IEU LEARNING PAPER

05/2022

SCALE, DEPTH AND DURATION - EXAMPLES OF TRANSFORMATIONAL CHANGE IN THE ENERGY AND PUBLIC HEALTH SECTORS

Arne R. Weiss, Neha Sharma, Jyotsna Puri, Martin Prowse, Atika Pasha, Ojiambo Kevin Ouma, Moses Ocan, Ekwaro A. Obuku, Thomas Katairo, Markus Frölich, Mathilda Featherston-Lardeux, Robert Apunyo, Zharas Aitmambet

GREEN CLIMATE FUND | Independent Evaluation Unit

CIF CLIMATE INVESTMENT FUNDS

GREEN CLIMATE FUND
INDEPENDENT EVALUATION UNIT

CLIMATE INVESTMENT FUNDS

Scale, depth and duration - Examples of transformational change in the energy and public health sectors

Arne R. Weiss, Neha Sharma, Jyotsna Puri, Martin Prowse, Atika Pasha, Ojiambo Kevin Ouma, Moses Ocan, Ekwaro A. Obuku, Thomas Katairo, Markus Frölich, Mathilda Featherston-Lardeux, Robert Apunyo, Zharas Aitmambet

05/2022

© 2022 Green Climate Fund Independent Evaluation Unit
175, Art center-daero
Yeonsu-gu, Incheon 22004
Republic of Korea
Tel. (+82) 032-458-6450
Email: ieu@gcfund.org
https://ieu.greenclimate.fund

All rights reserved.

First Edition

This paper is a product of the Independent Evaluation Unit at the Green Climate Fund (GCF/IEU). It is part of a larger effort to provide open access to its research and work and to make a contribution to climate change discussions around the world. This evidence review has been co-financed by the Climate Investment Funds (CIF) and completed through a collaboration with the Centre for Evaluation and Development (C4ED).

While the IEU, CIF and C4ED have undertaken every effort to ensure the data in this report is accurate, it is the reader's responsibility to determine if any and all information provided by the IEU is correct and verified. Neither the author(s) of this document nor anyone connected with the IEU or the GCF, CIF or C4ED can be held responsible for how the information herein is used.

Rights and Permissions
The material in this work is copyrighted. Copying or transmitting portions all or part of this report without permission may be a violation of applicable law. The IEU encourages dissemination of its work and will normally grant permission promptly. Please send requests to ieu@gcfund.org.

The IEU reserves the right to edit text for brevity and clarity in subsequent reprints.

Citation
The citation details for this paper are as follows:
Weiss, Arne R., and others (2022). Scale, depth and duration - Examples of transformational change in the energy and public health sectors. Learning paper, (May). Songdo, South Korea: Independent Evaluation Unit, Green Climate Fund.

Credits
Head of the GCF Independent Evaluation Unit a.i.: Mr. Andreas Reumann
Task manager: Dr. Martin Prowse, Evaluation Specialist, Independent Evaluation Unit; Neha Sharma, Senior Evaluation and Learning Specialist, Climate Investment Funds
Editing: Climate Investment Funds
Layout and design: Giang Pham
Cover photo: GCF funds an adaptation project to enhance Ugandan subsistence farmers' ability in the wetlands areas to deal with climate impacts, ©Green Climate Fund

A FREE PUBLICATION

Printed on eco-friendly paper

About the IEU

The IEU was established by the GCF Board as an independent unit, to provide objective assessments of the results of the Fund, including its funded activities, its effectiveness, and its efficiency. The IEU fulfils this mandate through four main activities:

Evaluation: Undertakes independent evaluations at different levels to inform GCF's strategic result areas and ensure its accountability.

Learning and communication: Ensures high-quality evidence and recommendations from independent evaluations are synthesized and incorporated into GCF's functioning and processes.

Advisory and capacity support: Advises the GCF Board and its stakeholders of lessons learned from evaluations and high-quality evaluative evidence, and provides guidance and capacity support to implementing entities of the GCF and their evaluation offices.

Engagement: Engages with independent evaluation offices of accredited entities and other GCF stakeholders.

About the IEU's Learning Paper series

The IEU's Learning Paper series is part of a larger effort to provide open access to the IEU's work and to contribute to global discussion on climate change. The overall aim of the series is to contribute to learning and to add to global knowledge on what works, for whom, why, how much and under what circumstances, in climate change action. The findings, interpretations and conclusions are entirely those of the authors. They do not necessarily reflect the views of the IEU, the GCF or its affiliated organizations or of the governments associated with it. Comments are welcome and should be sent to ieu@gcfund.org.

About the CIF

The Climate Investment Funds (CIF) were established in 2008 to mobilize resources and trigger investments for low-carbon, climate-resilient development in select middle-income and developing countries. CIF accelerates climate action by empowering transformations through its programmes on clean technology, energy access, climate resilience, and sustainable forests in developing and middle-income countries. The CIF's large-scale, low-cost, long-term financing lowers the risk and cost of climate financing. CIF tests new business models, builds track records in unproven markets, and boosts investor confidence to unlock additional sources of finance. CIF manages a collection of targeted programmes that enable climate-smart development planning and action through 325 projects in 72 developing and middle-income countries worldwide.

About the CIF's Evaluation and Learning Initiative

The CIF was established with a mandate to serve as a learning laboratory for scaled up climate finance, to generate evidence and stimulate learning from its operations to inform and enhance current and future climate finance investments. The CIF Evaluation and Learning (E&L) Initiative is helping to fulfil this mandate through a range of strategic and demand-driven evaluations covering some of the most important and pressing challenges facing climate finance funders and practitioners. The CIF E&L Initiative identifies strategic lessons across CIF's portfolio and enables learning that is timely, relevant, and applicable to climate programmes, strategies, and investments.

About this IEU Learning Paper

The study analyses robust, causal evidence of transformational change and its drivers. We approach this topic by systematically reviewing literature that has the potential to document causal evidence for transformational change across a broad set of interventions and outcomes. Our focus is on evidence found in developing countries. We approach this learning exercise directly by looking at

the energy sector and indirectly by reviewing the evidence on behavioural change in the public health sector. Both sectors show examples of transformational change with scale, depth and duration. We identify lessons on transformational change from these two sectors that is relevant to policymakers.

TABLE OF CONTENTS

Acknowledgements x

Abbreviations xi

EXECUTIVE SUMMARY XII

I. BACKGROUND 1

A. The need for transformational change in climate change mitigation and adaptation 1

B. Searching for evidence on drivers of transformational change in the energy and public health sectors 3

C. Drivers of climate change 4

D. Categorization of interventions and outcomes 6

1. Public health 6

2. Energy 9

E. Why it is important to do this evidence synthesis 10

II. OBJECTIVES 12

III. METHODS 13

A. Criteria for considering studies for this evidence synthesis 13

1. Types of participants (population) 13

2. Types of interventions 14

3. Types of outcome measures 14

4. Comparison 14

5. Study design 14

6. Exclusion criteria 15

B. Search methods for identification of studies 15

1. Databases, repositories, and individual journals 15

a. Public health 15

b. Energy 15

2. Search terms 16

C. Data collection and analysis 17

1. Screening and selection of studies for the Evidence and gap maps (EGMs) 17

2. Selection of studies for the systematic review 17

3. Data extraction and management 18

4. Data analysis steps 19

5. Evidence and gap maps 19
6. Meta-analysis 20
a. Measures of the treatment effect 20
b. Dealing with missing data 21
c. Assessment of heterogeneity 21
d. Data synthesis 22
e. Assessment of risk-of-bias in included studies 22

IV. RESULTS 24

a. Search results for energy and health sectors 24
2. Public health 24
3. Energy 25
B. Evidence and gap maps 28
1. Public health 28
2. Energy 32
C. Description of studies for the systematic reviews 34
1. Public health 34
2. Energy 37
D. Risk-of-bias in included studies in the systematic review 38
1. Public health 38
2. Energy 39
E. Synthesis of results: meta-analysis 39
1. Public health 39
a. Effects of intervention functions on all outcome types 40
b. Effects of intervention functions on health-seeking behaviour with largely private benefits 44
c. Effects of intervention functions on health-seeking behaviour with social benefits 46
d. Effects of intervention functions on consumption/purchasing behaviour with largely private benefits 48
e. Effects of intervention functions on consumption/purchasing behaviour with social benefits 50
f. Meta-regression 52
g. Unconverted effect sizes in the utilization of the health services subsector 55
2. Energy 60
a. Effects of electrification on employment 60
b. Effects of the Chinese pilot ETS on GHG emissions 70

V. CONCLUSIONS AND IMPLICATIONS 75

a. Summary 75
B. Quality of the evidence and limitations of the review process 76
C. Areas for further investigation 78
D. Cross-sectoral learnings and authors' conclusions 79

REFERENCES 81

INFORMATION ABOUT THIS EVIDENCE SYNTHESIS 100

A. Roles and responsibilities 100
B. Declarations of interest 100

APPENDICES 101

Appendix 1. Theory of change for the public health sector 102

Appendix 2. Theory of change for the energy sector 103

Appendix 3. Inclusion and exclusion criteria for the public health sector 104

Appendix 4. Inclusion and exclusion criteria for the energy sector 107

Appendix 5. Search strategy health sector—Web of Science database 112

Appendix 6. Search strategy energy sector—Web of Science database 117

Appendix 7. Studies included in the meta-analysis (public health) 124

Appendix 8. Studies included in meta-analysis (energy) 127

Appendix 9. Tables defining terms used in behavioural change wheel 129

Appendix 10. Online supplemental materials 131

TABLES

Table IV-1. Title and abstracts screening for health sector (12,310 studies screened in total) 24
Table IV-2. Full-text screening for the health sector (all 578 studies screened) 24
Table IV-3. Title and abstract screening in the energy sector 27
Table IV-4. Full-text screening in the energy sector 27
Table IV-5. Map of studies in the public health sector, by abstract outcomes and intervention functions 29
Table IV-6. Map of studies in the energy sector, by outcomes and interventions 32
Table IV-7. Descriptive statistics for the effect sizes 36

Table IV-8. Map of effect sizes in the public health sector, by abstract outcomes and intervention function....36
Table IV-9. Summary results of the risk-of-bias assessment....39
Table IV-10. Summary results of the risk-of-bias assessment....39
Table IV-11. Meta-analysis of the intervention function cells....41
Table IV-12. Meta-analysis of the intervention function cells within health-seeking behaviour with largely private benefits outcomes....44
Table IV-13. Meta-analysis of the intervention function cells within health-seeking behaviour with social benefits....47
Table IV-14. Meta-analysis for the intervention function cells within consumption/purchasing decisions with largely private benefits....49
Table IV-15. Meta-analysis for the intervention function cells within consumption/purchasing decisions with social benefits....51
Table IV-16. Regression results for the overall sample, as well as by each abstract outcome subgroup....53
Table IV-17. Regression results for each intervention function sample (with more than five degrees of freedom remaining)....55

FIGURES

Figure I-1. Behaviour change wheel....7
Figure IV-1. Flow diagram for the health sector's evidence and gap map....25
Figure IV-2. Flow diagram for the energy sector evidence and gap map....26
Figure IV-3. Number of studies, by intervention functions, in the health sector....29
Figure IV-4. Number of studies, by source of behaviour, in the health sector....30
Figure IV-5. Number of studies, by original outcomes, in the health sector....30
Figure IV-6. Number of studies by subsector, in the health sector....31
Figure IV-7. Number of studies, by region, in the health sector....32
Figure IV-8. Number of studies, by intervention category, in the energy sector....33
Figure IV-9. Number of studies, by outcome category, in the energy sector....33
Figure IV-10. Number of studies, per region, in the energy sector....34
Figure IV-11. Funnel plot for the Education intervention function sample....42
Figure IV-12. Funnel plot for the Enablement intervention function sample....43
Figure IV-13. Funnel plot for the Persuasion intervention function sample....43
Figure IV-14. Overall effect sizes for the intervention type, Persuasion, on health-seeking behaviour with largely private benefits....45
Figure IV-15. Funnel plot for the intervention function, Persuasion, for outcomes related to health-seeking behaviour with largely private benefits....46
Figure IV-16. Overall effect sizes of the intervention type, Enablement, on health-seeking behaviour with social benefits....47

Figure IV-17. Funnel plot for the intervention function, Enablement, for outcomes related to health-seeking behaviour with social benefits....48
Figure IV-18. Overall effect sizes for the intervention type, Enablement, on outcomes indicating consumption/purchasing decisions on largely private benefits....49
Figure IV-19. Leaving one out analysis for Enablement, within consumption/purchasing decision with largely private benefits outcomes....50
Figure IV-20. Overall effect sizes for the intervention type, Enablement, on consumption/purchasing decisions with social benefits....51
Figure IV-21. Leaving one out analysis for enablement on consumption/purchasing decisions with social benefits....52
Figure IV-22. Funnel plot for the intervention function, Enablement, for outcomes related to consumption/purchasing behaviour with social benefits....52
Figure IV-23. Overall effect size for non-standardized odds ratios (left) and standardized Cohen's d (right) for the utilization of the health services sector....57
Figure IV-24. Overall effect size for non-standardized regression beta coefficients (left) and standardized Cohen's d (right) for the utilization of the health services sector....58
Figure IV-25. Funnel plot of the raw effect sizes for odds ratio (left) and regressions (right)....59
Figure IV-26. Effects of electrification on employment (forest plot)....61
Figure IV-27. Effects of electrification on employment by gender (forest plot)....63
Figure IV-28. Subgroup analysis based on region....65
Figure IV-29. Funnel plot for the effects of electrification on employment....66
Figure IV-30. Subgroup analysis for testing differences between grey and published literature....67
Figure IV-31. Subgroup analysis based on the risk-of-bias assessment....68
Figure IV-32. Outlier inspection: estimating the overall effect size of electrification while excluding one study at a time....69
Figure IV-33. Effects of electrification on employment measured, in terms of Cohen's d (forest plot)....70
Figure IV-34. Effects of the Chinese pilot ETS on logarithmic GHG emissions (forest plot)....71
Figure IV-35. Test for "small-study effects": effects of the Chinese pilot ETS on logarithmic GHG emissions (funnel plot)....72
Figure IV-36. Outlier inspection: estimating the overall effect size of electrification while excluding one study at a time....73
Figure IV-37. Effects of the Chinese pilot ETS on GHG emissions, in terms of Cohen's d (forest plot)....74

ACKNOWLEDGEMENTS

We take this opportunity to thank the EPPI Center of the University College London, John Eyers, Sarah Khan, Jörg Peters, Frank Renkewitz, Martin Rudasingwa, Ashrita Saran, Hugh Waddington, Bernd Weiss, and Howard White for their invaluable input, right from the conceptualization to the conclusion and reporting of this project.

We would also like to express our gratitude to Alex Avdeenko, Giulia Montresor, Marc Gillaizeau, Tatyana Kulp, and several others at the Center for Evaluation and Development (C4ED) for taking the time to review this work and provide valuable suggestions on small and large details during the entire process. Their counsel made this report a much better knowledge product. Even more so, this report was only made possible through the continued support and diligence of the various research assistants who put in many long hours, often later than expected, over the course of three years.

Finally, we would also like to thank Deborah Kim and Wonju Kim from the Independent Evaluation Unit of the Green Climate Fund for their contributions when the report was nearing completion. The tasks managers for this evidence review have been Neha Sharma for the Climate Investment Funds and Martin Prowse for the Independent Evaluation Unit of the Green Climate Fund.

