*Corresponding author. Email: amberpeterman@gmail.com; apeterma@email.unc.edu.

Social safety nets (SSNs), or economic assistance including cash, in-kind and asset transfers, are widely used policy instruments to promote household economic security, resiliency to shocks and investment in human capital.^{1,2} Alongside their ability to cost-effectively reach and promote dignity and inclusion among marginalized populations, an increasingly recognized advantage of SSNs is their ability to address drivers of gender inequality.^{3,4} Advocates for a gender-sensitive approach have argued for provisions at the systems, policy and program levels that consider and seek to close gender disparities across a wide range of outcomes.⁵⁻⁷ Evidence supports these efforts, showing that carefully designed SSNs can have multi-faceted positive effects for women, including transforming their economic status via participation in the labor force, increasing income and building wealth stores, as well as increasing their sense of agency via participation in household and community decision-making.^{1,3,8,9} Nonetheless, governments have struggled to put these recommendations into practice. For example, among 3,099 global government social protection and labor market responses that were planned, adapted or implemented through 2021 in response to the COVID-19 pandemic, less than 20% took gender into account (defined as either targeting women's economic security or supporting unpaid care).⁶ As investment in SSNs increasingly responds to global shocks, including conflict and climate change, better understanding of their potential to reverse gender gaps is needed.

Do SSNs increase women's economic achievements and promote their agency within households and communities? Previous reviews suggest they can.^{1,3,8,9} However, to date, reviews have been largely narrative, lacking meta-analytic methodology able to aggregate evidence across SSN typologies, geographies and outcomes. Further, studies captured in previous reviews are dominated by cash transfers, rather than broader modalities. In addition, multiple reviews set out to aggregate evidence on the role of design and implementation components in delivering impacts for women, however numerous reviews conclude the evidence base is too mixed or too

thin to make firm conclusions.^{1,3} Finally, some studies point to potential adverse effects or unintended consequences of SSNs for women. A common example raised is the potential for cash transfers with conditions to increase women's burden of unpaid care, reinforcing her involvement in childcare or domestic work.^{10,11} This might occur if women are designated as responsible for attending mandatory trainings associated with the intervention or for monitoring children's schooling or health due to co-responsibilities. It is also possible that interventions which appear 'successful' in terms of household-level impacts could simultaneously result in no or adverse effects for women specifically, or increased inequality across household members, warranting further examination whether these trends hold across multiple studies.

The objective of this study is to synthesize evidence on the effectiveness of SSNs on women's economic achievements and agency in low- and middle-income countries (LMICs). We define SSNs broadly following the World Bank's Atlas of Social Protection Indicators of Resilience and Equity (ASPIRE) non-contributory programming to include seven instruments: 1) unconditional cash transfers (UCTs), 2) conditional cash transfers (CCTs), 3) food, vouchers or consumable in-kind transfers, 4) productive asset transfers, 5) public work programs, 6) fee waivers and subsidies and 7) social care services.¹² We include SSN with complementary programming (bundled), as well as stand-alone economic benefits programs. Employing a systematic review and meta-analysis of experimental evaluations, we answer the following questions: 1) What is the direction and magnitude of impact of SSNs on women's economic achievements and agency?; 2) How do these impacts vary according to SSN modality, outcome measure, as well as enablers and barriers related to intervention and evaluation designs, target group and context?; and 3) What are the cost-benefit calculations accompanying these interventions? We define economic achievements and agency broadly to capture multi-

dimensional aspects of each concept and provide a comprehensive view of evidence to inform future investment and policy.

We aggregate results from 1,307 effect sizes found in 115 publications, representing 93 studies and 218,828 women across 45 LMICs. Literature is relatively recent, with over half of included papers published in the last five years (2019 or after)—indicating a quickly emerging and dynamic field. We find robust evidence that SSNs improve women's economic achievements (hedges' $g = 0.113$, $p < 0.001$, $n = 843$) and agency (hedges' $g = 0.101$, $p < 0.001$, $n = 462$). Pooled effects show that impacts from UCTs, social care and asset transfers are the most robust (hedges g 's of 0.128, 0.117 and 0.115, respectively), while public works are comparative in magnitude, but of weaker statistical significance, likely in part due to lower power (hedges' $g = 0.127$, $p < 0.05$, $n = 106$). Finally, impacts from CCTs and food, vouchers and in-kind transfers are positive, but smaller, and for the latter, insignificant. Pooled effects for the domain of economic achievement show strong effects for assets, savings, expenditures, labor force participation and productive work intensity (e.g., hours worked or earnings). On the other hand, impacts for debt and loans and care work intensity are insignificant and care work participation has insufficient power to estimate effects. Pooled effects for the domain of agency show that voice, autonomy and decision-making are robust and significant, while aspirations and leadership have insufficient power to estimate effects. Meta-regressions investigating study, intervention and outcome-level factors show few significant predictors of impacts. Studies conducted in South Asia and in Europe and Central Asia are associated with higher effect sizes (as compared to sub-Saharan Africa), however this result is driven by one study and thus cannot be generalized. CCTs and those in pilot phases are associated with lower effect sizes (as compared to UCTs). We hypothesize that low power and high heterogeneity between studies may have contributed to these minimal results.

Our study builds on two bodies of evidence. The first seeks to understand which interventions are most promising to build women's agency and economic achievements. Our results align with a narrative review of 160 experimental and quasi-experimental evaluations from LMICs, aimed at identifying what works to enhance women's agency (including measures of economic, social and political standing).¹³ Among 16 different intervention types examined, four were selected as having 'strong or moderately strong' evidence across multiple forms of agency, including cash and in-kind transfers and the graduation approach, both which fall under our definition of SSNs. Our review also builds on meta-analyses examining effects of economic interventions more broadly on women's agency or economic standing. Results from previous meta-analyses show that our pooled effects are comparable to those found for vocational training interventions and their impacts on women's employment and earnings, are lower than those found for economic self-help groups on empowerment outcomes, and are consistently higher than those found for microcredit and savings interventions and their impacts on a range of similar economic and agency outcomes.^{14–18} The second strand of literature aims to unpack under what circumstances SSNs can improve women's outcomes and how they can be better designed and implemented to leverage gender equality impacts. Previous narrative reviews and practice-based recommendations have suggested that targeting women, ensuring a benefits are of sufficient value, incorporating key complementary programming and attention to gender in operations and implementation matters.^{1,3,8,9} However, like our study, previous reviews suggest that for many design and contextual factors, more research is needed to make firm conclusions. While our study shows that impacts of SSN on women's economic achievement and agency are promising, many gaps remain.