ABBREVIATIONS

ANC	Antenatal care
C4ED	Center for Evaluation and Development
CDM	Clean Development Mechanism
CIF	Climate Investment Funds
DiD	Difference-in-differences
E&L	CIF's Evaluation and Learning Initiative
EGM	Evidence gap map
ETS	Emissions trading system
GCF	Green Climate Fund
GHG	Greenhouse gas
GRADE	Grading of Recommendations Assessment, Development, and Evaluation
HIC	High-income country
IEA	International Energy Agency
IEU	Independent Evaluation Unit
IPCC	Intergovernmental Panel on Climate Change
LMIC	Low and middle-income country
LNG	Liquefied natural gas
OECD	Organization for Economic Co-operation and Development
PICOS	Population, intervention, comparison, outcome, and study design
PNC	Postnatal care
PPP	Public-private partnership
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PSM	Propensity score matching
RDD	Regression discontinuity design
SD	Standard deviations
SE	Standard error
TCLP	CIF's Transformational Change Learning Partnership
ToC	Theory of change
TWh	Terawatt hour
UNEP	United Nations Environment Programme
UNFCCC	United Nations Framework Convention on Climate Change
USD	United States dollar
WASH	Water, sanitation and hygiene
WB	World Bank

EXECUTIVE SUMMARY

BACKGROUND

Amidst growing concern about global warming and climate change effects, especially in developing countries, the international community has responded with climate change adaptation and mitigation measures, such as the Paris Agreement and National Adaptation Programmes of Action. Mitigation measures cover efforts to reduce greenhouse gas (GHG) emissions, such as through a transition to clean-energy sources and the absorption of gases already emitted. Adaptation, on the other hand, refers to actions needed to better cope with the impacts of climate change. Though considerable investments have been made in the last decades, current levels of investment remain well below what is required to limit global warming to 1.5 degrees.

The Intergovernmental Panel on Climate Change as well as the climate literature refer to the systemic changes required to prevent dangerous climate change as requiring a paradigm shift or as a process of transformational change. However, the evidence on the causal drivers of transformational change in general, and in particular, in relation to climate change mitigation and adaptation, is still limited. Various studies and approaches have attempted to define "transformational change" with some common features: it aims to transform the structural or foundational attributes of a system over a sustained period of time with an impact on a significant number of people. This synthesis conceptualizes the attributes of transformational change as large effects (depth of change) over a longer time frame (sustained change) targeting many beneficiaries or covering large areas (scale of change).

The study analyses robust and causal evidence of transformational change and its drivers. We approach this topic by mapping the available evidence, based on our criteria in defining transformational change to identify evidence gaps, and then systematically reviewing the literature that has the potential to document causal evidence for transformational change across a broad set of interventions and outcomes. Our focus is on developing countries. We approach this learning exercise directly by looking at the energy sector and indirectly by reviewing the evidence on behavioural change in the public health sector. The public health literature has the longest tradition of long-term causal studies on how behavioural science can facilitate transformational changes. We identify lessons on transformational change relevant to policymakers.

OBJECTIVES

This report combines evidence gap maps (EGMs) with systematic reviews and focuses on the evidence from developing countries to learn about the attributes, determinants, and contributors of transformational change in the energy and public health sectors.

SEARCH METHODS

We conducted research on different databases and websites of agencies and research institutes based on the relevance and comprehensiveness of their coverage of the literature on each of the two sectors (that is, energy and public health). Besides the systematic searches, we also applied the snowballing approach to identify additional studies, after the full-text screening was finalized.

SELECTION CRITERIA

This synthesis included quantitative studies from experimental and quasi-experimental designs that have the *potential* to *document* transformational change produced by a relevant intervention in a developing country, using the proxy of the LMIC categorization of the World Bank. Studies had to measure a relevant outcome at least one year (long-term) after the onset of an intervention covering at least 1,000 individual beneficiaries (large-scale) or targeting entire administrative units larger than a village. Effects had to be estimated based on a clearly identified control group that was comparable to the group of beneficiaries in the absence of the intervention (causality and attribution). The depth of change was not part of the inclusion criteria for the individual studies; instead, the question of which interventions led to large effects over time was answered through the meta-analysis. The nature of the comparison group was dependent on the type of research design used in the study and could include both active and passive comparison groups.

The screening of interventions and outcomes followed two broad theories of change for each sector. These theories of change simultaneously structured and defined the scope of this study. For the energy sector, we covered a broad set of interventions that either targeted or could have effects on climate change mitigation and adaptation. These interventions took place either at the level of institutional and market systems through incentives and standards, that is, "soft" interventions (nudges), or investments in infrastructure. Outcomes under the purview of this report capture climate change mitigation and adaptation (resilience of energy systems), labour market co-benefits from investments, or the transition into renewable energy. For the public health sector, we included interventions targeting behavioural change in five broad areas—nutritional (dietary) habits, physical activity, substance abuse, hygiene practices, and the utilization of health care services.

The inclusion and exclusion criteria were operationalized following the PICOS (population, intervention, comparison, outcome, and study design) model.

DATA COLLECTION AND ANALYSIS

The screening and coding of studies were done by a pair of coders (double coding) on the basis of the inclusion criteria in EPPI Reviewer 4—a specialized software for managing and analysing data in literature reviews. At least 20 per cent of studies were double screened by an independent reviewer. Disagreements were resolved through discussions within the review team. The results of the screening process were presented in PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) diagrams.

With the goal of searching for evidence on transformational change, the data analysis led to two EGMs—one for each sector, along with the corresponding meta-analysis. We constructed EGMs, with interventions listed along the Y-axis and outcomes along the X-axis, to document evidence and gaps within the scope of each sector. We then concentrated on the sufficiently populated areas within these maps to conduct systematic reviews in the form of meta-analyses on the available evidence using conceptually similar studies to estimate overall effect sizes.

On a statistical level, wherever necessary and possible, we made the studies comparable by calculating the same standardized effect sizes (Cohen´s d) and excluded studies that did not provide sufficient information to do so. We relied on the conventional literature to define the thresholds for effect sizes using Cohen's d—large (d = 0.8), medium (d = 0.5), and small (d= 0.2).

The meta-analysis was conducted using a random-effects model given the large heterogeneity in our studies. We also investigated reporting biases and sources of heterogeneity (variability in the intervention effects) and ran sensitivity analyses of the distribution of the effects. The main results

were presented using forest plots. Additionally, subgroup analyses, as well as outlier detection, were performed as a means of exploring sources of heterogeneity and as sensitivity tests. Funnel plots and the Egger test were used to test for "small-study effects" to explore publication bias. In the case of the health sector, basic meta-regressions were also conducted.

RESULTS

We retrieved 32,909 articles from 13 databases and screened 19,402 titles and abstracts. For the EGMs, we included 144 studies in public health and 96 studies in energy sectors, respectively. The evidence gaps for the public health sector were for interventions that relied on raising taxes or prices (terms as coercion) and restriction (via bans or regulations to limit use), with less than four studies in both intervention types. Each of the energy sector interventions had at least 18 studies. Enablement (n=66), defined as support that reduces barriers or increases capabilities, and "investments into infrastructure, equipment, and technologies" (n=45) had the most available studies for the public health and energy sectors, respectively. In terms of the public health sector outcomes, "consumption/purchasing decisions with largely private benefits" was the least populated (n=23), while "health-seeking behaviour with largely private benefits" yielded the greatest number of studies (n=57). Outcomes in the energy sector showed that "resilience of energy systems adaptation" (n=11) and "energy market development" (n=13) had the least studies, while "energy consumption and demand" came up the most frequently (n=51).

For the meta-analysis, we included 53 studies in public health and 31 studies in energy. In public health, the meta-analysis focused on education, persuasion, and enablement (when resources or reminders are used to support action) as the three intervention functions with the largest sample sizes. Of these three, only enablement approaches a Cohen's d value of between 0.2 and 1.12—passing the thresholds for small and large-effect sizes based on the conventional definition adopted in our review. The meta-regression coefficients further indicate larger changes due to enablement, particularly on consumption or purchasing decisions. However, between-study heterogeneity (variance) was considerable.

In terms of lessons learned for climate interventions, the findings from the evidence review on behavioural change in public health do indicate early hypotheses that may need to be further tested on the types of interventions that could work best for mitigation or adaptation interventions. Overall, the findings suggest that enablement (such as when resources or reminders are used to support action) may offer potential transformational effects for mitigation interventions, while persuasion may produce potential transformational effects for adaptation (in combination with other intervention types).

The findings of the evidence review on energy offer some interesting, yet ultimately, mixed results. The meta-analysis focused on two areas: first, the effects of electrification on formal employment; and second, the effects of a pilot emission trading scheme in China. First, with regard to the effects of electrification on formal employment, the review offers estimates of a percentage point increase of two per cent, driven mainly by the greater employment of women. However, the effect of two per cent was lowered when considering heterogeneity and publication biases. The estimated Cohen's d value of 0.03 is considerably below a small effect size threshold of 0.2 and very far away from the transformational change threshold of 0.8. This determination does not take into account contextual factors; as such, other factors that we had not explored through this study may be causing these results, which would require further investigation.

On the other extreme is the reduction of emissions from Chinese pilot emissions trading system (ETS). This shows a reduction of 17 per cent from the control group. With a Cohen´s d value of -1,

which is considerable, it offers promise for transformational change within the set of studies evaluated as part of this report. However, due to the potential risks of publication bias and possible flaws in internal validity, this effect size is questionable, and we are not able to draw a firm conclusion.

CONCLUSIONS

This project applied a novel approach and combined two different evidence syntheses into one learning exercise on transformational change. Though the search for interventions that can have transformational changes will certainly continue, this report emphasizes how difficult it is to synthetize evidence on causal drivers of transformational change. This approach severely restricted the available evidence solely to those coming from quantitative experimental and quasi-experimental studies. The nature of these methods means that such studies usually focus on singular interventions on outcomes at the individual or household level within a relatively short time frame. Therefore, long-term changes over decades, rather than years occurring at the system level as a result of a complex interplay between many incremental changes from various interventions, are not captured. A key lesson from our report, therefore, is that transformational change is difficult to achieve with interventions that are studied through experimental and quasi-experimental study designs.

I. BACKGROUND

A. THE NEED FOR TRANSFORMATIONAL CHANGE IN CLIMATE CHANGE MITIGATION AND ADAPTATION

The Intergovernmental Panel on Climate Change (IPCC) estimates that if greenhouse gas (GHG) emissions maintain their rise at the current rate ("business-as-usual"), then by the end of the 21st century, the average temperature will have increased by 2.6 to 4.8 degrees Celsius and sea levels will have risen by 0.45 to 0.82 meters (Intergovernmental Panel on Climate Change, 2018).

The international community is responding with climate change adaptation and mitigation measures, such as the Paris Agreement and National Adaptation Programmes of Action. Mitigation measures cover efforts to reduce GHG emissions, such as through a transition to clean-energy sources and the absorption of gases already emitted. Adaptation, on the other hand, refers to the actions needed to better cope with the impact of climate change (Intergovernmental Panel on Climate Change, 2014; Watts and others, 2018). The recent report from the Global Commission on Adaptation (2019) highlights how priority areas in developing countries include food production, water management, city infrastructure, and the natural environment.

Ongoing global efforts are, however, not sufficient to meet the goals of the Paris Agreement. Although climate finance has risen considerably over the past years, it is still deemed too low compared to the level required to achieve a global-warming scenario of 1.5 degrees Celsius (Buchner and others, 2019). The United Nations Environment Programme (UNEP, 2021) highlights how immediate and large-scale investments in mitigation globally could limit global adaptation costs by up to 75 per cent.

However, the current levels of investments into low-carbon technologies fall short of what is required to meet the mitigation target, according to a report by the International Energy Agency (International Energy Agency, 2019b). The same agency also highlights the enormous scale of the challenge in transforming economies to a net-zero basis by 2050 (International Energy Agency, 2020b). Shortfalls in funding are also present in adaptation finance (Global Commission on Adaptation, 2019) that does not meet the needs expressed in the nationally determined contributions (Neufeldt and others, 2018). UNEP (2021) highlights how the costs of adapting to climate change impacts in non-Organization for Economic Co-operation and Development (OECD) countries could be between USD280–500 billion per year by 2050, as impacts become more severe. Current flows amount to around USD27 billion a year from both OECD and non-OECD sources.

To meet needed targets in both mitigation and adaptation, climate investments need to be increased substantially and their impacts per dollar spent need to be considerably higher. Furthermore, the longer current mitigation and adaptation measures fall short, the higher the overall impact is required by future interventions. It is, therefore, imperative to usher in interventions that have a transformational effect.

Both the Independent Evaluation Unit (IEU) of the Green Climate Fund (GCF) and the Climate Investment Funds' (CIF) Evaluation and Learning (E&L) Initiative have ongoing work in this area.

CIF's Transformational Change Learning Partnership (TCLP)[1] has been advancing the understanding of transformational change through global collaboration, facilitated learning and analysis with a diverse set of organizations and individuals from around the world since 2017. The

[1] https://www.climateinvestmentfunds.org/tclp.

partnership uses a working definition of "transformational change"—a fundamental change in systems relevant to climate action with large-scale positive impacts that shift and accelerate the trajectory of progress towards climate-neutral, inclusive, resilient, and sustainable development pathways. This definition accounts for the complex systems in which change is occurring and stresses fundamental and large-scale change. The work by TCLP posits that there are five key dimensions for transformational change—relevance, systemic change, speed, scale, and adaptive sustainability (Climate Investment Funds, 2021). This definition is an adaptation from the earlier CIF work (Williams and others, 2020) that includes all the aforementioned dimensions, except for speed.

TCLP's work is built upon others' definitions and conceptualizations of "transformation", including those of Fazey and others (2017), Olsen and Fenhann (2016), Westphal and Thwaites (2016), along with O'Brien and Synga (2013) (Ross Strategic and Community Science, 2017). Additional work that has played a key role in understanding transformational change included the International Climate Fund's (2014) midterm evaluation, CIF (2015), the Independent Evaluation Group (2016), and the Initiative for Climate Action Transparency (2017).

CIF has also published extensively on "transformational change", including an evidence synthesis, an evaluation, a publication on signals for transformational change, along with several other papers and thought-pieces (Climate Investment Funds, 2019, 2020; Williams and others, 2021). This work demonstrates that transformation, at the level of depth and breadth needed to address the climate crisis, is a very ambitious global goal, requiring changes spanning natural and human systems as well as at all scales. Such transformations that will take place at different time frames and speeds will entail evolving focus and targets as contexts change.

The work from the IEU of GCF, such as Puri (2019), argues that while the concept of "transformational change", or shifts in paradigm, has been used in multilateral development agencies, an operational definition still remains elusive. This work highlights how the three elements—significant impact, sustained, and large scale—are common in how major multinational agencies operationalize transformational change, though the way they are defined and approached may differ.[2]

International agencies have generally employed proxy indicators to measure transformational change. The International Fund for Agricultural Development collected high-quality baseline data and regular project-level data that show the constancy of programme implementation, using ex-post impact assessment methods. The examination of transformational change by the Global Environment Facility's Independent Evaluation Office was conducted through a desk review of final evaluation reports on the measures of relevance, depth of change, and sustainability, along with the presence and quality of evaluative evidence. World Bank's (WB) Independent Evaluation Group based its work on the following definition of "transformational engagements": "interventions that support deep, systemic and sustainable change with the potential for large-scale impact—fundamental change in a system, having a large-scale impact at the national or global level, and are economically financially and environmentally sustainable." Accordingly, WB's dimensions of "transformational change" are identified as relevance, depth of change, scale of change, and sustainability.

[2] For a growing literature on the substantive definitions of, as well as the theoretic mechanisms behind, "transformational change", see Feola (2015), Few and others (2017), and Kates and others (2012). While these mechanisms enter neither our definition nor our analysis of the evidence for transformational change, they do inform the selection of interventions, and thereby, the types of studies included in our synthesis. See also Ajibade and Adams (2019), Fedele and others (2019), Feinstein (2019), Mapfumo and others (2017), Termeer and others (2017), Thomalla and others (2018), Thornton and Comberti (2017), van den Berg and Cando-Noordhuizen (2017), and Wienges and others (2017).

Puri (2019) furthers this idea by arguing for the importance of building last-mile considerations and behavioural insights into discussions and the conceptualizations of transformational change. She notes that the current paradigmatic thinking that knowledge and attitudes are enough to change behaviour is untrue. Puri (2019) highlights the existence of "last-mile" problems that need to be fully addressed in order for programmes to be transformative. Also, it is important to consider high-quality evidence to be an agent of transformation, as it constitutes part of the replicability that culminates in the build-up of credible evidence.

In a recent work by IEU, Puri and others (2021) broaden this definition by using eight components to assess the likelihood of transformational change in the GCF portfolio: scale (breadth of impact); depth (impact per beneficiary unit); permanence (sustained change through time); innovation (geographical, sectoral, or institutional disruption); behaviour change (that beneficiaries reduce an intention-action gap or use other means to alter behaviour); demonstration ability (that influences other actors); policy change (whether the project catalyses policy change, strengthens policies, or increases spending on policy initiatives); as well as complementarity and coherence (with other relevant actors).

This work reviewed project documents from 125 GCF investments through March 2020 used bivariate statistics and multivariate cluster analysis to examine the likelihood of transformational change from GCF's project portfolio. The analysis found that adaptation and cross-cutting projects are more likely to be transformational than mitigation projects. Projects with the highest potential for transformational change showed the highest scores for permanence, policy change, demonstration ability, and complementarity and coherence. Overall, the GCF portfolio as of March 2020 consisted of projects with a mixed ability to contribute to transformational change. The main recommendation from this review was that the GCF can optimize its comparative advantage in adaptation finance by focusing on investments that display transformational change attributes.

B. SEARCHING FOR EVIDENCE ON DRIVERS OF TRANSFORMATIONAL CHANGE IN THE ENERGY AND PUBLIC HEALTH SECTORS

Transformational change has become the "holy grail" in climate change and development assistance. While there are examples of interventions with large effects (such as the Clean Air Act in the United States), many interventions do not replicate it when scaled up, or work well in one context and fail elsewhere (Banerjee and others, 2017; Deaton, 2010; Madrian, 2014; Muralidharan and Niehaus, 2017). The lack of systematic high-quality evidence may be particularly dire in the field of climate change interventions, where rigorous evidence has only recently started to grow (Prowse and Snilstveit, 2010).

Our synthesis is a step towards learning more about transformational change. We do so by searching for robust and causal evidence of transformational change and its drivers. For the purpose of this synthesis, we followed the aforementioned work by GCF-IEU and operationalized "transformational change" as large effects (depth of change) over time (sustained change) targeting many beneficiaries or covering large areas (scale of change). We applied these criteria to a systematic search of two sectors that showed considerable transformational potential. On the one hand, we approached the search *directly* by systematically reviewing experimental and quasi-experimental literature[3] that has the *potential* to document transformational change across a broad set of interventions and outcomes, with the focus on the energy sector in developing countries (for which we use the proxy of low and middle-income countries). On the other hand, we also approached this learning exercise *indirectly*

[3] See section III.A.5 for experimental and quasi-experimental methods that satisfy the study design criterion for inclusion in this synthesis.

by reviewing the evidence on behavioural change in the public health sector. The public health literature has the longest tradition of long-term causal studies on behavioural science.

The goal behind our study is to identify lessons about transformational change in energy and behavioural change in public health (in terms of interventions that led to large and sustained change at scale) so that these lessons can inform broader mitigation and adaptation investments. This synthesis, therefore, combines in a novel way two different syntheses into one learning exercise on transformational change. However, in doing so, it focuses only on experimental and quasi-experimental studies.

The primary research question guiding this report is: What are the attributes, determinants, and contributors of transformational change in the energy and public health sectors? The effort is to map and systematically meta-analyse multisectoral evidence from experimental or quasi-experimental impact evaluations.

Transformational change, as such, is difficult to assess for at least five reasons:

First, transformational change consists of several elements that are typically not the outcomes measured by empirical studies. Instead, evidence for transformational change may be found across a wide range of potential outcomes, some of which are not possible to quantify.

Second, restricting the examination of studies that "find" transformational change (that is, large effects at scale and sustained over time) risks finding only statistical outliers, rather than an unbiased reflection of the available evidence. Therefore, it is important to search for evidence across a wide range of interventions and outcomes in studies that have the *potential* to document transformational change, regardless of whether the individual study indeed highlights large effects over time.

Third, causal drivers for transformational change have several limitations that are inherent to its methodological requirements. For a start, transformational change can be advanced in multiple ways, including through incremental changes that eventually combine with other factors to reach a tipping point, as well as through more rapid or sudden shifts. It is, thus, not always clear to what extent an intervention has been, or is, currently on a pathway to being transformational. Relatedly, experimental and quasi-experimental methods, seeking to explain causality in such complex intervention contexts and multiple outcome areas, are not always well-suited to capturing such changes, in terms of the breadth, depth, and level of nuance required.

Fourth, despite the availability of high-quality evaluations, finding evidence of transformational change—as defined here—put additional requirements on the data, such as the scope and time frame, which only a limited number of studies satisfy.[4] At times, these requirements are not well-described, even if they are met. Lastly, the focus on developing countries means missing out on evaluating such successful interventions as the Clean Air Act in the United States. Therefore, as in any systematic review, the gaps in evidence that this report reveals need to be seen in light of its inclusion criteria. These are described in more detail in section III.A.

C. DRIVERS OF CLIMATE CHANGE

In the following paragraphs, we describe why studying the energy sector in developing countries is key to future mitigation efforts and also highlight the need for the energy sector to adapt to climate change.

[4] As we note in section III.A.3, we did face a trade-off between the need to find sufficiently long-lasting effects and the minimum required number of studies from a statistical perspective. A comprehensive recent discussion of the difficulties in finding or assessing transformational change can be found, for example, in Williams and others (2021).

Carbon dioxide, methane, nitrous dioxide, and fluorinated gases (F-gases) are the key GHGs emitted by human activities, with carbon dioxide (CO_2) contributing to 76 per cent of emissions alone. Of the overall GHG emissions, fossil fuels and industrial processes account for 86 per cent.[5] In terms of economic sectors, energy accounts for around 35 per cent of GHG emissions, including emissions that occur in the middle stages of energy production, such as fuel extraction, refining, processing, and transportation (Intergovernmental Panel on Climate Change, 2014). Global warming is a consequence of the lagged, cumulative effect of GHG emissions. Such gases stay in the atmosphere for up to a century, such that on a *per capita*, historical basis, industrialized countries [that is, Annex-1 countries party to the United Nations Framework Convention on Climate Change (UNFCCC)] bear the preponderant responsibility for such pollutants.

That said, nearly all the future growth in energy demand, and consequently, fossil-fuel use and GHG emissions, is predicted to come from developing countries (Wolfram and others, 2012). Part of this increase may, in itself, be driven by climate change. With rising temperatures, developing countries, for example, are expected to increasingly use air conditioners, with the demand for residential air conditioning projected to rise from 500 terawatt hour (TWh) in 2000 to around 4,000 TWh in 2050 (World Energy Council, 2015). The reliance of many countries on fossil fuels for energy production means the projected increase in energy demand will, without strong counter-measures, result in even higher GHG emissions (Ebinger and Vergara, 2011). For the period of 1994 and 2014, Falconí and others (2019) found considerably higher growth rates of per capita CO_2 emissions in middle-income compared to high-income countries (HICs), with -0.2 per cent for the latter compared to 2.8 per cent for upper-income countries and 1.4 per cent for LMICs. Similarly, the per capita energy-use growth rates of upper- and lower-middle-income countries are nearly 24 times and nine times that of HICs, respectively. The contrast between the responsibility of Annex 1 countries for historical emissions and that of non-Annex 1 countries for future emissions is why climate change is such an intractable problem and multilateral climate cooperation so fraught. It also shows why the energy sector in developing countries plays such a key role for mitigation measures.

At the same time, the energy sector itself is vulnerable to climate change. Changing precipitation and weather patterns directly affect renewable energy plants that are dependent on natural activities. Hydropower plants can suffer from drying rivers, while wind power plants produce less energy if there is a windless drought, and the output of solar panels will decrease due to higher precipitation and increased cloud cover (Ebinger and Vergara, 2011). Since developing countries are predicted to be the main victims of climate change, because of the strength of the impacts and their limited capacities to adapt (Cole, 2008), adaptation measures in the energy sector are, therefore, particularly important for these countries.

Despite their vulnerability to climate change, developing countries also have opportunities to implement effective adaptation and mitigation strategies. For example, according to IEA's analysis in its *World Energy Outlook 2019*, Sub-Saharan Africa could achieve significant industrialization and economic growth, while keeping emissions relatively low, by increasing the share of renewable energy in the energy mix. To achieve this, IEA calls for investments on-grid expansion, reinforcement, and maintenance, as well as renewable energy-generating capacity, particularly solar photovoltaics (PV). Moreover, IEA (2020b) offers a scenario for net-zero economies by 2050—the epitome of transformational change. Turning to adaptation, the Global Commission on Adaptation (2019) highlights how investments in climate change adaptation could generate high rates of return and pay out a "triple dividend" of avoided losses, economic benefits (for example, through reduced

[5] Other gases are less dominant, however, still heavily present: 16 percent of emissions are methane, with six percent nitrous dioxide and the remaining two percent consisting of F-gases. Relative contributions to GHG emissions are based on the measurements of gigaton carbon dioxide equivalent per year, weighed by Global Warming Potentials with a 100-year time horizon (GWP_{100}) from the IPCC Second Assessment Report (Edenhofer and others, 2014).

climate risk), as well as social and environmental co-benefits. Through our synthesis, we study which interventions indeed show robust and causal evidence across individual studies of transformational change in energy (that shows a direct connection to climate interventions) and behavioural change in the public health sector (from which lessons can be drawn to inform climate interventions).

D. CATEGORIZATION OF INTERVENTIONS AND OUTCOMES

The interventions and outcomes covered in this synthesis are categorized within two broad theories of change (ToCs) (see Appendix 1 and Appendix 2) for each sector. These ToCs simultaneously structured and defined the scope of this study, which was important in light of our goal to broadly search for interventions that could produce transformational change.

For the energy sector, we covered a broad set of interventions that either target, or could have effects for, climate change mitigation and adaptation. They take place either at the level of institutional and market systems through incentives and standards, that is, "soft" interventions (nudges), or investments in the infrastructure. Outcomes under the purview of this report capture climate change mitigation, adaptation (resilience of energy systems), labour-market co-benefits from investments, or the transition into renewable energy. They are described in more detail in the next section.

For the public health sector, we included interventions targeting behavioural change in five broad areas: nutritional (dietary) habits, physical activity, substance abuse, hygiene practices, and the utilization of health care services.[6] The current scope of the targeted areas for the interventions in the health sector is as shown in the inclusion and exclusion criteria (Appendix 3). They were redefined to align with the preferred scope of the study and coded according to a behaviour change framework from Michie and others (2011) that we turn to now (section 1 below) before addressing the energy sector.

The two ToCs include moderating factors that interventions function through and the assumptions that influence the overall relation between the interventions and their potential outcomes. Specific assumptions on which the causal chain between interventions and long-term goals rest are presented below:

- Individuals are responsive and engage with the intervention as envisioned.
- Interventions are relevant for the context or have been contextualized appropriately.

Finally, the two ToCs map out the consequent outcomes and long-term goals targeted by interventions, as shown in Appendix 1 and Appendix 2. The sector-specific ToCs are now described below.

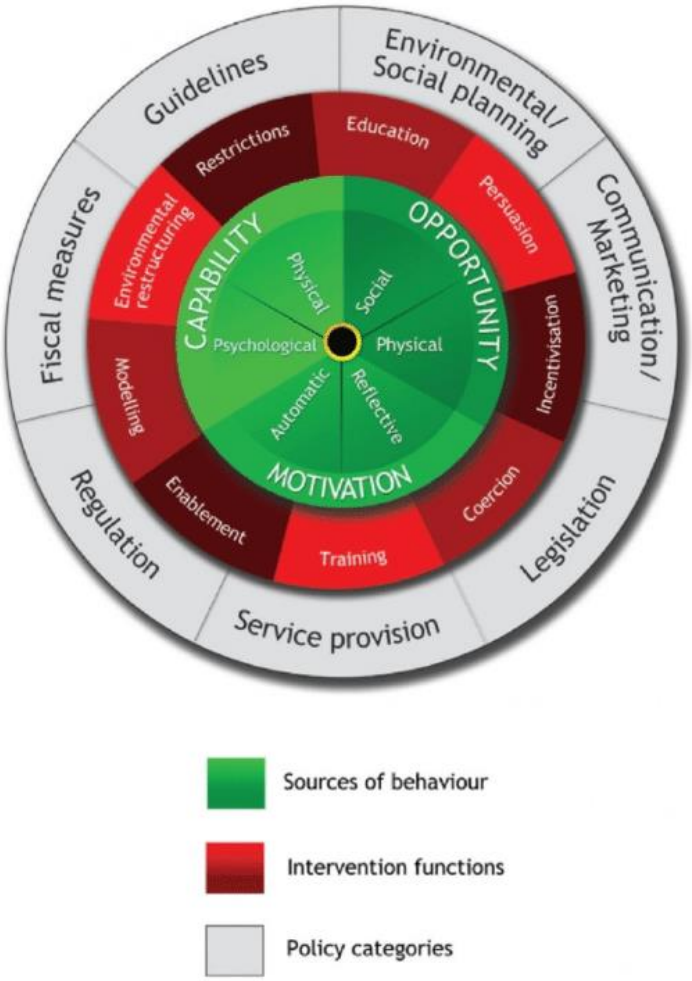
1. PUBLIC HEALTH

The public health sector has the longest tradition of using causal methods to investigate interventions that may produce large and sustained behavioural change at scale. Within this synthesis, we used this tradition to highlight the key interventions that may elicit sustained behaviour change in individuals, within the five areas of *nutritional (and dietary) habits, physical activity, substance abuse, hygiene practices,* and the *utilization of health care services.*

Each intervention is categorized under the behavioural framework from Michie and others (2011), as illustrated in the figure below.

[6] These five broad areas are not final; they are subject to change, based on the final search results and their appropriateness for the objective of the review.

Figure I-1. Behaviour change wheel



Source: Michie and others (2011)

Michie and others (2011) rely on expert consultations, as well as a review of a range of other behavioural frameworks, to define a framework for categorizing interventions and policies. This framework they call the "behaviour change wheel" groups interventions along nine *intervention functions*.[7] Behind these functions lie three essential sources of behaviour change: capability, opportunity, and motivation, or the *COM-B system*. These source functions are effectively the drivers of behaviour change: without one (or more) of these being targeted, behaviour change is not possible. Each of these sources is broken into two additional categories. Within capability, there exist *psychological* and *physical* capacities to allow the individual to engage in the activity promoted/inhibited by the intervention. Similarly, without *social* and *physical* opportunities, which lie outside of the control of the individual, behavioural change may not be possible. Both capability and opportunity also provide the necessary stimuli to the brain processes that motivate behavioural change, either by *reflection* or *automatically*. All these sources informed the design of the intervention, as depicted within the causal chain. The nine categories of intervention functions included within the behaviour change wheel are meant to contribute towards long-term change in health behaviours. The categories of intervention functions are not mutually exclusive, and in many cases, they naturally complement and reinforce one another in a particular intervention design. To

[7] They also categorize seven policy categories (communication/marketing, guidelines, fiscal, regulation, legislation, environmental/social planning, and service provision), but these are not included in our theoretical causal chain of results.

contextualize these intervention functions and increase the clarity to the reader, we offer relevant examples of climate interventions.