Results

Sample

Inclusion and exclusion criteria for qualifying studies are defined in a pre-specified systematic review and meta-analysis protocol (PROSPERO #CRD42022382158) (Table S1). We follow the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines to search databases in English, French and Spanish for randomized control trials (RCTs) published and grey literature from 2003 through December 2024. Figure 1 presents the PRISMA flow diagram for the study (see Supplementary Material for detailed description). The final sample for the analysis originated from 5,120 paper hits (4,146 from databases and 974 from other sources) and includes 1,307 effect sizes from 115 papers, representing 93 studies and 218,828 women. Figure S1 shows the evolution of papers across years by publication type (top panel) and by region (bottom panel).

Table 1 reports sample descriptives for included studies at the publication-level ($n = 115$), the intervention arm-level ($n = 135$) and the effect-level ($n = 1,307$), overall and by type of SSN (note that percentages across SSN types do not sum to 100%, as some publications evaluate interventions with multiple treatment arms studying different types of SSNs, or layer multiple SSNs within an intervention package). The majority of publications examine UCTs ($n = 65$ or 57%) and CCTs ($n = 24$ or 21%), with fewer publications evaluating asset transfers ($n = 17$ or 15%), social care ($n = 9$ or 8%), public works programs and food, vouchers or in-kind transfer programs ($n = 9$ and 7, respectively). Publications cover a total of 45 countries, with effects concentrated in sub-Saharan Africa (57%), South Asia (21%) and Latin America and the Caribbean (11%) (Fig. 2). Panel A shows that most publications focus on lower-middle income settings (52%), followed by low-income (33%) and upper-middle income (15%) settings. A minority of studies take place in fragile settings (22%) or urban settings, with overall approximately a third of studies including at least some or all the sample in urban areas.

Table 1 (Panel B) summarizes characteristics of study intervention arms (n = 135).

Overall, governments were the primary implementers (38%), followed by NGOs or UN agencies (39%) and research teams (23%). Approximately half of intervention arms were assessed as mid-level in scale, while the remaining were split between pilot (33%) and at-scale implementation (21%). The average total value of economic benefits over implementation periods studied is USD \$474, which was relatively larger in asset transfer and public works arms (just over USD \$711 and USD \$631, respectively) and relatively smaller in CCT arms and social care arms (USD \$274 and USD \$387, respectively). Most intervention arms were poverty targeted in some way (68%), as well as gender targeted (75%), including targeting adolescent girls or women as mothers, caregivers, entrepreneurs, or female-headed households. Most interventions also have some form of plus component (59%, 21% with a gender-neutral plus component and 37% with a gender sensitive plus component) with asset transfers and CCTs showing greater likelihood of complementary programming (84% and 72%, respectively), while public works and care services show lower likelihood of the same (20% and 42%, respectively). Plus components were the most likely to be training or informational (44% of all interventions), or include other types of economic (13%), health or protection (9%) or other programming (10%). Most women in the study sample were in the reproductive ages, with only 16% of average samples being younger than 24 years old (youth) and 13% of average samples being over the age of 40 years.

We aimed to categorize holistic measures of women's economic achievement and agency following existing frameworks, excluding measures which were not entirely woman specific. For each category and domain, aggregate indicators or indices were also considered as long as they were woman specific and included a majority of qualifying indicators. Effects are split 64% (35%) across economic achievement (agency). Within the economic achievement domain, the most common indicators are productive work intensity (28% of all effects), followed by labor

force participation (12%), assets (6%) and savings (6%). Within the agency domain, the most common indicators are decision-making (18%), followed by autonomy and self-efficacy (12%) and voice (4%). Very few studies measure care work participation, debt or loans, aspirations and goals, or leadership. Further details detailing outcome indicator definitions and parameters and included in Supplemental Materials. The average effect relates to interventions that last 12 months (range 1 – 60 months), with 56% of interventions lasting 12 months or longer. The average effect relates to follow-up measures taken approximately 14 months after interventions ended (range: 0 – 140, or nearly 12 years), with 50% of effects being collected 12 months or longer post-intervention. Table S3 gives further details of each publication, including authors, year of publication, country of study, intervention type and name, number of aggregate impacts included in the meta-analysis (total and by domain) and overall quality assessment score.

Pooled effects

Figure 2 summarizes overall pooled effects sizes disaggregated by SSN modality (top, Panel A) and by indicators (bottom). Pooled effects are reported in standardized units (hedges' g) using robust variance estimation (RVE) modeling, which corrects standard errors for dependency within studies that present multiple relevant effect estimates per outcome type. Tables S6 - S8 give details for these estimations. Impacts overall for all SSN modalities (top) are positive and highly significant with a hedges' g of 0.107 (95% confidence interval [CI] 0.085 – 0.129, p<0.001). When disaggregated by type of SSN, pooled effects show that impacts from UCTs (n = 817), social care (n = 91) and asset transfers (n = 216) are the most robust (hedges' g of 0.128, 0.117 and 0.115, respectively) with significance levels at the p<0.001. In addition, the pooled coefficient on public works (n = 106) is large in magnitude, but significant at a lower level (hedges' g 0.273, at p<0.05). Finally, impacts from CCTs (n = 167) and in-kind transfers (n = 112) are positive, but relatively smaller and the latter is insignificant. Overall, heterogeneity

between studies is high with most I^2 values in the 80% to 95% range and most τ^2 values exceeding 0.02, including substantial diversity in interventions, settings, study designs, or operationalization of outcomes. Table S9 replicates effects by domain of outcome, showing relatively consistent impacts across economic achievement and agency, with some exceptions. For example, effects for both CCT domains are insignificant, but of comparable magnitude to the pooled effect; impacts for in-kind transfers driven by the agency domain and impacts for care services and public work programs are driven by the economic achievement domain.

Figure 2 displays impacts by indicator type and domain (bottom, Panel **B**), showing similar magnitude of hedges' g for economic achievement (0.113) and agency domains (0.101, $p<0.001$ for both measures). Table S8 shows details underlying these estimates. Strong effects for economic achievement are seen for savings (hedges' g 0.229 [95% CI 0.147 – 0.311]), assets (hedges' g 0.235 [95% CI 0.125 – 0.345]), expenditures (hedges' g 0.177 [95% CI 0.072 – 0.282]), labor force participation (hedges' g 0.106 [95% CI 0.058 – 0.154]) and productive work intensity (hedges' g 0.075 [95% CI 0.044 – 0.106]), all at the $p<0.001$ level. Impacts for debt and loans and care work intensity are insignificant and care work participation has insufficient power to estimate effects. Robust and significant effects are seen for agency outcomes of voice (hedges' g 0.172 [95% CI 0.048 – 0.297]), autonomy and self-efficacy (hedges' g 0.105 [95% CI 0.046 – 0.165]) and decision-making (hedges' g 0.087 [95% CI 0.036 – 0.37]) at the $p<0.001$ level. Meanwhile, aspirations and goals, as well as leadership have insufficient effect sizes and statistical power to estimate pooled effects. Similar to previous pooled estimates, most I^2 values are in the 80% to 95% range, and most τ^2 values exceeding 0.02, thus pointing to substantial heterogeneity and variability between studies.

We provide a number of sensitivity analyses and robustness checks for these main results and report these in Supplementary Material. First, we replicate results with an alternative model,

a multi-level random-effects model with robust standard errors clustered at the study level. We find very similar results with only slight changes to point estimates and 95% confidence intervals at the second or third decimal. In addition, we conduct a robustness check removing influential effects using visual inspection of Cook's D , again finding very similar effects (Fig. S3, Table S12).¹⁹ Finally, we provide a visual representation of pooled effects using the Orchard plot, an alternative to the Forest plot, for estimations with a large number of effects. We report these in Fig. S5 (by outcome domain) and S6 (by SSN type).

Meta-regression

Table 2 reports RVE meta-regression results by background characteristics, grouped into three levels: publication, intervention and effect (Table 1, S3). The left panel uses a set of basic controls consisting of paper and study quality factors, while the right panel uses a set of extended controls at each level (study, intervention and effect levels). Overall, significant correlations are largely consistent between basic and extended control models. At the study level, studies in Europe and Central Asia are associated with higher effect sizes, as compared to sub-Saharan Africa (the omitted category), while studies in Latin America and the Caribbean are associated with smaller effect sizes (only in the basic control model). However, the result for Europe and Central Asia is driven by a one study with a high benefit level, thus should be interpreted with caution. At the intervention level, there is consistent evidence that CCTs are associated with smaller effect sizes as compared to UCTs (the omitted category). Likewise, pilot interventions are associated with smaller effect sizes (as compared to mid-level interventions). At the effect-level, as compared to productive work intensity, labor force participation, savings, assets, expenditure and voice are associated with higher magnitude effects in extended control models. No other effect-level outcomes show significance. We conduct a variety of robustness checks varying the cut-offs and functional form of these indicators, however, find no additional

meaningful correlations. I^2 values exceed 90% in all cases, pointing to substantial variability between studies and suggesting that covariates do not sufficiently explain this variation. We provide sensitivity analysis of these findings using the aforementioned multi-level random effects model in Supplementary Material.

Cost-benefit analysis

A total of 25 papers (or 22%) representing 23 studies (or 25%) reported on some measure of cost-benefit within studies (Table S10). Due to the low percentage of studies reporting these measures, as well as the diversity of measures reported, we summarize these in narrative form. The most common measure reported were benefit-cost ratios (BCR) (19 studies), followed by internal rate of return (IRR) (12 studies), cost-effectiveness (5 studies), net present value (4 studies) and economic multipliers (2 studies). Table S10 presents these studies sorted by size of BCR (the most reported measure), followed by remaining measures. Studies reporting cost-benefit results are overwhelmingly UCTs or asset transfers (or a combination of the two), while the remaining are CCTs (two studies), public works (one study) or social care services (1 study). Overall, BCRs (or range of estimates) are positive and include within bounds a number greater than one in all but one study—ranging from a high of 16.9 for a UCT in Tunisia²⁰ to -1.98 for a combined asset and in-kind transfer in Honduras.²¹ Studies use a variety of methods and assumptions in calculations; however, only two (in Egypt and Nicaragua) take into account women's labor force participation or earnings in the benefit calculations, while most base calculations exclusively on household consumption and asset accumulation. As interventions typically deliver a wide range of additional benefits, including women's economic activity, which contributes directly to households' ability to increase wellbeing and increase resilience against shocks, BCRs summarized here are likely lower bounds. Ten studies report exclusively positive IRR estimates, ranging from a high of 73% in a UCT in Niger²², to a low of 6% in an

asset transfer and UCT intervention in Ghana.²¹ The two remaining studies report IRR lower bounds that include negative values, however these occur under assumptions that impacts would dissipate immediately or after two years.^{22,23} The IRRs for most reported interventions indicate most interventions would pay for themselves (break even points) after a nominal number of years. Five studies reported cost-effectiveness measures, typically comparing different study treatment arms. For example, a UCT versus micro-franchising experiment in urban Kenya targeting young women reported UCT arms were generally more cost effective.²⁴ However, in two other UCT experiments, the cost-effectiveness of a psychological plus component alone²⁵, or a ‘full package’ arm including numerous plus components was found to be 1.6 to 1.7 times higher than the UCT alone.²² Finally, economic multipliers in two different UCTs in Zambia were found to be 1.6 to 1.7, indicating households spent or saved 60 to 70% more than they received via transfers.²⁶ Despite these promising results, as previously mentioned, few studies explicitly included gender-specific benefits in their measures (five studies included at least one woman’s outcome in calculations and three focused on women’s outcomes). This indicates the benefits to women aggregated in this review, and gender equality considerations more generally, are largely absent from cost-benefit calculations.

Discussion and policy implications

We show highly significant pooled effects of SSNs on women’s economic achievement and agency. These effect sizes are similar to those found in previous meta-analyses of vocational training on women’s employment and earnings (50 studies, pooled effect size: 0.109),¹⁷ however smaller than those found for economic self-help groups considering empowerment outcomes (e.g., political, social, economic agency; 7 studies, pooled effect size: 0.18),¹⁴ and larger than those found for microcredit and savings interventions on a range of similar outcomes to ours (17 [12] studies, pooled effect sizes: 0.027 economic empowerment [0.019 agency])¹⁵ (Table S11).

They are also similar to effect sizes found for the impact of cash transfers on subjective wellbeing and mental health (45 studies, pooled effect size: 0.13)²⁷ or for economic empowerment interventions on intimate partner violence (16 studies, pooled effect size: 0.09).²⁸ Our analysis reflects a substantial increase in the evidence-base from previous reviews and reinforce that SSNs benefit women, increasing their agency and economic standing.

When examining effects by outcome category, we see strong support for SSNs boosting and benefiting women's labor force participation, productive work intensity, savings, assets and expenditures. Favorable impacts align with previous reviews and analyses of cash transfers in LMICs—showing little evidence of 'dependency' effects, whereby participants may reduce work effort due to benefits (e.g., there is a negative relationship between the intervention and either labor force participation or work intensity).^{29,30} In addition, we show no evidence of increased care work intensity and debt or loans (both with null effects). While it would be encouraging to see *reductions* in care work intensity, signaling a possible shift in norms around men's involvement in domestic work—the number of studies which measure care work on the intensive margin is still relatively low (16 studies). We also see strong evidence for positive impacts on voice, autonomy and self-efficacy and decision-making—yet lack sample sizes to be able to estimate effects on aspirations and goals, or leadership. Decision-making indicators make up a relatively smaller proportion of all agency indicators (52%) as compared to previous reviews, nonetheless suffer from measurement limitations, including lack of specificity, ambiguity regarding jointness in decisions and variability across indicator construction, among others.^{31–33} There are also few estimates for combined (aggregate) measures of agency, indicating most experimental studies SSNs have not attempted to holistically measure agency as an overall construct.