Interventions under the category of *education*, such as awareness and knowledge campaigns, are used to increase knowledge or understanding, not only to inspire a particular behaviour, but also to provide knowledge about competing behaviours. Education is central to many adaptation interventions, such as improving the resilience of agricultural production techniques or housing in flood-prone locations.

The second category of interventions falls under *persuasion*, whereby through various methods of communication, such as reminders or warnings via phone or other information and communication technologies, positive or negative feelings are induced to stimulate action. Such approaches can be used to increase the uptake and use of climate information portals and systems.

Incentivization, in the form of monetary and in-kind rewards, is the third category of interventions. It is meant to create reward expectations for complying with a particular behaviour or abstaining from it (for example, rewarding energy efficiency measures through a financial incentive structure).

Fourth, the category of interventions termed *coercion*—the opposite of incentivization—creates an expectation of punishment, such as by raising prices or increasing taxes (for example, through the imposition of higher road taxes for older, inefficient vehicles).

The fifth type of intervention is *training*, where individuals are imparted skills to encourage the behaviour of activity being trained. As with education, training is central to adaptation interventions, including the use of climate information systems, climate-smart agriculture, and standard operating procedures for disasters.

Restriction—prohibiting engagement in the targeted behaviour with the use of rules, such as bans or regulated uses—is the sixth category of interventions. By discouraging competing behaviours, these can also be used to encourage a particular behaviour. An example here is the restriction on the use of chlorofluorocarbons—production inputs that not only contribute to the thinning of the ozone layer but also to climate change.

Another set of interventions falls under the category of *environmental restructuring*, where modifying the *physical* context around an individual, such as improving the infrastructure or technologies related to the targeted behaviour, can encourage or discourage behavioural change. An example here is the use of the smart design of cities to change everyday behaviour, especially in relation to transport and energy use. Another subset of interventions under this category captures the modification of the *social* context around the targeted behaviour, such as prompts that guide behaviour change. Here, social norms and civic pride can play an important role in interventions, for example, protecting natural resources at risk from climate impacts.

The penultimate category of interventions is *modelling*, where behavioural change is stimulated by depicting what the model behaviour should be. This is the method of leading by example, that is, by showcasing the model behaviour. An example here is the use of “lead” farmers when implementing climate-resilient agricultural activities.

Finally, under the category of *enablement*, any type of support that increases the means, reduces the barriers, or increases the capability to act on targeted behaviour will be included. An example here is providing access to on-grid electricity in rural areas.

Within our synthesis, we further divided the intervention function of environmental restructuring into its two subcategories—*physical restructuring* and *social restructuring*, thus giving us 10 intervention functions in total.

Each intervention function affects one or more source function(s), and thereby leads to the required modification of health behaviours, attitudes, and practices—depicted as the *concrete outcomes* in

our theory of change (ToC). Overall, these interventions aim at sustained improvements in health behaviour, infrastructure, and practices.

These interventions that were targeted at changing behaviour in the outcome categories also defined the scope of the public health sector. For the purpose of this evidence synthesis, we are not interested in all possible health outcomes, but rather in what we can learn from these health behaviours for behaviours related to climate change mitigation and adaptation. We, therefore, defined the scope of the health outcomes along the following dimensions: *action/health-seeking behaviour* and *purchasing/consumption behaviour*. These two dimensions can have a private benefit (quitting smoking), or might alternatively, also affect health outcomes for other individuals (because of less exposure to passive smoke).

2. ENERGY

A large number of the interventions targeted at climate mitigation can be found in the energy sector. Due to the implementation of the Paris Agreement, 197 countries are required to have national GHG emission-reduction policies and plans for their post-2020 agenda (World Resources Institute, 2018). Fostering low-carbon technologies is, therefore, projected to be a major issue for governments (Bouyé and others, 2018).

The long-term goal of the ToC, as it pertains to the energy sector, is that even as energy production and consumption remain sustainable and resilient, they do not contribute to climate change. Moreover, the increase of energy supply and demand also seeks to contribute to higher employment that can be a social co-benefit of energy investments. For this sector, we based the ToC mainly on different assessment reports and systematic reviews concerning climate change mitigation and adaptation, especially IPCC's synthesis report on climate change (Intergovernmental Panel on Climate Change, 2014); frameworks and reports by the WB Group, IEA, and the European Union Energy Initiative Partnership Dialogue Facility (Ebinger and Vergara, 2011); as well as extensive discussions with CIF E&L Initiative and GCF's IEU.

We grouped interventions into four broad categories: they jointly contribute to achieving the long-term goals according to the ToC.

The first category is *institutional and market systems*, that is, interventions that change the institutional structure of energy systems or markets. The subcategories are public administration reforms, industry coordination and industry self-regulation, privatization, liberalization, and the introduction of market-based mechanisms, as well as de-privatization and de-liberalization.

The second category is *incentives and standards*. This category consists of three subcategories that are directly linked to the behavioural framework from Michie and others (2011), as described in the public health sector above: incentivization (such as transfers), coercion (such as taxes and fees), and restrictions (such as bans and limits).

The third category is the *"soft" interventions* that do not change incentives. The subcategories therein are again taken from Michie and others (2011): education, persuasion, training, environmental restructuring (such as social norms), modelling (such as presenting model behaviour in TV shows), and enablement (such as defaults).

Lastly, *investments in energy infrastructure, equipment, and technologies* are the fourth category. Subcategories are investments in energy transmission, distribution, and the storage of electric energy systems, as well as investments into renewable energy-generating equipment.

These interventions may lead to outcomes that are grouped under seven categories.

First, mainly through investments into energy infrastructure, such as grid extensions, *access to energy and the supply of (renewable) energy* may increase.

Second, *energy market development* may be spurred through institutional and market-system interventions (International Finance Corporation, 2019).

Third, *energy consumption and demand* (differentiated between renewable, non-renewable, and on-grid electricity), along with fourth, the *adoption of more energy-efficient technologies (including the transition to renewables)* may change due to targeted interventions in all intervention categories.

Fifth, the *resilience of energy systems* to climate change may increase due to investments into energy systems, such as smart grids and energy storage capacities (Ebinger and Vergara, 2011; Stuart and Escudero, 2017).

Sixth, as a result of incentives and standards (such as energy efficiency standards), as well as cleaner energy supply and demand and the adoption of more energy-efficient technologies (such as improved cookstoves), *GHG emissions and indoor air pollution* may decrease.

Lastly, as a *labour market co-benefit* from investments into renewable energy, formal employment may increase.

In order to facilitate cross-sector learning, the behavioural outcomes within these seven outcomes were coded in terms of whether they are *action behaviours* or *purchasing/consumption behaviours*.

E. WHY IT IS IMPORTANT TO DO THIS EVIDENCE SYNTHESIS

The concept of a transformational change is embodied in Article 2 of the Governing Instrument of the Green Climate Fund as "the Fund will promote a paradigm shift towards low emission and climate-resilient development pathways". However, to our knowledge, there appears to be an absence of the mapping of systematic evidence on the causal drivers of transformational change in general, and in particular, in relation to mitigation and adaptation. The study that is closest to our review is done by Lee and others (2013): it systematically reviews the literature on organizational transformation, mainly in health care. Nonetheless, there are two key differences. First, their definition of “transformational change” is focused on organizational practices, whereas we looked at a broader range of outcomes. Second, most of the studies included in their review are qualitative, whereas this study was purposefully confined to quantitative experimental and quasi-experimental studies. As such, this evidence synthesis will address this gap within the literature on the available causal evidence from experimental and quasi-experimental literature on a broad set of interventions and their effects on outcomes.

We contribute to the literature on the drivers of transformational change in the following ways:

- We discuss the attributes of transformational change by offering a precise definition of “transformational change” for the purpose of our review of experimental and quasi-experimental literature (see next section).
- While our evidence synthesis is broad in scope, we covered a precise, but extensive, list of interventions and outcomes within each sector as well as within clearly structured categories. This approach allowed us to search for evidence for transformational change across fields of studies, while at the same time, keeping the scope manageable.
- Our synthesis combines EGMs with systematic reviews in the two sectors. EGMs are a convenient and simple-to-use tool for policymakers to quickly inform themselves about existing quantitative causal evidence to make informed decisions on further experimental and quasi-experimental research activities. This exercise highlights where research is comprehensive and where there appears to be a lack of such evidence.
- In order to learn about causal evidence on transformational change, the systematic reviews consist of a quantitative synthesis through meta-analyses only on quantitative studies with an

experimental or quasi-experimental study design. Furthermore, our inclusion criteria are based on our own precise definition of "transformational change". More specifically, we only included studies that had the potential to document transformational change according to these criteria. For instance, we only included studies where data collection was done at least one year after the intervention. Whether transformational change indeed happened was the empirical question that we addressed through our meta-analysis, acknowledging nonetheless the limitations noted above and throughout the paper.

- We conducted meta-analyses with the data extracted from the selected quantitative studies for sufficiently populated areas of the EGMs. This is another exercise that has not been found to be common in the literature on transformational change. We do this with caution, given the limitations of this approach and its application to a complex concept such as "transformational change". The results of the meta-analysis can help determine where robust quantitative evidence from experimental and quasi-experimental literature exists, that is, across individual studies and contexts, for transformational change. Doing so minimizes the risk that large effects of interventions are simply statistical outliers. By using the results from this meta-analysis, we present only those intervention and outcome combinations where the evidence for large effects at scale and over time exists.
- We combine two systematic reviews in two different sectors—public health and energy—into one learning exercise about transformational change and explore how the lessons learned may inform climate change mitigation and adaptation investments.

II. OBJECTIVES

This evidence synthesis focuses on the evidence from developing countries to learn about the attributes, determinants, and contributors of transformational change in the energy and public health sectors.

For the purpose of our study, we had to operationalize transformational change into clearly measurable criteria. We did so with the following three criteria:

1) A large **depth** of change: Transformational change requires a sizeable change. For the purpose of this synthesis, it was measured, in terms of the effect of the size an intervention had produced on the outcomes.

   To define what a large impact is, we relied on previous literature that had attempted to standardize these definitions. Sawilowsky (2009) defines rule-of-thumb effect sizes for Cohen's d as large if d =.8 (and very large, if d = 1.2, and huge for a d = 2), based on a literature review and a contextualization of the effect sizes, as defined in Cohen (1988).[8] Cohen's d is a relative measure of effect size—by putting the effect size in relation to its standard deviation. In addition, where the measurement of effects (in the whole sample or a subsample of studies) is done based on the same outcome measures and permits direct comparisons, we also compared and discussed effect sizes in original units.

2) A large **scale** of change: The likelihood for an intervention to become an important contributor to transformational change increases with scale, that is, targeting many beneficiaries or covering large areas.

   Given the variety of interventions within the two sectors, we considered interventions as large-scale, if there are at least 1,000 individual beneficiaries,[9] or if they targeted an entire administrative area larger than a village (such as district, region, or state).

3) **Sustained** change: For a change to be transformational, it has to persist over time. The definition of "sustained" is found to vary considerably over the literature we reviewed (between six months and several years). In order to maintain coherence across studies, we considered an effect to be sustained, if it has persisted at least one year after the first full implementation of the intervention. Note that this is a lower bound; thus, impacts arriving later than one year were also considered to pass this threshold.

   The question of whether impacts are likely to wane or increase over time may be different between the two sectors. In public health, the time required to form a habit may be relatively short (Lally and others, 2010) while stopping adoption of behaviours is a strong concern. Therefore, behavioural change is unlikely to happen, if it is not already present after one year. By contrast, many energy interventions may take a long time to demonstrate impacts, with many ups and downs. We, therefore, acknowledge that the absence of large impacts in the energy sector after one year does not imply that large impacts may not arrive later. This is a concern for studies that measure outcomes not longer than one year after the onset of the intervention. While studying longer time frames than one year would certainly be useful, setting the threshold higher would risk leaving out many studies.

The report discusses these thresholds for the depth of change, scale of change, and sustained change, in light of the available evidence that we found.

[8] For relative risk, common in the medical literature, the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) uses a scale separating relative risks of at least 2 as large (and those greater than 5 as very large) (Guyatt and others, 2011). We will use these two definitions to define large impacts based on effect sizes in outcomes.
[9] The effect being measured here is the so-called "treatment effect on the treated".

III. METHODS

A. CRITERIA FOR CONSIDERING STUDIES FOR THIS EVIDENCE SYNTHESIS

To preview the precise inclusion criteria of individual studies described through the PICOS model (population, intervention, comparison, outcome and study design model), we only looked at studies that meet the following criteria:

- measured effect sizes (allowing us to search for interventions that produce a large **depth** of change)
- sufficiently large in their scope or coverage (reflecting the possibility for a large **scale** of change)
- collected data over a prolonged period in time (reflecting the possibility of **sustained** change)

Note that these three criteria are closely linked to the definition of "transformational change". However, the criteria only helped in finding studies that have the *potential* to document transformational change.

Crucially, individual studies were also selected for inclusion, if the evaluated intervention did not lead to large effects over time. If we only included studies with large effects over time, we would run a strong risk of picking statistical outliers, instead of finding an unbiased picture of the available evidence from experimental and quasi-experimental studies.

Following Petticrew and Roberts (2006), we used the PICOS model to precisely describe the inclusion and exclusion criteria. The tables, including the summary of the inclusion and exclusion criteria for both sectors, can be found in Appendix 3 and Appendix 4. Pilot screening led to minor adjustments compared to the approach paper and protocol.

The inclusion and exclusion criteria apply to both the EGMs and the systematic reviews. Studies for the systematic reviews were selected from the populated cells presented in the EGMs. More information can be found in section C.2.

1. TYPES OF PARTICIPANTS (POPULATION)

We included interventions rolled out in developing countries and use LMICs as defined by the current WB categorization (2020) as a proxy. Thus, we excluded studies of interventions in HICs, or that included developing countries but did not separately report results for these countries. In the energy sector, we excluded interventions targeted solely at children (below the age of 12) because they are generally not the main agents of climate change mitigation and adaptation. Included interventions, therefore, targeted adults, adolescents, and any form of social systems (such as firms). In the health sector, interventions that targeted the behavioural change of adolescents or children (below the age of 18) were excluded because we wished to study long-lasting behavioural change, about which we may learn more general lessons from adults who have more solidified personalities than adolescents.[10]

[10] While it would have been interesting to compare adults and adolescents, it is beyond the scope of this report.

2. TYPES OF INTERVENTIONS

The types of interventions we studied were informed by the sector-specific theories of change described in the background section and enclosed in Appendix 1 and Appendix 2. We focused on studies that seek to evaluate the causal effect of an intervention that was purposefully implemented.

Furthermore, we focused only on interventions that were sufficiently large in scale in order to draw meaningful conclusions. Results should be representative of a large-scale intervention, following Muralidharan and Niehaus (2017). First is the scale of the intervention: there needs to be at least 1,000 beneficiaries (automatically fulfilled, if there are more than 1,000 "treated" individuals in the study sample). If the number of beneficiaries is not given, or in case the intervention is disseminated through radio and other media, it needs to target an entire administrative area larger than a village (such as a district, region, or state). Second is the scale of the population represented: we aimed for studies, whereby the sample of treated individuals is representative of a sampling frame of at least 1,000 treated individuals, or an administrative area larger than a village. While truly "large scale" (to reduce the familiar upward "bias" of small-scale interventions and studies) would mean a threshold higher than 1,000 beneficiaries, the criterion was purposefully set low initially in order to minimize the risk of excluding too many studies.

3. TYPES OF OUTCOME MEASURES

Since our major outcome - transformational change - cannot be directly measured, we looked at a range of outcomes and measured change therein to be later synthesized through the meta-analysis. At this stage, we determined whether evidence for transformational change from experimental or quasi-experimental studies could be found. Our list of outcomes is based on the ToCs described in the background section (I.D.1 and I.D.2) and provided in the Appendix 1 and Appendix 2. In order to see whether the effects were long-lasting, the outcomes of the studies must have been measured at least one year after the full delivery of the intervention. While collecting data one year after the intervention is, in many cases, not sufficiently long to be certain of a sustained change (for example, by enduring changes in political or administrative leadership), a higher threshold would have led to the exclusion of too many studies.

4. COMPARISON

We considered only quantitative studies that aim to evaluate the causal effect of an intervention on the outcome (that is, experimental or quasi-experimental studies). In addition, we covered studies that have a clearly defined comparison group to allow for the evaluation of the effects of the intervention. The nature of the comparison group depends on the type of research design used in the study and could include both active and passive comparison groups. As noted previously, taken together, these parameters significantly limited the range of available evidence included in this synthesis, particularly in relation to energy systems interventions; as such, the findings and conclusions must be considered with these limitations in mind.

5. STUDY DESIGN

Based on the research design, we categorized the studies under two major groups:

- Experimental designs: This type of study specifically uses the random assignment of intervention to the treatment group and evaluates the effect by comparing the outcome with a comparison group by using an appropriate methodology.
- Quasi-experimental designs: These methods include and are restricted to regression discontinuity design (RDD), instrumental variable, difference-in-differences (DiD), propensity

score matching (PSM), and Heckman selection models. For the health sector, in addition to these aforementioned methods, we also used interrupted time series and controlled before-after studies, given their relevance in the health literature.

6. EXCLUSION CRITERIA

We excluded studies conducted outside the 1990–2020 time frame in the energy sector and before 2000 in the public health sector, or those that do not include a separate sample from developing countries. We also excluded studies that do not attempt to evaluate the causal effects of the intervention on the outcome, particularly those that do not follow the methods explained in the study design. As mentioned above, we excluded studies that are not sufficiently large-scale or long-term (as defined before). Studies that have not met all our inclusion criteria (such as both covering a relevant intervention and a relevant outcome) were also excluded.

B. SEARCH METHODS FOR IDENTIFICATION OF STUDIES

1. DATABASES, REPOSITORIES, AND INDIVIDUAL JOURNALS

The choice of databases was guided by the relevance and comprehensiveness of their coverage of the sectoral literature. Naturally, these databases are not identical for the two different sectors; and thus, we searched these sources based on their appropriateness.

a. Public health

1. *Databases:*
    a. EconLit (via EBSCO)
    b. Global Health (CAB- Ovid)
    c. Medline (Ovid)
    d. Web of Science (Social Sciences Citation Index)
2. *Websites of agencies and research institutes:*
    a. Campbell Collaboration https://campbellcollaboration.org/
    b. Cochrane Database of Systematic Reviews https://www.cochranelibrary.com/cdsr/
    c. Collaboration for Environmental Evidence https://www.environmentalevidence.org/
    d. International Initiative for Impact Evaluation: 3ie Development Evidence Portal https://developmentevidence.3ieimpact.org/

b. Energy

3. *Databases:*
    a. Academic Search Complete (via EBSCO)
    b. CAB Abstracts (via EBSCO)
    c. EconLit (via EBSCO)
    d. GreenFILE (via EBSCO)
    e. Web of Science (Social Sciences Citation Index, Science Citation Index Expanded, Emerging Sources Citation Index)
    f. World Bank eLibrary (via EBSCO) [11]

[11] We used simplified search strings for the EBSCO World Bank eLibrary database to get non-zero results.

4. *Websites of agencies and research institutes:*[12]
    a. African Development Bank https://www.afdb.org/en
    b. Asian Development Bank https://www.adb.org/
    c. Campbell Collaboration https://campbellcollaboration.org/
    d. Collaboration for Environmental Evidence https://www.environmentalevidence.org/
    e. 3ie Development Evidence Portal https://developmentevidence.3ieimpact.org/
    f. National Bureau of Economic Research https://www.nber.org/
    g. Swedish International Development Cooperation Agency[13] https://www.sida.se/English/
5. *Key journals:*[14]
    a. Energy Economics
    b. Energy Journal
    c. Energy Policy
6. *Snowball searches:*

We included targeted snowball searches (cited and citing the literature of included studies) to ensure that the selection of papers included in our meta-analysis was complete.

2. SEARCH TERMS

The search terms were organized in six different categories that reflected the inclusion criteria and the sector-specific theories of change. The search terms within each category were combined with the "OR" Boolean operator, whereas the "AND" operator was used to combine the different categories of search terms.

1) Long-term or large-scale: This category encompasses terms used to describe studies carried out over a longer timespan or at a large scale.
2) Methodology: These terms capture the experimental and quasi-experimental methods (for more details, see the inclusion/exclusion tables in Appendix 3 and Appendix 4).
3) Countries: As a proxy for developing countries, all LMICs, as well as general terms describing LMICs, are listed here.
4) Interventions: Terms were based on sector-specific ToCs.
5) Outcomes: Terms were based on sector-specific ToCs.
6) Sector-specifying terms: This category contains terms used to describe the respective sectors.

In the energy sector, five categories (numbers 2 to 6 in the list above) were combined through the "AND" operator because the total number of search results was below the target of 7,500. In public health, the number of studies was well above the target of 15,000; therefore, the long-term or large-scale search terms were applied (category 1), in addition to the four categories (numbers 2 to 5 in the list), to narrow down the searches.

[12] We ran simplified searches on institutional websites, as they do not permit the use of sophisticated search operators.
[13] We included only one bilateral agency website, because during preliminary searches, we identified it as having potentially relevant impact evaluation studies. We excluded the websites of other shortlisted bilateral agencies due to a lack of relevant studies.
[14] We ran a database search in the Web of Science platform with the simplified set of search terms: three of the six categories (methodology, countries, and interventions) were combined through the "AND" operator with the Publication Name terms string: SO=("energy economics" OR "energy journal" OR "energy policy") to find studies in three energy journals with the highest impact factors. The highest-impact journals relevant for this synthesis were selected from the list available at the Scimago Journal & Country Rank for energy.

The search strategies were tested against a set of benchmark studies in each sector.[15] If more than two-thirds of the benchmark papers could be retrieved through the database searches, the search strategy was deemed satisfactory. For the energy sector, the final version of the search strategy captured 83.33 per cent of the benchmark studies; in public health, the target was just slightly above the threshold at 68.75 per cent.

The exact comprehensive search strategies and search results for one of the databases—the Web of Science—are provided in Appendix 5 and Appendix 6. This is the same search that was then adapted to the search algorithms of the other databases.[16]

C. Data collection and analysis

1. Screening and selection of studies for the Evidence and gap maps (EGMs)

The Africa Centre for Systematic Reviews & Knowledge Translation carried out the screening process of the two populations of studies that were included in the sector-specific EGMs.

Note that screening was done in several steps for each sector such that there were two separate screening processes:

First, pilot screening aimed to make sure that the coding tools were well understood and adapted where necessary. Titles, abstracts, and keywords were screened to exclude any irrelevant studies. The screening and coding of each study was done by a pair of coders (double coding) in EPPI Reviewer 4—a web-based software programme for managing and analysing data in literature reviews (https://eppi.ioe.ac.uk/EPPIReviewer-Web/). Two independent coders each screened 200 studies. The results of the pilot screening were considered satisfactory once the overlap between the inclusion decisions of both screeners after reconciliation was above 80 per cent.

Second, the post-pilot screening of studies was conducted. In order to save time, given the wide scope of the literature search, this stage was assisted by the priority-screening algorithm embedded in EPPI Reviewer 4. The pilot screening entries of the 200 studies initially trained the algorithm; it was then progressively updated in the course of screening the studies.

Third, we applied the specified inclusion/exclusion criteria to the full text and determined whether the study should be included for analysis. We recorded all search results, including the reasons for exclusion, at the full-text screening stage. These results are presented in the PRISMA diagrams. At least 20 per cent of the studies were double screened by a second reviewer. Disagreements were resolved through discussions with the core Center for Evaluation and Development (C4ED) team that also provided additional clarifications.

Systematic reviews were screened on the basis of their inclusion criteria. If the systematic review met all of our inclusion criteria, it was passed on for data extraction and shown in our EGMs.

2. Selection of studies for the systematic review

One crucial step of our methodology was pooling similar studies for the meta-analysis. Given the wide scope of the two sectors, the similarity was difficult to define ex-ante. For the purpose of the

[15] Benchmark studies are studies that, according to their titles and abstracts, are clearly relevant for our review. As the search strategy relied on searching in title, abstract, and keyword fields, finding these studies is a good measure of the quality of the search strategy. The large majority of benchmark studies also satisfied all of our inclusion criteria (as specified in the inclusion and exclusion tables in Appendix 3 and Appendix 4), and would therefore, also pass the full-text screening stage.

[16] As noted above, for the WB e-Library database, the search strings had to be simplified to get non-zero results.

statistical synthesis, the similarity between the studies was defined conceptually and statistically. In order to be selected, studies need to share an underlying conceptual hypothesis (Cooper, 2010) and measure outcomes in statistically comparable and convertible ways.

As our synthesis covers a wide range of interventions, outcomes, and research questions, we did not know in advance which of the studies that passed our inclusion criteria would share a conceptual hypothesis. We, therefore, approximated similarity through the location of the studies within our EGMs, assuming that the studies that fall into the same cell are conceptually similar and statistically comparable.

As the EGMs show, there is, however, considerable heterogeneity between studies even within cells. Furthermore, there are different possibilities for classifying studies, thus leading to different potential EGMs. Therefore, instead of just mechanically selecting populated cells, they only served as a starting point. We then investigated the cells and assessed the ones where we found a sufficient number of studies that were similar—conceptually and statistically.

Here, we describe the necessary steps on how the sets of studies were pooled for the meta-analysis.

From a conceptual point of view, we asked whether a given set of studies could be grouped under one overarching conceptual hypothesis, that is, we tested the effectiveness of the same well-defined intervention on comparable outcomes. The starting point was the studies within populated cells. In the case of the energy sector, we thought that cells at the subcategory level had the best chance of finding a common conceptual hypothesis. As for the health sector, two levels were considered—first, the level of the subsector within health, and thereafter, within those subsectors, all intervention function levels. At this level, we first identified populated cells, and then compared studies based on their titles, to see if they address a similar question. We then identified the studies that were potential candidates for inclusion in the meta-analysis.

At the same time, we needed to identify how many studies could be pooled on statistical grounds. While all conversions among different effect measures (for example, based on Cohen's d or odds ratio) are based on statistical assumptions that may be violated in practice, it is recommended to still use conversions, rather than omit those studies that use a different effect measure (Borenstein and others, 2011). Statistical difficulties arise when the intervention effects are measured as regression coefficients because statistical models are rarely the same across different studies. Therefore, whenever possible, we favoured effect measures that are based on a direct comparison between the treatment and the control group. How we approached possible cases of synthesizing regression coefficients is discussed in section 6.

After conducting both steps, we identified all combinations of interventions and outcomes with at least 10 studies that could be pooled in a meta-analysis.

3. Data extraction and management

The final studies that were included after the full-text screening into the EGMs were then coded using a pre-piloted extraction form (see online Appendices A and B) by two coders. Disagreements in coding were resolved through discussions and third-member involvement. Information collection included those necessary to generate the EGMs, such as potential filters. Study results were not extracted at this stage.

Data extraction for studies selected for the meta-analysis was done by one coder, with extensive quality checks by another coder from C4ED. The forms used to extract data for the meta-analyses can be found in online appendices C and D. They underwent several phases of piloting.

4. Data analysis steps

The goal of the analysis is to search for evidence of transformational change. The analysis proceeded in several steps for each sector, with technical details described in other paragraphs.

- We constructed EGMs, with interventions listed along the Y-axis and outcomes along the X-axis, to document evidence and gaps within the scope of each sector.
- We then concentrated on the sufficiently populated areas within the map to do systematic reviews in the form of meta-analyses on the available evidence on conceptually similar studies, as described in section 2, and estimated overall effect sizes.
- We ran meta-analyses among comparable studies. On a statistical level, we made the studies comparable by calculating the same standardized effect sizes across studies wherever necessary and possible. We also investigated publication bias and ran sensitivity analyses on the distribution of the effects. We excluded studies that do not provide sufficient information to do so.
- We then searched for combinations of interventions and outcomes where evidence of transformational change has been found, that is, there is a large-effect size at least one year after the intervention, following the thresholds defined in section A5.1. It is this step where the results of the studies, that is, the depth of change and sustained change, are used as selection criteria. Selection was not done at the level of the individual study, but rather at the level of intervention-outcome combinations. Based on the general advice in the literature (Borenstein and others, 2011) and simulation results, 10 studies were set as the lower bound in the test for heterogeneity, and therefore, the assessment of the generalizability of the results.

5. Evidence and gap maps

In order to draw the EGMs, the following procedure was applied:

- **Categorization of studies:** We followed Rankin and others (2016) to determine the categorization of studies in the EGMs. In the case where several different interventions were grouped together, each intervention was coded separately in order to be able to show all available evidence related to a particular intervention. For example, if a study looked at the effects of a programme that included a cash transfer intervention and an awareness intervention on two different outcomes, then the two associated outcomes would be coded separately for each intervention. In some studies, it might be that only some elements of the programme or evaluation were relevant to this EGM (such as a specific intervention or outcome); then, only these aspects were extracted and coded. Systematic reviews were coded based on the PICOS of the review. If a systematic review covered more than one relevant intervention and outcome, it appeared in each cell that was applicable.
- **Generate categories based on the outcomes related to each sector:** Within an EGM, the outcomes, presented on the X-axis of the map (columns), indicate a cluster of multiple studies. These categories were generated on the basis of the outcomes, as described in the sector-specific theories of change.
- **Generate categories based on the interventions related to each sector:** The Y-axis (rows) of the EGM lists all the specific interventions that are part of the sector-specific ToC. In public health, the categories are based on the intervention functions of the behaviour change wheel. All impact evaluations and systematic reviews that use a specific intervention function would be included within that row.

6. META-ANALYSIS

The systematic reviews are based on quantitative synthesis through a meta-analysis. It is a statistical procedure used for synthesizing effect sizes from multiple studies by weighing them by their precision. A meta-analysis follows a two-stage process.

In the first stage, a summary statistic was calculated for each study to describe the observed intervention effect in the same way for every study.

In the second stage, an overall (combined) intervention effect estimate was calculated as a weighted average of the intervention effects estimated in the individual studies (Gurevitch and others, 2018).

We reported the effect estimates from the meta-analysis graphically on a forest plot: It illustrates the effect size and corresponding confidence interval for each study, along with the overall effect size of all included studies and the corresponding overall confidence interval. The examination of heterogeneity (variability in the intervention effects) and reporting biases in the study results is also a critical outcome of a meta-analysis.

a. Measures of the treatment effect

In order to do a systematic review of the studies, we used two approaches for the two sectors, given the variations in the types of studies that were included at this stage.

In the case of the energy sector, since all the papers in the two cells selected for the meta-analysis are from regression models with similar outcomes, the treatment effect is comparable for all variables and coefficients, even when unstandardized. In order to determine whether the overall effect is large or small, we also converted the effect sizes (and standard errors) to a standardized mean-difference effect size, the Cohen's d (a common effect-size metric across studies), using a standard t-stat transformation (as, for instance, done by Waddington and others (2019)).

Within the health sector, the effect sizes (and thereby, the standard errors) reported in the paper are based on different types of models, where none of the abstract outcomes are reported by one effect type. Essentially, the sample of effects, included at the meta-analysis stage, comprises either comparable variables (that is, handwashing before eating versus the frequency of handwashing in one day) measured differently (that is, odds ratio versus difference in means), or comparable methods for non-comparable effects (that is, handwashing versus immunization). Therefore, to meta-analyse the effects, we converted all effect sizes to Cohen's d, using a t-stat transformation in the case of regression coefficients. Due to the large range of measurement methods used, the conversion into Cohen's d was needed for adaptation to the measurement method.