Our second research objective was to examine variation in impacts by SSN type, as well as by design and contextual factors. The most striking differences emerging for CCTs and food, voucher or in-kind transfers—for which pooled effects are smaller in magnitude and for the latter, insignificant. The differences for CCTs emerge in the meta-regression, including in models which control for other intervention-level factors which might explain this difference (i.e., value of benefits, gender targeting). One hypothesis may be that that CCTs (like in-kind transfers) may restrict women's choices regarding use of benefits, thus giving her less autonomy over spending. In addition, due to conditionalities, there may be some disempowering aspects of programs, leading to reinforcement of gender roles, as has been found in other studies.^{10,11} For example, if women are responsible for attending required trainings or ensuring children attend health check-ups, this may both reduce her time to engage in productive work (particularly if meetings are far away or at inconvenient times), as well as reinforce her role as a primary caretaker. For these reasons, there have been increasing calls to drop conditionalities, in order to ensure the most marginalized women can maintain eligibility and to avoid overburdening them with program compliance requirements.⁶ Our evidence supports these recommendations, however we note we cannot fully disentangle conditions from other design features of such programs.

Our final research objective was to assess the cost-benefit evidence associated with included studies. A minority of studies included costing estimates and among those that did, few considered women's impacts as part of benefit calculations—instead relying on household-level impacts. Thus, while our summary confirms that interventions largely appear to be smart investments, additional shifts are needed for cost-benefit analysis to incorporate a gender lens. This shift would necessitate an approach which could account for intra-household benefits or count (aggregate) non-economic benefits like agency measures.

A key limitation of the current analysis is this lack of concrete program design and contextual factors which appear to meaningfully predict effectiveness. This could be in part due to power issues in the meta-regression, as well as generalized high heterogeneity of studies. For example, the current SSN literature stresses the importance of targeting women and providing complementary programming to deliver holistic impacts – yet we do not observe these design components are associated with larger effect sizes. These findings align with two recent meta-analyses: One on early childhood outcomes which concludes that cash plus was rarely more effective than cash alone,³⁴ and the second which finds that many educational gains may be made for girls through nontargeted programs.³⁵ Nonetheless, our results do not lend themselves to strong conclusions, partially because of the heterogeneity of plus programming in our sample on one hand, and the low variability in gender targeting on the other. Relatedly, although we find little evidence of publication bias, we are not able to fully investigate researchers' decisions to measure women's outcomes or analyze outcomes disaggregated by gender in the first place. It could be that the studies that produce impacts on women are those most likely to have objectives related to and design considerations appropriate for benefiting women. If this is the case, we may be both underestimating the importance of design factors – for example, the contribution of gender-informed design – as well as overestimating the potential for gender equality impacts. Additionally, it is well established that impacts of SSNs can vary over time, including the fade out of initial impacts or the accumulation of longer-term impacts, including intergenerational impacts. Our sample is primarily short-term—on average interventions are 12-months long and follow-up surveys are 14-months later. Thus, impacts may evolve as more studies conduct longer-term follow-ups, and it is unclear which direction this may change conclusions. Finally, the fact that we only include experimental studies and despite searches, did not identify studies in French and Spanish, limits the generalizability of our findings.

Our review shows that SSNs improve women's economic achievements and agency in LMICs—a conclusion which holds for most intervention types. There is significant momentum at the global level to invest in systems and programs which ensure women benefit equally and to ensure gender gaps do not widen.^{5,6,36} Our results indicate SSNs have the greatest potential to benefit women if they are designed using unconditional modalities with cash, asset or care-based benefits. In addition, we encourage policymakers and practitioners to adopt existing practical recommendations, including eliminating barriers that may limit women's access to SSNs, extending coverage to previously excluded groups, investing in linkages with complementary services and prioritizing women's leadership and political voice in decision-making structures.⁶ We encourage researchers to continue to close evidence gaps, with respect to understudied geographies, including in fragile settings, SSN typologies and outcomes. Future research should rigorously test design and operational components, unpack the role of contextual factors, including gender norms in delivering benefits for women and expand cost-benefit analysis to incorporate a gender lens.

Methods

We use RVE meta-analysis to pool effect sizes across all outcomes as well as by SSN type and outcome. We first standardized treatment effects into standardized mean differences (SMDs), transformed so improvements in economic and agency are always positive. We convert SMDs into Hedges' g values, adjusted for potential bias due to small sample sizes. RVE meta-analysis has the advantage of correcting standard errors for dependency within studies that report on multiple estimates per outcomes type. Our correlated-effects RVE model accounts for both the variation within and between study clusters, as well as for number of effect sizes per cluster by using weights. To unpack the contribution of characteristics of the target group and setting, evaluation, and of the intervention design, we conducted RVE meta-regression. We first ran a

separate regression for each covariate independently, controlling for publication type, year of publication splines (published before 2015 or published after 2019), quality assurance (if medium quality, Table S4) and if the publication was an individual RCT (instead of cluster). In a second model, we additionally controlled for factors at each level, running three regressions for each of study-, intervention- and effect-level characteristics. We did not explicitly run sample size calculations for this analysis, however included all eligible effects based on our inclusion and exclusion criteria. However, meta-analyses and -regressions were only performed if the degrees of freedom (df) were greater than 4, as simulations have shown that the approximation of the variance distribution is not valid if $df < 4$.³⁷ Full details on methods are included in Supplementary Material.

Data and materials availability: Data and analysis replication files are available as supplementary materials and will be deposited in a public repository if accepted for publication. The data extraction form is not publicly available.

Acknowledgments: We thank colleagues at the Center for Global Development (CGD) within the larger portfolio on gender ‘cash’ and ‘care’ for helpful comments and discussions at the inception of this work, including: Megan O’Donnell, Shelby Bourgault, Mayra Buvinic, Eeshani Kandpal and Kehinde Ajayi (CGD), Victoria Kiasyo Isika (Innovations for Poverty Action) and Maren Duvendack (University of East Anglia). Participants at the International Association for Feminist Economics (IAFFE) 2023 annual conference in Cape Town and at the Association of Public Policy and Management (APPAM) 2023 annual conference in Atlanta provided helpful comments. We thank all authors who were willing to share additional information and data on

their primary studies included in this review. The views expressed in this paper are those of the authors and not the policies or views of affiliated institutions or funders.

Funding:

Bill and Melinda Gates Foundation grant to the Center for Global Development (AP, JW, KKS). Funders gave feedback on the review protocol and a previous version of this paper, however had no role in the decision to publish results or the interpretation of evidence.

Author contributions:

Conceptualization: AP, JS

Methodology: AP, JW, KKS, JS

Investigation: AP, JW, KKS, JS

Visualization: AP, JS

Project administration: AP

Writing – original draft: AP, JS

Writing – review & editing: AP, JW, KKS, JS

Competing interests: Authors declare that they have no competing interests.