In general, the effect types can be separated into two types: those based on a regression and those based on binary data that can be described in a two-by-two table. Instead of finding a formula for each measure individually, and risking conversions not being equivalent, we tried to use the same conversion as frequently as possible. For all measures describing binary data, we calculated the two-by-two table from the information given in the paper (and accepting that there would be rounding errors), and from there, calculated the odds ratio. We then followed Borenstein and others (2011) for the conversion of the (log) odds ratios into Cohen's d. For the effect measures based on regressions, we followed the same method as described for the energy sector at the beginning of this section. The formulas for the conversion of effect sizes into Cohen's d for both sectors are provided in the online appendix E. For the conversion, we excluded studies that do not provide sufficient information, whether it is due to missing sample sizes or insufficient statistical testing. An elaboration on the number of effect sizes and studies that were excluded, going from the EGM to the meta-analysis, is provided in section 0.

For synthesizing the raw coefficients from regressions, the following approach was implemented. We only included studies where the intervention was coded as a dummy variable, because this allowed us to have a comparably defined treatment variable across studies. Second, we then searched for the regression model that would estimate the unconditional intervention effect most precisely and with the least bias. In order to evaluate which regression model is preferable, we were guided by the opinion of the study authors. If the study authors do not make such a claim, we used the most informative model (judged by information criteria such as R-squared) that does not include any co-variates that explain the intervention effect (such as moderator variables).

For the public health sector, one additional step was carried out after the conversion into Cohen's d. After conversion, the comparable/similar outcome effect sizes within the same treatment arm were aggregated to arrive at a final data set. It includes effects measured at the treatment arm level (instead of at the outcome or study level). This step is necessary for two reasons, both related to avoiding bias in the meta-analysis later, that is to ensure that 1) we do not double-count large effects for two outcomes from the same treatment arm and study, and 2) we do not inflate the sample size related to one type of outcome versus another. To combine the comparable effect sizes, the converted estimates were averaged (to keep the sample size constant, instead of inflating it per effect size in the treatment arm) and the entire data set was collapsed to keep only the combined values per treatment arm. Therefore, in the public health sector, the effects are not considered at the study level, but at the treatment arm level.

b. Dealing with missing data

It was determined that all the studies finalized for the data extraction had to contain information that permitted the calculation of comparable effect sizes. This was a larger challenge in the public health sector, where there are a variety of different measures, meaning that the analysis needed to be done on standardized effect sizes. Even when working with non-standardized effect sizes, standard errors are necessary for calculating an overall effect. In practically all Cohen's d conversions, the sample size of treatment and control group played a role.

In order to account for this missing information, we used the following assumptions:

- Sample size in treatment or control group: In certain cases (the most problematic being the DiD regression studies), a clear control or treatment group size is not provided. In this case, assumptions were made on the basis of the sampling procedure, or by simply halving the overall sample size, where there is no indication of an unbalanced treatment or control sample.
- Back-calculation of the standard errors from t-statistics, p-values, standard deviations, and confidence intervals: Due to the heterogeneous reporting in the papers, the standard error required for the meta-analysis was missing in many observations. In order to generate this information, a back-calculation was required from the other information presented in the paper. For instance, in the case where the t-statistic is provided, the standard error was derived by dividing the beta coefficient with the t-statistic.

c. Assessment of heterogeneity

We tested for heterogeneity across studies and reported the amount of heterogeneity by the usual Tau statistic and the I-squared statistic. Tau denotes the standard error of true effect sizes in the original units, whereas I-squared measures the percentage of variability across studies that is not due to sampling error, but rather the differences in the study population, intervention, and implementation. Thus, Tau indicates the stability of an average true effect size across studies, while the I-squared allows for a rough categorization of heterogeneity (Borenstein and others, 2011). As I-squared is heavily dependent on the extent of the sampling error, and many studies we included have large samples (consequently, low sampling error), a high I-squared does not tell us much about

the heterogeneity in true effect sizes. We, therefore, focused our assessment of heterogeneity on Tau.

If substantial heterogeneity is present, we investigated what factors explained it by conducting a moderator analysis, including a subgroup meta-analysis and a meta-regression. For sufficient statistical power in meta-regressions, we followed Borenstein and others (2011) who recommend that each covariate (the coding of studies) contains at least 10 studies. Where studies are sufficiently similar to be comparable, we ran meta-regressions across sufficiently populated cells in both EGMs.

Finally, we used contour-enhanced funnel plots and corresponding regression methods (Stanley and Doucouliagos, 2014) to assess small-study effects that can arise due to the publication bias, for instance.

d. Data synthesis

The main results are presented using forest plots. In terms of software, we used the Stata meta-analysis package for the entire analysis, and the *meta forest plot* command to generate the forest plots. The meta-analysis was conducted, using a random-effects model (the *reml* option within the aforementioned Stata package), given the large heterogeneity in our studies.

For the energy sector, the analysis was carried out for the two main cells selected from the EGM stage. The forest plots were generated for the non-standardized and standardized values for the effect sizes in both cells. Additionally, subgroup analyses, outlier detection, and funnel plots were performed in order to explore the sources of heterogeneity and conducted as sensitivity tests.

For the health sector, the analysis was conducted at the level of the abstract outcomes in order to explore the impact of the intervention functions on the abstract outcomes, as defined in the ToC. Therefore, the overall effect sizes are first reported for the total sample, then for all four abstract outcomes, and then for the eight intervention functions for which data was extracted after the EGM stage. We explored heterogeneity across the various intervention functions by abstract outcome subgroups.

For the public health sector, the large number of studies also allows for a meta-regression (using the *meta regress* command in Stata) of the overall sample, samples grouped by abstract outcomes, and intervention function samples. The moderators used in each of the regressions are similar. A dummy variable for each of the regions and for the outcome measure (odd's ratio, means, hazard ratio, etc.) were included as moderators. In addition, moderators for all intervention functions were also used. For the intervention-function-specific samples (that is, those focused on education, enablement, modelling, incentives, persuasion, training, social restructuring, or physical restructuring), the other seven intervention functions (that is, all intervention functions except education, in the case of the education sample) were included as moderators.

e. Assessment of risk-of-bias in included studies

The main reporting biases we discuss in this report are those arising from publication bias. In order to test for publication bias, we used the Stata *meta bias* command that applies a refined version of the Egger's test (Egger and others, 1997), where the standardized effect sizes are regressed on their precision, as given by standard errors in this case. A significant correlation between the two shows a "small-study effect" that could be due to publication bias.

Other potential biases are tackled within the risk-of-bias assessment for each study selected for the meta-analysis—the results of which are presented in section IV.D. The assessment was undertaken, using a tool modified and pilot-tested from the Campbell Collaboration framework for assessing the quality of studies included in reviews (online appendix F). This framework applies the following six items to assess the quality of evidence in included studies: study design (to ensure that potential

confounders are considered); clarity of intervention definition; clarity in the definition of outcome measures and reliability; baseline balance tests; sample representativeness of large-scale interventions; and the precision of estimates.

IV. RESULTS

a. SEARCH RESULTS FOR ENERGY AND HEALTH SECTORS

2. PUBLIC HEALTH

The codes used for the screening in the health sector, for both the title and abstract, as well as the full text, are provided in Table IV-1 and Table IV-2, respectively. Only one of the exclusion criteria was marked to exclude each study. As shown, a total of 578 studies were finalized for full-text screening.

In the health sector, from a total 21,477 studies (column 1), 12,310 (column 2) were screened. After screening over 12,000 studies, the inclusion rate of studies dropped significantly, where even after 400 screened papers, no additional studies were coded as included.

Table IV-1. Title and abstracts screening for health sector (12,310 studies screened in total)

CODE IN EPPIR4	COUNT	COUNT (STUDIES SCREENED AT TITLE & ABSTRACT AND RECONCILED)
INCLUDE	**578**	**578**
EXCLUDE	11,715	11,715
ExDUPLICATE	17	17
None of the codes above	9,167	0
Total	**21,477**	**12,310**

Source: Authors

Table IV-2. Full-text screening for the health sector (all 578 studies screened)

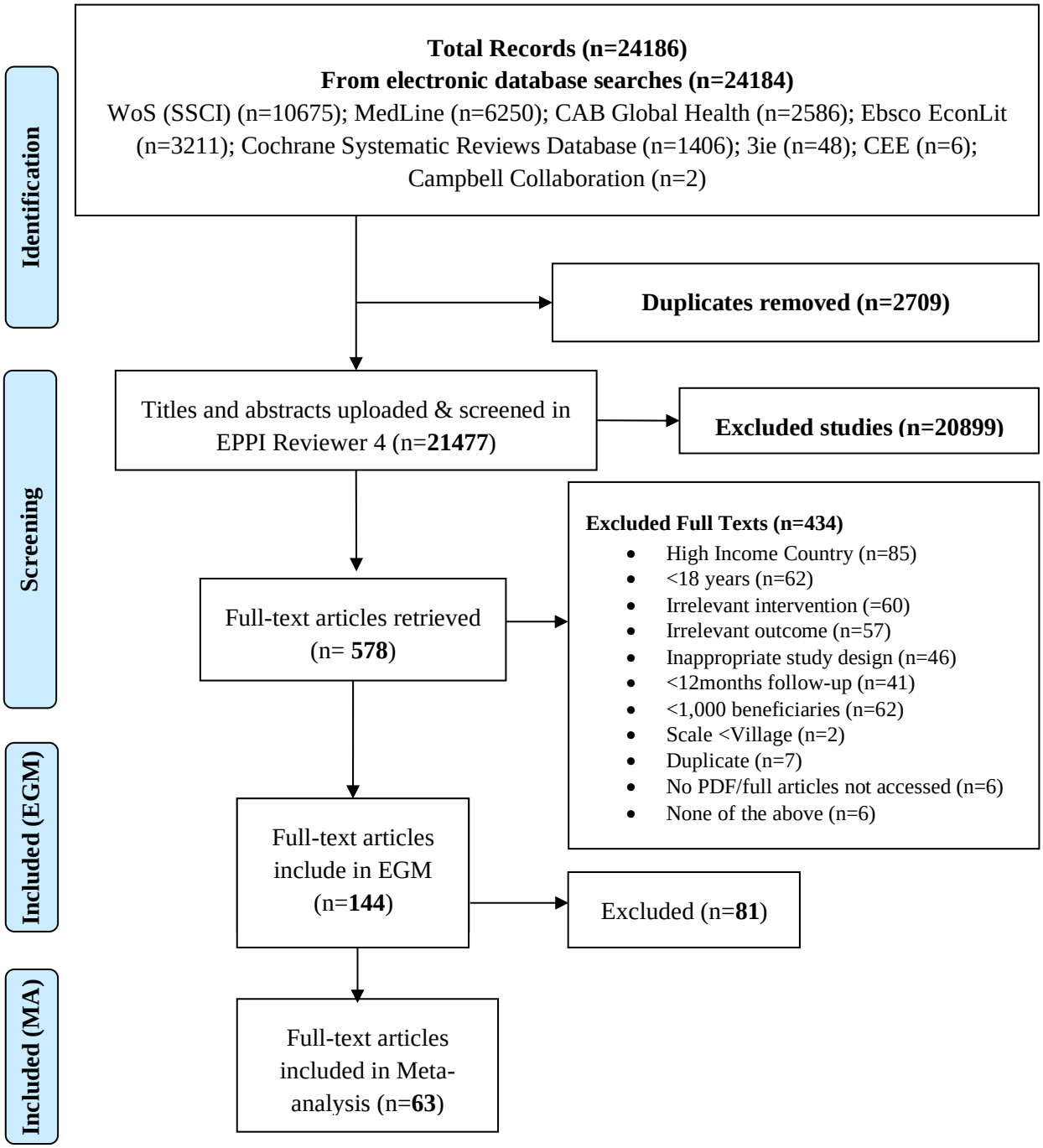
CODE IN EPPIR4	COUNT (FULL TEXTS SCREENED AND RECONCILED)
INCLUDE on full study	**144**
No PDF	6
Ex-HIC	85
ExAGE<18 years	62
ExINTERV—Irrelevant/notPHsector	60
ExOUTCOM—Irrelevant	57
ExSTUDYDESIGN—Inappropriate	46
ExSTUDYDURATION<12months	41
ExINTSCALE<1,000 beneficiaries	62
ExINTSCALE<Village	2
ExDUPLICATE	7
None of the codes above	6
Total	**578**

Source: Authors

Figure IV-1 provides the final numbers after full-text screening and reconciliation at both levels, where a total of 144 papers were included within the health EGM.

Of the total 144 selected studies, 120 were impact evaluations, while the remaining 24 were meta-analyses or systematic reviews. The list of final included studies can be found in References. The set of all 21,477 screened studies is provided in online appendix G.

Figure IV-1. Flow diagram for the health sector's evidence and gap map

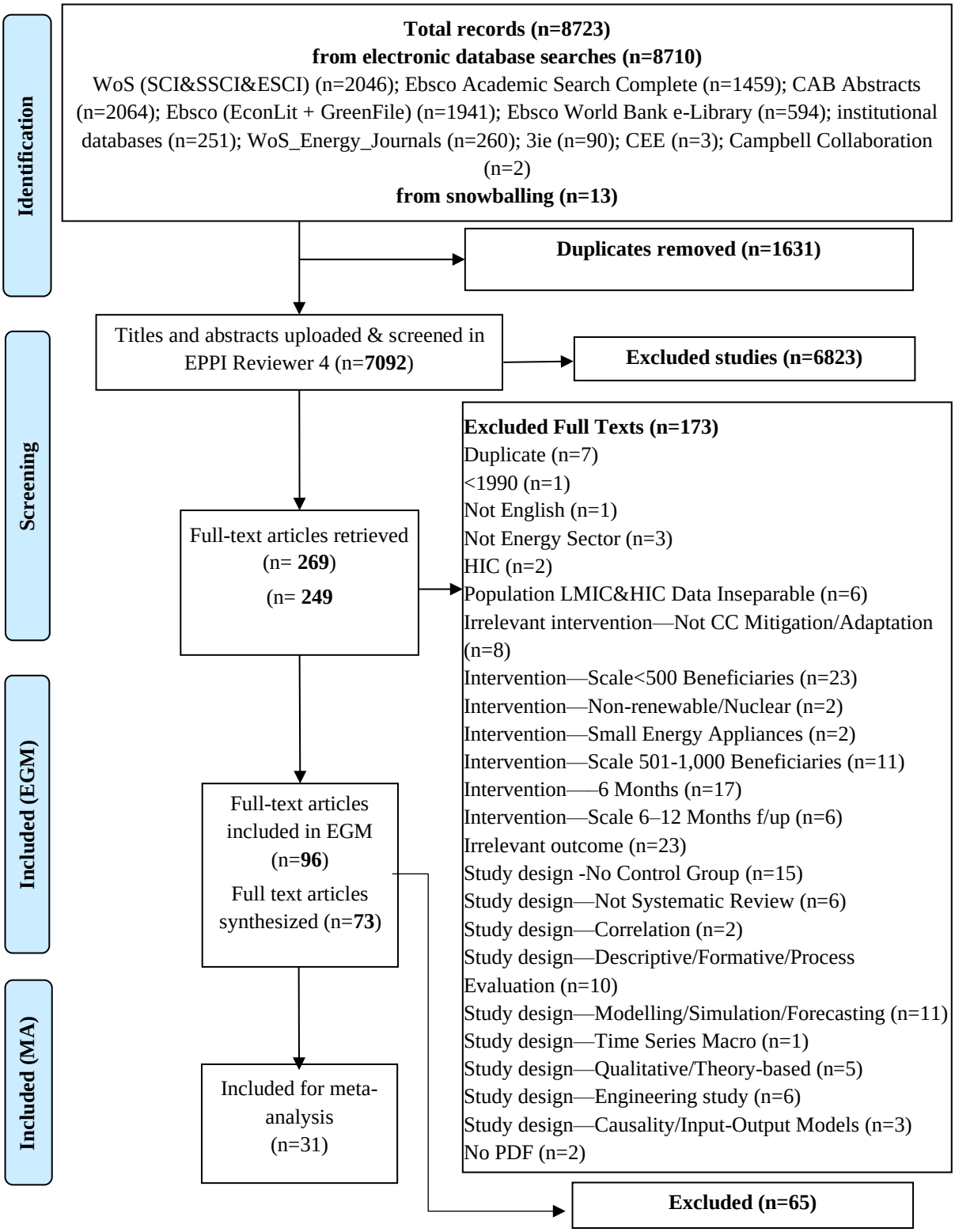


Source: Authors

3. ENERGY

In energy, a total of 8,710 studies were found in all databases combined, of which 96 finally entered the EGM. The flow of studies is shown in the PRISMA diagram (Figure IV-2).

Figure IV-2. Flow diagram for the energy sector evidence and gap map



Source: Authors

Of the 7,092 studies screened, 269 were included after title and abstract screening. Out of these 269 studies, 96 were included after the full-text screening and passed on to data extraction for inclusion in the EGM. The two tables below break down the reasons for exclusion at each stage. Note that many studies failed to meet multiple inclusion criteria; however, during screening, only one of these

criteria was marked for the exclusion of the study. The full list of included studies is provided in References. The set of all 7,092 screened studies is provided in online appendix H.

Table IV-3. Title and abstract screening in the energy sector

CODE IN EPPIR4	COUNT (STUDIES SCREENED AT TITLE & ABSTRACT AND RECONCILED)
INCLUDE on title and abstract	269
ExDUPLICATE_TiAb	152
ExLINGUAGE - not English	33
ExDATEPub<1990	2
ExPOP - HIC or Child<12 years	351
ExINTERV - Irrelevant	3,568
ExOUTCOME - Irrelevant	2,040
ExSTUDYDESIGN or Protocol/Guideline	677
Total	**7,092**

Source: Authors

Table IV-4. Full-text screening in the energy sector

CODE IN EPPIR4	COUNT
INCLUDE on full study	96
ExDUPLICATE - full text	7
ExDATEPub < 1990	1
ExLANGUAGE - not English	1
ExIRRELEVANT - not energy sector	3
ExPOP - HIC	2
ExPOP - LMIC&HIC data inseparable	6
ExINTERV-Irrelevant - not CC mitigation/adaptation	8
ExINTERV - scale<500 beneficiaries	23
ExINTERV - non-renewable/nuclear	2
ExINTERV - small energy appliances	2
ExINTERV - scale 501–1,000 beneficiaries	11
ExINTERV - 1–6 months	17
ExINTERV - scale 6–12 months f/up	6
ExOUTCOM - irrelevant	23
ExSTUDYDESIGN - no control group	15
ExSTUDYDESIGN - not systematic review	6
ExSTUDYDESIGN - correlation	2
ExSTUDYDESIGN - descriptive/formative/process evaluation	10
ExSTUDYDESIGN - modelling/simulation/forecasting	11
ExSTUDYDESIGN - time series macro	1

CODE IN EPPIR4	COUNT
ExSTUDYDESIGN - qualitative/theory based	5
ExSTUDYDESIGNEngineeringStudy	6
ExSTUDYDESIGNCausality/Input-OutputModels	3
No PDF	2
None of the codes above	0
Total	**269**

Source: Authors

B. EVIDENCE AND GAP MAPS

In this section, we briefly describe the evidence from the EGMs generated for both sectors. We describe the intervention and outcome disaggregation in both sectors, and then present a few graphs on the intervention types, the outcome types, and the regions where these studies were carried out in.

Interactive EGMs, with the links to included studies, can be accessed at the links below:

For the public health sector: https://ieu.greenclimate.fund/sites/default/files/page/public-health-egm-july2021.html.

For the energy sector: https://ieu.greenclimate.fund/sites/default/files/page/energy-egm-july2021.html.

The intervention and outcome categories that both define the scope of each sector and were used to code the studies are described in the inclusion and exclusion tables (see Appendix 3 and Appendix 4). Additional graphs, with descriptive statistics for the EGMs, can be found in online appendix I.

1. PUBLIC HEALTH

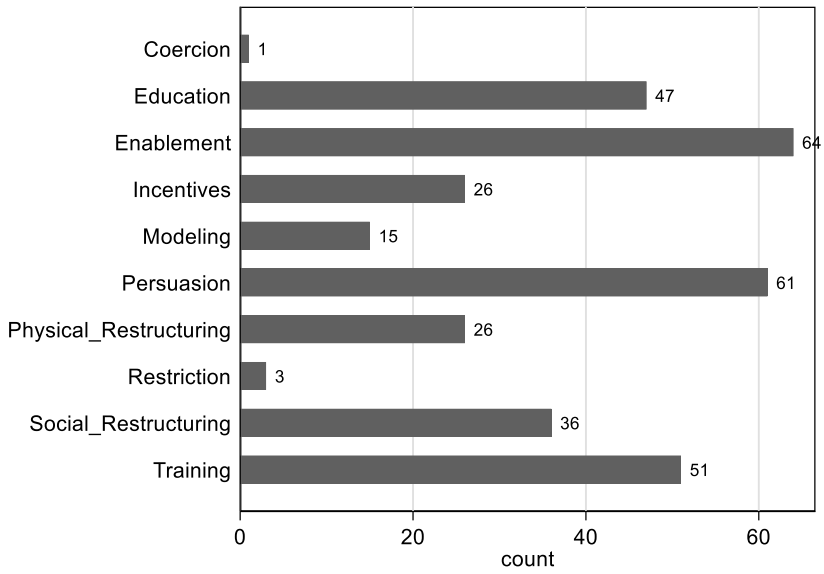
In the aggregate map below, we present the characteristics of the 120 studies (of the 144, 24 were systematic reviews) included in the EGM. The row totals report the number of times the intervention function was reported in the total sample, while the column totals indicate the number of studies reporting the particular abstract outcomes behaviour. Since the intervention functions were often combined with other intervention functions, and several papers report on many (and even all four) abstract outcomes, the row and column totals are overlapping and do not sum up to the total of 120. Each cell in the aggregate map depicts the study level (where the coding was done) overlaps between each intervention function and the particular abstract outcome.

Table IV-5. Map of studies in the public health sector, by abstract outcomes and intervention functions

INTERVENTION FUNCTION*	HEALTH-SEEKING BEHAVIOUR-PRIVATE BENEFITS	HEALTH-SEEKING BEHAVIOUR-SOCIAL BENEFITS	CONSUMPTION / PURCHASING DECISIONS-PRIVATE BENEFITS	CONSUMPTION / PURCHASING DECISIONS-SOCIAL BENEFITS	TOTAL
Coercion	0	0	1	0	1
Education	23	11	15	15	47
Enablement	27	10	25	17	64
Incentives	13	9	8	3	26
Modelling	4	6	4	5	15
Persuasion	31	12	12	20	61
Physical restructuring	11	1	4	11	26
Restriction	1	1	0	1	3
Social restructuring	21	5	3	10	36
Training	26	7	15	14	51
Total	**57**	**24**	**34**	**30**	**120**

Source: Authors
Note: *Each of the intervention function terms is defined and described in section I.D.1. Kindly refer to Appendix 9 which contains the definitions of all terms used in the behaviour change wheel.

Figure IV-3. Number of studies, by intervention functions, in the health sector



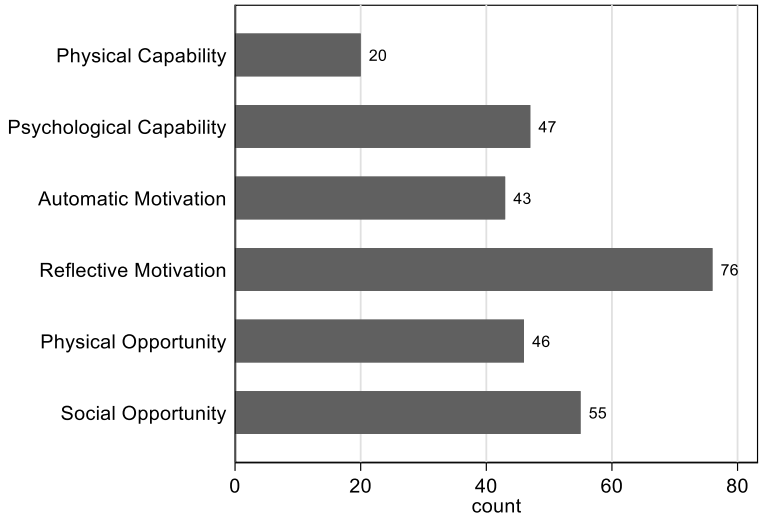
Source: Authors

As can be seen in Figure IV-3, in the public health sector, the most prevalent **intervention function** used to change behaviours within the health sector is enablement, with 64 out of the total 120 studies using this function in some form. Enablement is a relatively broad category that includes any kind of behavioural support, such as medication and surgeries. At 61 studies, persuasion ranked the second highest. This is not hard to imagine, given that all visual or oral communication aimed at

persuading participants is part of this intervention method. The other prominent intervention functions are training and education (for example, information campaigns).

Given the distribution of studies per intervention function, the results are also reflected within the COM-B model **sources of behaviour**, as shown in Figure IV-4 below (the definitions of these sources of behaviour are described in section I.D.1). Therefore, reflective motivation, that is, increasing knowledge or understanding, or eliciting positive (or negative) feelings about the behavioural target, is the most common source of behavioural change coded. Imagery and information campaigns, as well as training, are all interventions that target reflective motivation.

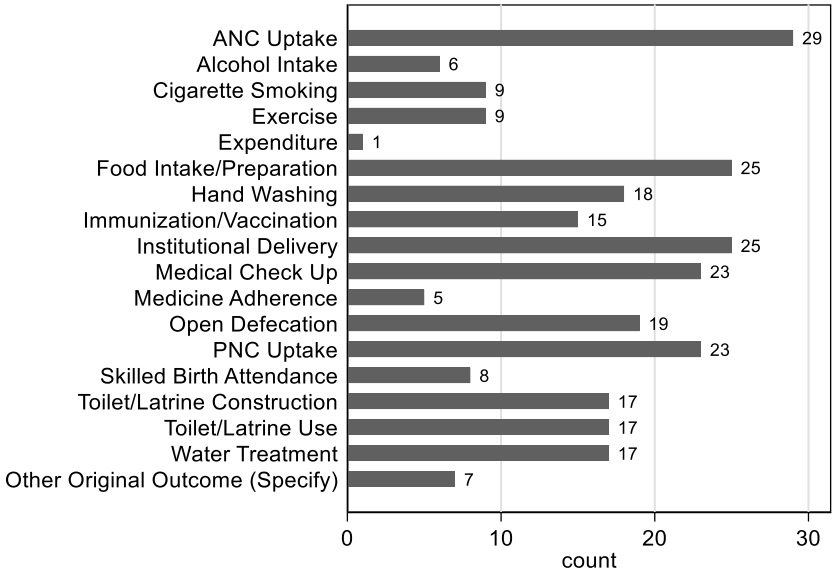
Figure IV-4. Number of studies, by source of behaviour, in the health sector



Source: Authors

In Figure IV-5 below, we can observe that the most commonly measured (behavioural) outcome in the public health sector is antenatal care (ANC) uptake. The succeeding rank outcomes are institutional delivery, postnatal care (PNC) uptake, and food intake/preparation.

Figure IV-5. Number of studies, by original outcomes, in the health sector

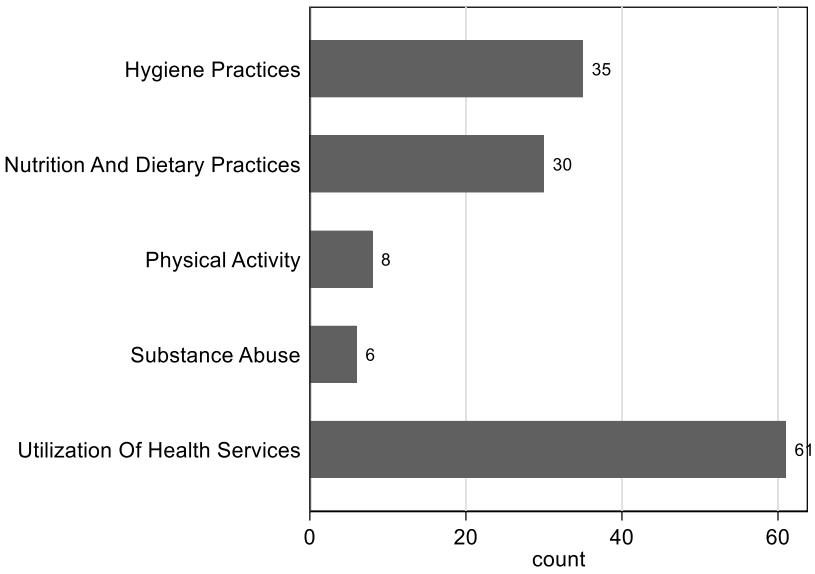


Source: Authors

The distribution of original outcomes within the study sample is correspondingly reflected within the **abstract outcomes** coded within the EGM (please see section I.D.1 that describes how the outcome categories are clustered and the definitions of abstract outcomes). Here, health-seeking behaviour, with largely private benefits, is the category that has the highest number of studies, given that improvements in ANC/PNC uptake, institutional delivery, along with food intake and preparation, are action behaviours with largely private benefits. Consumption and purchasing decisions that have largely private benefits fall under the second-highest category, as reflected from all the hygiene practice outcomes, such as toilet construction or water treatment.

Finally, the sets of studies **per sector** show that the highest number of outcomes is measured from the utilization of the health services subsector (Figure IV-6). There are also a considerable number of studies found within the cells of hygiene practices, as well as nutritional and dietary practices. From the graph, it appears that the subsectors of physical activity and substance abuse do not fulfil the criterion of a minimum of 10 studies and are, therefore, not suitable for a rigorous meta-evidence analysis.

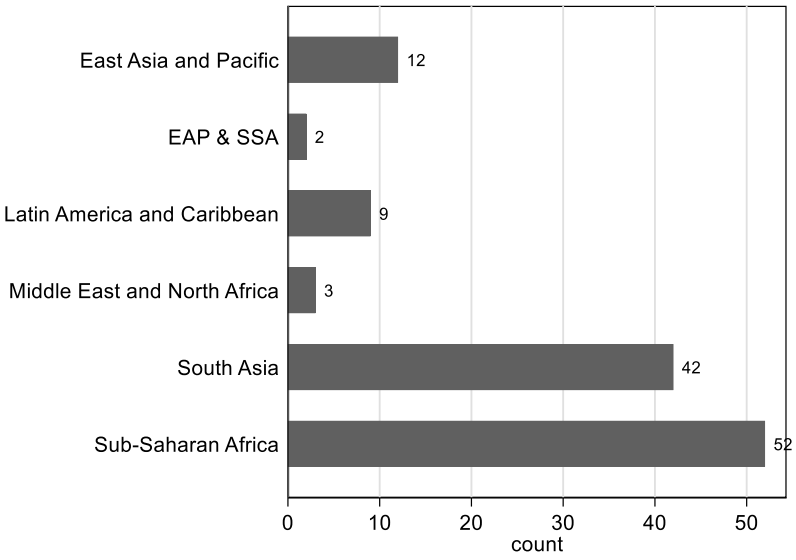
Figure IV-6. Number of studies by subsector, in the health sector



Source: Authors

From the five regions coded, the largest number of studies are from Sub-Saharan Africa (52) and followed by South Asia (42 studies), as shown in Figure IV-7 below.