	All	UCTs	CCTs	Food, voucher or in-kind	Asset transfer	Public works	Social care
<i>Panel A: Publication-level</i>							
N	115	65	24	7	17	9	9
<i>Year of publication</i>							
Before 2015	0.12	0.11	0.25	0.00	0.12	0.11	0.00
Between 2015-2019	0.38	0.40	0.50	0.57	0.65	0.11	0.22
After 2019	0.50	0.49	0.25	0.43	0.24	0.78	0.78
<i>Region</i>							
Sub-Saharan Africa	0.52	0.62	0.42	0.14	0.29	0.22	0.44
South Asia	0.19	0.29	0.04	0.14	0.53	0.11	0.11
Latin America & the Caribbean	0.17	0.06	0.42	0.57	0.12	0.11	0.33
Middle East & North Africa	0.06	0.02	0.04	0.14	0.00	0.33	0.11
East Asia & Pacific	0.04	0.02	0.08	0.00	0.06	0.11	0.00
Europe & Central Asia	0.01	0.00	0.00	0.00	0.00	0.11	0.00
<i>Income group</i>							
Low-income	0.33	0.42	0.13	0.14	0.29	0.11	0.33
Lower-middle income	0.52	0.52	0.58	0.43	0.65	0.56	0.44
Upper-middle income	0.15	0.06	0.29	0.43	0.06	0.33	0.22
Fragile setting (any)	0.22	0.22	0.25	0.57	0.12	0.33	0.00
Urban setting (any)	0.30	0.23	0.38	0.43	0.12	0.33	0.67
<i>Panel B: Intervention arm-level</i>							
N	135	75	25	9	19	10	12
<i>Implementer</i>							
Government	0.38	0.37	0.48	0.11	0.16	0.50	0.42
NGO, UN or other	0.39	0.40	0.36	0.67	0.63	0.20	0.17
Researchers	0.23	0.23	0.16	0.22	0.21	0.30	0.42
<i>Scale of implementation</i>							
Pilot	0.33	0.29	0.40	0.33	0.26	0.50	0.17
Mid-level	0.47	0.48	0.48	0.44	0.32	0.30	0.58
At scale	0.21	0.23	0.12	0.22	0.42	0.20	0.25
<i>Value of economic benefit</i>							
Total value (USD)	474	533	274	400	711	631	387
First tercile	0.39	0.39	0.40	0.22	0.32	0.40	0.50
Second tercile	0.34	0.27	0.52	0.56	0.16	0.20	0.42
Third tercile	0.27	0.35	0.08	0.22	0.53	0.40	0.08
<i>Targeting</i>							
Poverty targeting	0.68	0.73	0.76	0.56	0.79	0.50	0.42
Gender targeting	0.75	0.72	0.84	0.67	0.47	0.60	1.00
<i>Plus components</i>							
Gender-neutral	0.21	0.20	0.28	0.22	0.42	0.00	0.33
Gender-sensitive	0.37	0.41	0.44	0.44	0.42	0.20	0.08
Economic	0.13	0.17	0.16	0.00	0.26	0.00	0.08
Health or protection	0.09	0.09	0.08	0.00	0.21	0.00	0.17
Training or information	0.44	0.43	0.60	0.67	0.53	0.20	0.17
Other	0.10	0.15	0.00	0.11	0.32	0.00	0.17
<i>Female sample age</i>							
Age: <=24 years	0.16	0.12	0.32	0.11	0.05	0.20	0.08
Age: 25-39 years	0.68	0.69	0.48	0.89	0.74	0.40	0.92
Age: 40+ years	0.13	0.15	0.20	0.00	0.16	0.30	0.00
	All	UCTs	CCTs	Food, voucher or in-kind	Asset transfer	Public works	Social care
<i>Panel C: Effect-level</i>							
N	1,307	817	167	112	216	106	91
<i>Domain: economic</i>							
Labor force participation	0.64	0.62	0.68	0.25	0.53	0.79	0.92
Productive work intensity	0.12	0.07	0.16	0.09	0.06	0.41	0.16

Care work participation	0.01	0.00	0.00	0.00	0.00	0.00	0.11
Care work intensity	0.05	0.04	0.02	0.01	0.00	0.13	0.13
Savings	0.06	0.08	0.03	0.01	0.06	0.00	0.01
Debt or loans	0.02	0.02	0.02	0.01	0.01	0.00	0.00
Assets	0.06	0.08	0.02	0.01	0.17	0.00	0.00
Expenditure	0.04	0.05	0.04	0.02	0.01	0.02	0.01
Aggregate economic	0.01	0.01	0.00	0.02	0.02	0.00	0.00
<i>Domain: Agency</i>	0.35	0.38	0.32	0.75	0.47	0.21	0.08
Decision making	0.18	0.20	0.05	0.38	0.35	0.08	0.03
Autonomy & self-efficacy	0.12	0.12	0.11	0.29	0.05	0.08	0.04
Aspirations & goals	0.00	0.00	0.00	0.00	0.00	0.04	0.00
Voice	0.04	0.03	0.14	0.06	0.06	0.00	0.00
Leadership	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Aggregate agency	0.01	0.02	0.01	0.02	0.00	0.00	0.00
<i>Duration of intervention</i>							
12+ months	0.56	0.69	0.43	0.46	0.68	0.04	0.58
<i>Time post-intervention at survey</i>							
12+ months	0.50	0.55	0.46	0.43	0.85	0.09	0.22

Table 1. Sample Characteristics. Sample sizes across SSN types do not sum to the full sample, as interventions may include more than one SSN. CCT = conditional cash transfer; UCT = unconditional cash transfer.

	Basic controls			Extended controls		
	Coef	SE	I ²	Coef	SE	I ²
<i>Panel A: Study-level</i>						
<i>Region (omitted = sub-Saharan Africa)</i>						
South Asia	0.033	(0.029)	91.78%	0.046	(0.029)	91.95%
Latin America & the Caribbean	-0.063*	(0.026)	“	-0.043	(0.046)	“
Middle East & North Africa	-0.009	(0.041)	“	0.017	(0.047)	“
East Asia & Pacific	-0.070	(0.037)	“	-0.045	(0.044)	“
Europe & Central Asia	0.519***	(0.023)	“	0.526***	(0.043)	“
<i>Income group (omitted = lower-middle)</i>						
Low-income	0.038	(0.029)	91.63%	0.044	(0.034)	“
Upper-middle income	-0.020	(0.025)	“	0.008	(0.041)	“
Fragile setting	-0.010	(0.037)	91.89%	-0.019	(0.041)	“
Urban setting	-0.033	(0.026)	91.89%	-0.014	(0.030)	“
<i>Panel B: Intervention arm-level</i>						
<i>Intervention type (omitted = UCT)</i>						
CCT	-0.055 ⁺	(0.032)	91.79%	-0.075*	(0.031)	92.64%
Food, voucher or in-kind transfers	-0.046	(0.049)	“	-0.045	(0.055)	“
Asset transfers	0.002	(0.031)	“	-0.014	(0.039)	“
Public work programs	0.011	(0.049)	“	0.056	(0.050)	“
Social care services	-0.009	(0.029)	“	-0.009	(0.042)	“
<i>Implementer (omitted = NGO, UN or other)</i>						
Government	-0.008	(0.027)	91.88%	-0.026	(0.033)	“
Researchers	-0.002	(0.035)	“	-0.025	(0.041)	“
<i>Scale of implementation (omitted = mid-level)</i>						
Pilot	-0.058*	(0.026)	91.79%	-0.079*	(0.030)	“
At scale	-0.028	(0.028)	“	-0.028	(0.034)	“
<i>Value of economic benefit (omitted = first tercile)</i>						
Second tercile	-0.018	(0.025)	91.88%	-0.004	(0.028)	“
Third tercile	-0.008	(0.031)	“	-0.005	(0.035)	“
<i>Targeting approach</i>						
Poverty targeted	-0.001	(0.026)	91.93%	-0.010	(0.034)	“
Gender targeted	0.012	(0.024)	“	0.036	(0.040)	“
<i>Plus components (omitted = no Plus)</i>						
Gender-neutral	0.022	(0.050)	92.12%	0.049	(0.053)	“
Gender-sensitive	0.026	(0.043)	“	0.037	(0.053)	“
Economic	0.051	(0.032)	“	0.059	(0.044)	“
Health or protection	0.005	(0.045)	“	0.007	(0.048)	“
Training or information	-0.008	(0.037)	“	-0.010	(0.043)	“
Other	-0.011	(0.041)	“	-0.021	(0.041)	“
<i>Female sample age (omitted = ≤ 24 years)</i>						
Age: 25-39 years	0.002	(0.031)	91.96%	-0.019	(0.039)	“
Age: 40+ years	0.042	(0.041)	“	0.053	(0.051)	“
<i>Panel C: Effect-level</i>						
<i>Outcome category (omitted = productive work intensity)</i>						
Labor force participation	0.045 ⁺	(0.023)	91.26%	0.044 ⁺	(0.025)	91.40%
Care work participation	0.227	(0.162)	“	0.229	(0.158)	“
Care work intensity	-0.049	(0.029)	“	-0.050	(0.029)	“
Savings	0.138**	(0.037)	“	0.134**	(0.037)	“
Debt or loans	0.021	(0.052)	“	0.027	(0.054)	“
Assets	0.125*	(0.041)	“	0.127*	(0.041)	“
Expenditure	0.110 ⁺	(0.054)	“	0.107 ⁺	(0.055)	“
Decision making	0.031	(0.034)	“	0.025	(0.034)	“
Autonomy & self-efficacy	0.028	(0.034)	“	0.025	(0.032)	“
Voice	0.113*	(0.046)	“	0.115 ⁺	(0.048)	“
<i>Duration of the intervention (omitted = <12 months)</i>						

12+ months	0.016	(0.028)	91.883%	0.019	(0.026)	"
<i>Time post intervention at survey (omitted = <12 months)</i>						
Over 12+ months	-0.011	(0.023)	91.88%	-0.001	(0.022)	"

Table 2. RVE Meta-regression Results Across all Outcomes by Publication,

Implementation and Effect Characteristics. CCT = conditional cash transfer; SE = standard error; UCT = unconditional cash transfer; ⁺p<0.10, *p<0.05, **p<0.01, ***p<0.001. All estimates are from regressions with n = 1,307 effects and 93 studies; I² estimates are entered once per model and apply to the remaining estimates in each category; Basic control estimates are separate regressions run by background characteristic, controlling for publication type, year of publication splines, indicator for low or medium quality assurance and if the publication was an individual RCT (instead of cluster). Extended control estimates additionally control for all other covariates at the same level (study, intervention or effect).

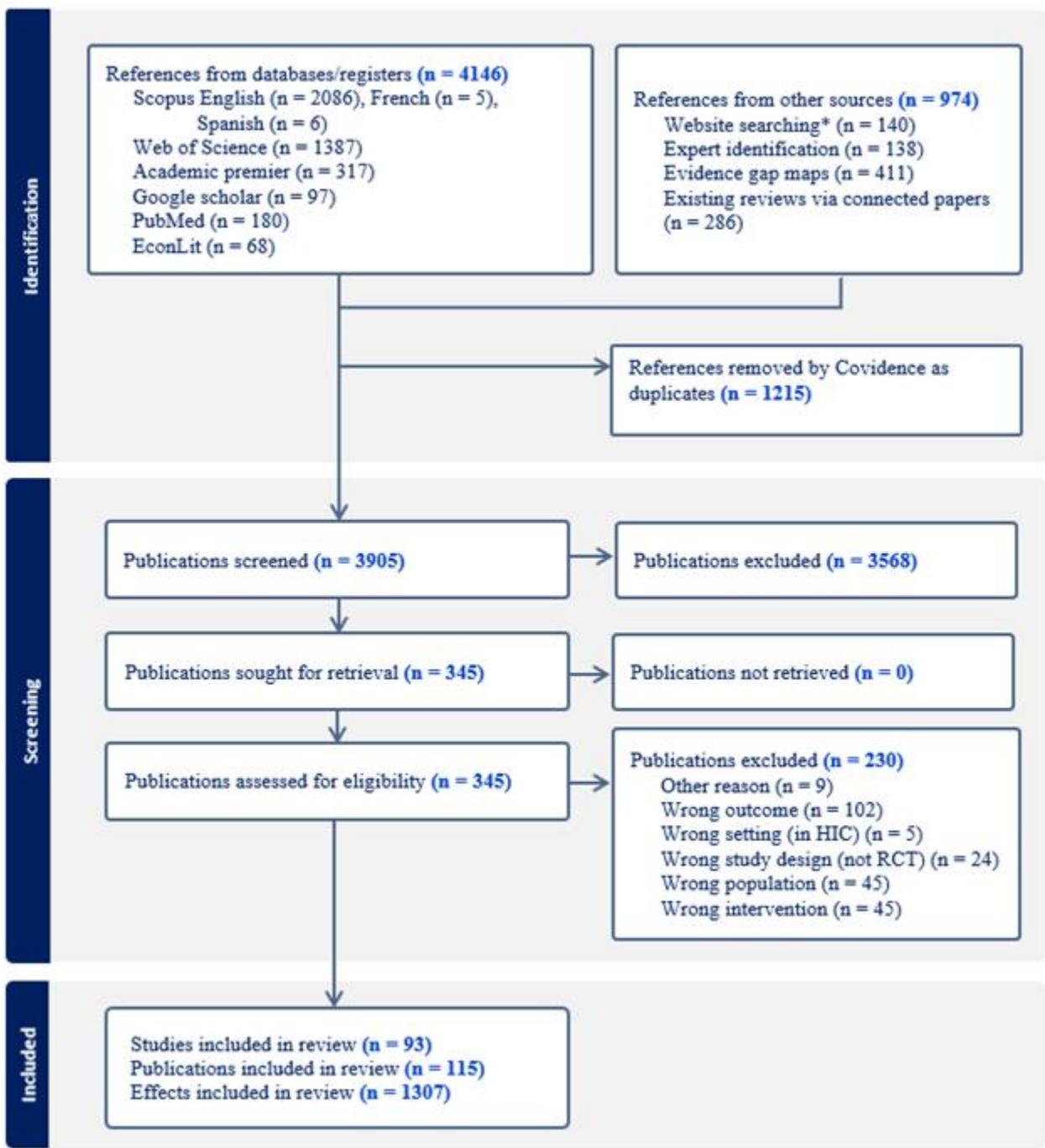


Fig. 1. PRISMA flow diagram. Websites included the World Bank open knowledge repository, World Bank Gender Innovation Labs, Asian Development Bank, Inter-American Development Bank, African Development Bank, JPAL, Innovations for Poverty Action, UNICEF Innocenti, Oxford Policy Management, the International Food Policy Research Institute, Socialprotection.org, and the Transfer Project.



Fig. 2. Geographical distribution of overall effect sizes. Distribution of effects (n=1,307) across 45 countries; including 57% of all impacts in sub-Saharan Africa, 21% in South Asia, 11% in Latin America and the Caribbean, 9% in the Middle East and North Africa, 2% in East Asia and the Pacific and less than 1% in Europe and Central Asia.

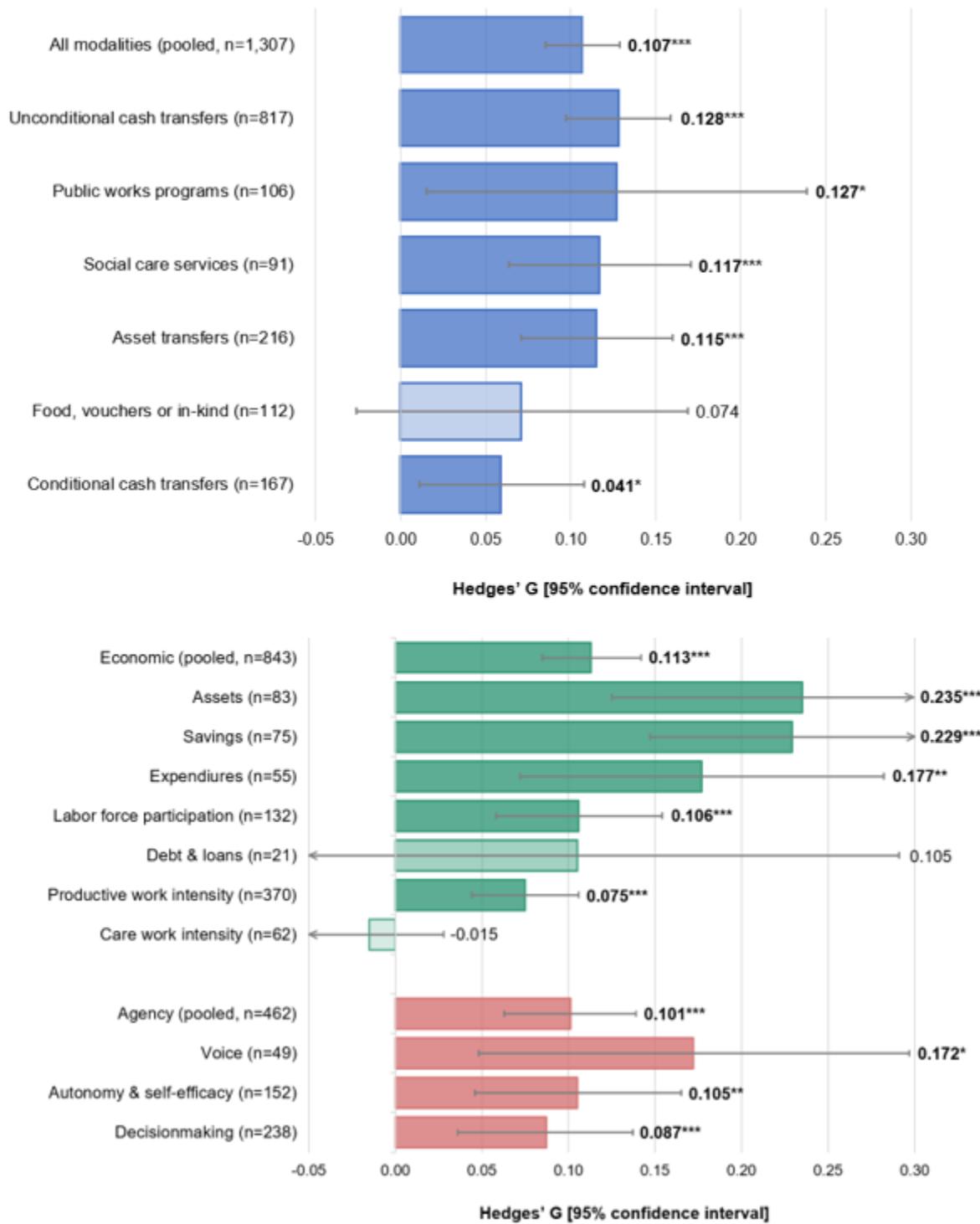


Fig. 3. Pooled effects by SSN type (top, panel A) and outcome indicator (bottom, panel B).
 Pooled effect sizes are calculated using standardized impacts and reported with 95% confidence intervals; Full results are reported in Tables S6-S8; Estimate for fee waivers (top) and for care work participation, aspirations and goals and leadership (bottom) are not shown due to insufficient power to calculate effects; ⁺p<0.10, *p<0.05, **p<0.01, ***p<0.001.

References

1. Bastagli, F. *et al.* The Impact of Cash Transfers: A Review of the Evidence from Low- and Middle-income Countries. *Journal of Social Policy* **48**, 569–594 (2019).
2. Hidrobo, M., Hoddinott, J., Kumar, N. & Olivier, M. Social Protection, Food Security, and Asset Formation. *World Development* **101**, 88–103 (2018).
3. Peterman, A., Kumar, N., Pereria, A. & Gilligan, D. *Towards Gender Equality: A Review of Evidence on Social Safety Nets in Africa*. (Intl Food Policy Res Inst, 2020).
4. UNICEF Office of Research-Innocenti. *Gender-Responsive Age-Sensitive Social Protection: A Conceptual Framework*. <https://www.unicef-irc.org/publications/1116-gender-responsive-age-sensitive-social-protection-a-conceptual-framework.html> (2020).
5. Cookson, T. P., Ebner, N., Amron, Y. & Kukreja, K. Social protection systems and gender: A review of the evidence. *Global Social Policy* 14680181231180507 (2023)
doi:10.1177/14680181231180507.
6. Gavrilovic, M. *et al.* Gender-responsive social protection post–COVID-19. *Science* **375**, 1111–1113 (2022).
7. UN Women. Social Protection Systems, Access to Public Services and Sustainable Infrastructure for Gender Equality and the Empowerment of Women and Girls; 2019
Commission on the Status of Women, Agreed Conclusions. (2019).
8. Halim, D., Perova, E. & Reynolds, S. Childcare and Mothers' Labor Market Outcomes in Lower- and Middle-Income Countries. *The World Bank Research Observer* **38**, 73–114 (2023).
9. Perera, C. *et al.* Impact of social protection on gender equality in low- and middle-income countries: A systematic review of reviews. *Campbell Systematic Reviews* **18**, e1240 (2022).

10. Cookson, T. P. *Unjust Conditions: Women's Work and the Hidden Cost of Cash Transfer Programs*. (2018).
11. Molyneux, M. Mothers at the Service of the New Poverty Agenda: Progresas/Oportunidades, Mexico's Conditional Transfer Programme. *Social Policy & Administration* **40**, 425–449 (2006).
12. World Bank, W. *The State of Social Safety Nets 2018*. (Washington, DC: World Bank, 2018). doi:10.1596/978-1-4648-1254-5.
13. Chang, W. *et al.* What Works to Enhance Women's Agency: Cross-Cutting Lessons from Experimental and Quasi-Experimental Studies.
https://www.povertyactionlab.org/sites/default/files/research-paper/gender_womens-agency-review_2020-march-05.pdf (2020).
14. Brody, C. *et al.* Economic Self-Help group Programs for Improving Women's Empowerment: A Systematic Review. *Campbell Systematic Reviews* **11**, 1–182 (2015).
15. Duvendack, M., Leon, M. D. A. & Filopoulos, J. *The Impact of Credit and Savings Interventions on Women's Economic Empowerment and Agency: A Systematic Review and Meta-Analysis*. (2023).
16. Lwamba, E. *et al.* Strengthening women's empowerment and gender equality in fragile contexts towards peaceful and inclusive societies: A systematic review and meta-analysis. *Campbell Systematic Reviews* **18**, e1214 (2022).
17. Stöterau, J., Kemper, J. & Ghisletta, A. The Impact of Vocational Training Interventions on Youth Labor Market Outcomes: A Meta-Analysis. (2022).
18. Vaessen, J. *et al.* The Effects of Microcredit on Women's Control over Household Spending in Developing Countries: A Systematic Review and Meta-analysis. *Campbell Systematic Reviews* **10**, 1–205 (2014).

19. Cook, R. D. Cook's Distance. in *International Encyclopedia of Statistical Science* (ed. Lovric, M.) 301–302 (Springer, Berlin, Heidelberg, 2011). doi:10.1007/978-3-642-04898-2_189.
20. Gazeaud, J., Khan, N., Mvukiyehe, E. & Sterck, O. With or without him? Experimental evidence on cash grants and gender-sensitive trainings in Tunisia. *Journal of Development Economics* **165**, 103169 (2023).
21. Banerjee, A. *et al.* A multifaceted program causes lasting progress for the very poor: Evidence from six countries. *Science* **348**, 1260799 (2015).
22. Bossuroy, T. *et al.* Tackling psychosocial and capital constraints to alleviate poverty. *Nature* **605**, 291–297 (2022).
23. Botea, I., Brudevold-Newman, A., Goldstein, M., Low, C. & Roberts, G. Supporting Women's Livelihoods at Scale: Evidence from a Nationwide Multi-Faceted Program. SSRN Scholarly Paper at <https://papers.ssrn.com/abstract=4560552> (2023).
24. Brudevold-Newman, A., Honorati, M., Jakielo, P. & Ozier, O. *A Firm of One's Own : Experimental Evidence on Credit Constraints and Occupational Choice*.
<https://documents.worldbank.org/en/publication/documents-reports/documentdetail/428361487270218330/A-firm-of-ones-own-experimental-evidence-on-credit-constraints-and-occupational-choice> (2017).
25. Orkin, K. *et al.* Aspiring to a Better Future: How a Simple Psychological Intervention Increases Investment. Preprint at
https://www.dropbox.com/s/u17dl4dbpx5bp4p/Aspiring_to_a_better_future_rev.pdf?dl=0 (2022).

26. Handa, S., Natali, L., Seidenfeld, D., Tembo, G. & Davis, B. Can unconditional cash transfers raise long-term living standards? Evidence from Zambia. *Journal of Development Economics* **133**, 42–65 (2018).
27. McGuire, J., Kaiser, C. & Bach-Mortensen, A. M. A systematic review and meta-analysis of the impact of cash transfers on subjective well-being and mental health in low- and middle-income countries. *Nat Hum Behav* **6**, 359–370 (2022).
28. Eggers del Campo, I. & Steinert, J. I. The Effect of Female Economic Empowerment Interventions on the Risk of Intimate Partner Violence: A Systematic Review and Meta-Analysis. *Trauma, Violence, & Abuse* **23**, 810–826 (2022).
29. Banerjee, A., Hanna, R., Kreindler, G. E. & Olken, B. A. Debunking the Stereotype of the Lazy Welfare Recipient: Evidence from Cash Transfer Programs. *The World Bank Research Observer* **32**, 155–184 (2017).
30. Handa, S. *et al.* Myth-Busting? Confronting Six Common Perceptions about Unconditional Cash Transfers as a Poverty Reduction Strategy in Africa. *The World Bank Research Observer* **33**, 259–298 (2018).
31. Ambler, K., Doss, C., Kieran, C. & Passarelli, S. He Says, She Says: Spousal Disagreement in Survey Measures of Bargaining Power. *Economic Development and Cultural Change* **69**, 765–788 (2021).
32. Peterman, A., Schwab, B., Roy, S., Hidrobo, M. & Gilligan, D. O. Measuring women's decisionmaking: Indicator choice and survey design experiments from cash and food transfer evaluations in Ecuador, Uganda and Yemen. *World Development* **141**, 105387 (2021).
33. Seymour, G. & Peterman, A. Context and measurement: An analysis of the relationship between intrahousehold decision making and autonomy. *World Development* **111**, 97–112 (2018).

34. Little, M. T. *et al.* Effectiveness of cash-plus programmes on early childhood outcomes compared to cash transfers alone: A systematic review and meta-analysis in low- and middle-income countries. *PLOS Medicine* **18**, e1003698 (2021).
35. Evans, D. K. & Yuan, F. What We Learn about Girls' Education from Interventions That Do Not Focus on Girls. *The World Bank Economic Review* **36**, 244–267 (2022).
36. Aliga, A. *et al.* Smart investment in global childcare requires local solutions and a coordinated research agenda. *BMJ Global Health* **8**, e012827 (2023).
37. Tipton, E. Small sample adjustments for robust variance estimation with meta-regression. *Psychological Methods* **20**, 375–393 (2015).