Figure IV-7. Number of studies, by region, in the health sector



Source: Authors
Note: EAP (East Asia and Pacific), SSA (Sub-Saharan Africa)

2. ENERGY

Within the EGM, the unit of analysis is an individual study, whereby every entry represents a combination of an intervention and an outcome. A single study can have several outcomes or interventions, and therefore, may be entered into several categories simultaneously. The interlinkages between the different categories can be seen in the interactive EGMs online.[17] It is important to note that the following graphs depict the results only for the impact evaluation studies and not the systematic review or the meta-analysis studies, as is the case for the public health sector.

Table IV-6. Map of studies in the energy sector, by outcomes and interventions

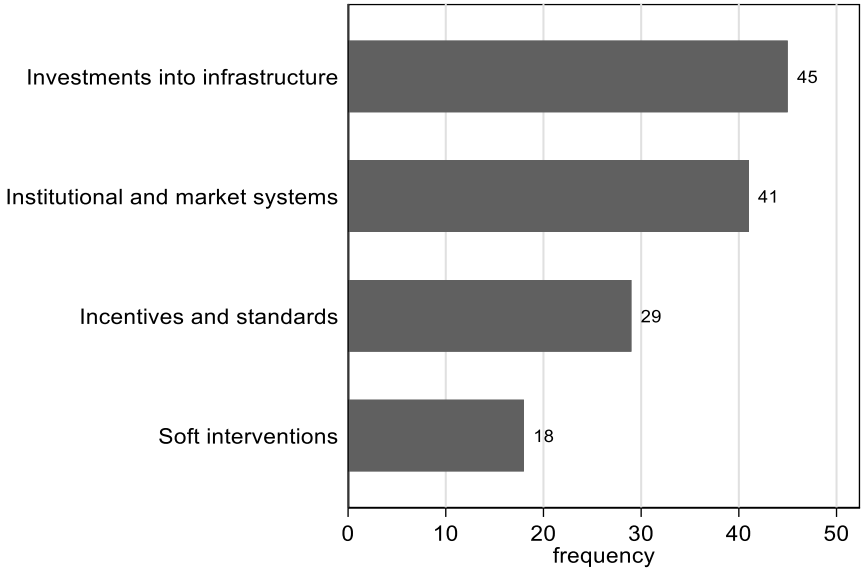
OUTCOMES / INTERVENTIONS	ACCESS TO ENERGY AND SUPPLY OF ENERGY	ENERGY MARKET DEVELOP-MENT	ENERGY CONSUMPTION AND DEMAND	ADOPTION OF MORE ENERGY- EFFICIENT TECHNOLOGIES	RESILIENCE OF ENERGY SYSTEMS (ADAPTATION)	GHG EMISSIONS AND POLLUTION	LABOUR MARKET CO-BENEFITS	TOTAL
Institutional and market systems	11	8	13	7	4	26	3	41
Incentives and standards	10	7	20	13	3	12	4	29
Soft interventions	7	1	14	9	6	8	7	18
Investments into infrastructure, equipment, and technologies	24	5	30	13	7	11	28	45
Total	**35**	**13**	**51**	**25**	**11**	**42**	**30**	**96**

Source: Authors

[17] The links to the EGM reports are provided in Appendix 10.

As can be seen, the studies in the energy sector are relatively evenly split across most intervention categories, as shown in Figure IV-8.

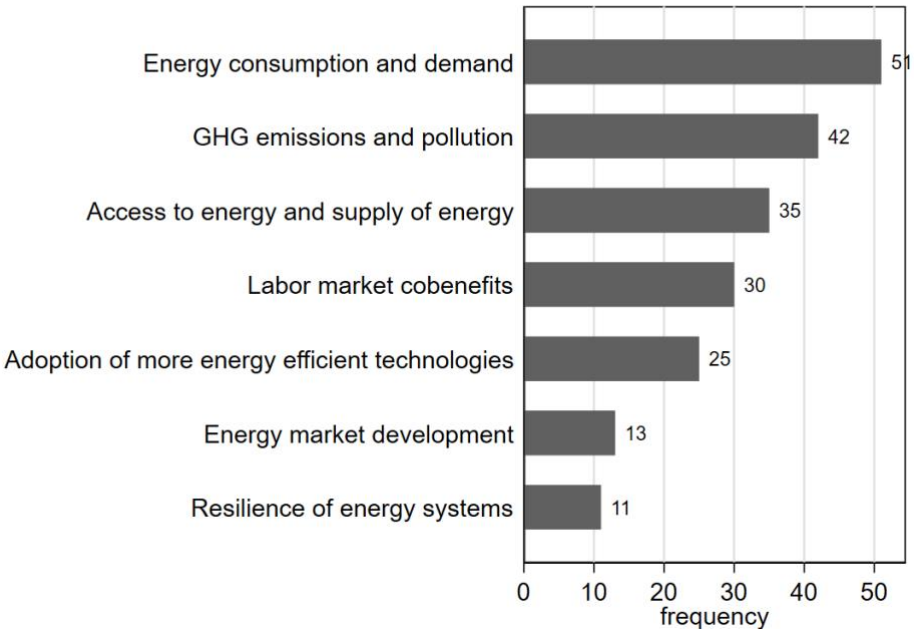
Figure IV-8. Number of studies, by intervention category, in the energy sector



Source: Authors

In terms of outcomes, the largest category is “energy consumption and demand”, followed by “GHG emissions and pollution”. By contrast, only a few studies covered outcomes in the categories of “energy market development” and the “resilience of energy systems (adaptation)”. This clearly shows that the large majority of studies fall into the area of climate change mitigation, with only a few covering resilience or the adaptation of energy systems.

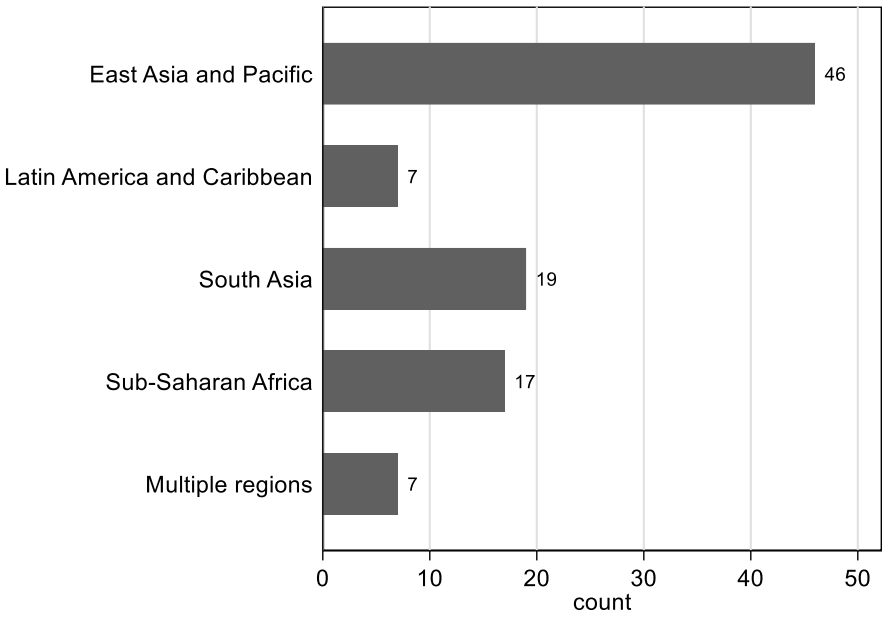
Figure IV-9. Number of studies, by outcome category, in the energy sector



Source: Authors

In the energy sector, most interventions took place in the region of "East Asia and Pacific", which is largely due to China being, by some distance, the country with the highest frequency of included studies.

Figure IV-10. Number of studies, per region, in the energy sector



Source: Authors

C. DESCRIPTION OF STUDIES FOR THE SYSTEMATIC REVIEWS

This section describes the studies included in the systematic reviews. They were quantitatively synthesized through meta-analysis within each sector. As described in subsection III.C.2, it is important that the studies are both conceptually and statistically similar so that their comparisons and combinations were possible.

1. PUBLIC HEALTH

For the public health sector, the large number of studies implies considerable heterogeneity in study design, outcomes, and measurement (methods). This was particularly obvious when tabulating the different types of treatment arms coded for the 120 papers, whereby the most frequent treatment arm (enablement plus incentives) was reported in only 13 studies, followed by 10 studies for enablement alone. In order to facilitate the comparison between studies, we decided to proceed with a subsample of studies with the largest intersection in intervention functions in their treatment arms. Therefore, all the included studies have to be part of a sample of a minimum of 10 studies reporting the same combination of three intervention functions.[18] At the end of this process, we selected a total of 63

[18] We first coded all 720, that is, 10 x 9 x 8 combinations of intervention functions (excluding combinations of the same intervention functions, such as education x education x education & education x education, we got 280 less combinations). After summing up the frequency of each combination, those combinations that had at least 10 studies were kept, while all studies that were part of the combinations with less than 10 studies were dropped. In case there were more than 10 studies with only one intervention-function type (enablement, for instance), or two intervention-function types (enablement and incentives, for example), these studies were automatically included.

studies with relatively similar intervention function combinations in their treatments.[19] Through this method, we were certain that we were comparing similar types of study designs.

Due to the low frequency of evidence from two subsectoral areas of physical activity and in particular, substance abuse, they were excluded. Therefore, the meta-analysis focuses only on the three subsectoral areas—the utilization of health services, hygiene practices, along with nutrition and dietary habits—to gather evidence on transformational change in behaviour. Consequently, the evidence we examined was further organized, based on a lower variation in outcome types across the subsectors. The meta-analysis by subsector was restricted by the low sample sizes, when considering the additional intersection between the intervention function and abstract outcome types. Nonetheless, we attempted to explore the subsector of the utilization of health services in more detail, in order to provide a glimpse into the subsectoral effect sizes in section E.1.g.

The goal of the meta-analysis is to examine the role of the various intervention types on the four abstract outcomes. Ultimately, our hope is to determine which of the intervention functions is most likely to lead to a transformational (or approaching large) effect on the abstract outcome. For this analysis, we, therefore, focused on those intervention functions that are most frequent and likely to yield an adequate amount of effect sizes per cell.

Of the total 63 studies selected based on the commonality in the intervention function, 53 studies were included in the meta-analysis. This reduction stems from the difficulties we faced regarding the comparison of the study results. One of the main issues is that some studies did not focus on measuring behaviour but included it as an intermediate result. This frequently led to a lack of statistical information that necessitated the exclusion of the paper. A prime example for this is Memon and others (2015), whereby the aim is to reduce perinatal and neonatal death through a community-based intervention that also targets the pregnant women's behaviour. Unfortunately, the lack of a measure of precision for the effect size led to this study being dropped from our sample. Another consequence of the same fundamental issue is that the behavioural outcomes are not directly reported in the first place. To illustrate, Triyana and Shankar (2017) investigate the effectiveness of a conditional cash transfer programme on antenatal coverage and quality. Although the intervention, in principle, is aimed at increasing the use of antenatal care, this is not directly measured. Instead, the paper reports statistics on which services the women received. In general, the greater amount of care received coincides with more women accessing antenatal care, but for the purpose of this synthesis, these outcomes do not directly measure behaviour.

Of the 412 relevant effect sizes extracted, the analysis was restricted to a total of 299 effect sizes. From these effect sizes, those that are similar within each treatment arm (such as handwashing before handling food and handwashing after coming home) were averaged to create one single effect size per treatment arm. The synthesis of the outcomes within one treatment arm led to the number of effect sizes being reduced from 299 to 136 distinct effect sizes (within the treatment arm). Table IV-7 describes the main characteristics of the final 136 effect sizes used in the analysis below.

Among the included effect sizes, health-seeking behaviour with largely private benefits and health-seeking behaviour with social externalities are the most frequently observed abstract outcome types. To evaluate the effects of the interventions, regressions (coefficients) were used the most frequently, followed by the proportion of participants who experienced change, as well as odds ratios and risk differences with standard errors.

[19] In actuality, there were 56 studies, to which seven benchmark papers, which also went through the intervention function (combination) similarity check, were later added.

Table IV-7. Descriptive statistics for the effect sizes

VARIABLE	OBS.	FREQUENCY	MEAN
Total number of papers	52		
Type of econometric method used			
Number of participants who experienced change	136	7	0.05
Proportion of participants who experienced change	136	15	0.11
Regression	136	58	0.43
Relative Risk (Risk Ratio) + SE	136	7	0.05
Risk Difference + SE	136	16	0.12
Abstract outcomes			
Health-seeking behaviour with largely private benefits	136	55	0.40
Health-seeking behaviour with social externalities	136	31	0.23
Consumption or purchasing decision with largely private benefits	136	23	0.17
Consumption or purchasing decision with social externalities	136	27	0.20
Intervention functions			
Education	136	57	0.42
Enablement	136	66	0.49
Incentives	136	38	0.28
Modelling	136	18	0.13
Persuasion	136	57	0.42
Physical restructuring	136	13	0.10
Social restructuring	136	41	0.30
Training	136	40	0.29

Source: Authors

Table IV-8. Map of effect sizes in the public health sector, by abstract outcomes and intervention function

INTERVENTION FUNCTION	HEALTH-SEEKING BEHAVIOUR-PRIVATE BENEFITS	HEALTH-SEEKING BEHAVIOUR-SOCIAL BENEFITS	CONSUMPTION / PURCHASING DECISIONS-PRIVATE BENEFITS	CONSUMPTION / PURCHASING DECISIONS- SOCIAL BENEFITS	TOTAL
Education	26	12	9	10	57
Enablement	26	17	9	14	66
Incentives	14	8	11	5	38
Modelling	8	5	3	2	18
Persuasion	24	16	4	13	57
Physical restructuring	6	3	0	4	13
Social restructuring	21	10	4	6	41
Training	19	11	4	6	40
Total	**55**	**31**	**23**	**27**	**136**

Source: Authors

The aggregate map on the effects level (per treatment arm) is provided in Table IV-8. As shown in the aggregate map for health at the effect size level, there are multiple combinations (cells) with a large enough number of outcomes. We see a high concentration of effects in the first two columns, meaning that we have a larger number of outcomes measuring health-seeking behaviours, especially those with largely private benefits. Since intervention functions, as defined earlier in accordance with the behaviour wheel, are not mutually exclusive, we can see quite a large number of outcomes, with enablement, persuasion, or education as part of their interventions. These are exactly the cells that allow for a meta-analysis without sample size restrictions. Due to the paper selection based on the frequent combinations of the intervention functions as described above, none of the included papers has coercion or restriction as part of their intervention. Therefore, these two intervention functions were dropped.

2. ENERGY

On a conceptual level, two cells in the energy EGM stand out as candidates for the meta-analysis. The first is the intersection of the intervention subcategory - "investments into energy transmission, distribution and storage of electric energy systems" - with the outcome subcategory, "employment in the formal sector". In this cell, there are 22 studies on the effects of grid investments on formal employment. Among them, almost all studies address the benefits of rural electrification, and are therefore, closely linked. The second cell is the intersection of the intervention subcategory - "privatization, liberalization, and the introduction of market-based mechanisms" - with the outcome subcategory, "CO_2 emissions". Here, we find many studies (22) on the effects of China's pilot carbon trading scheme on GHG emissions.

In addition to our initial searches, we conducted targeted snowball searches for these two cells. For the first cell, we found one additional paper, and for the second cell, we found 12. These numbers do not reflect the limitations of our original search, as all but one of the papers found were published after we conducted our main searches.

In the first cell, the papers all aim to measure the employment effects of electrification. The degree of electrification in the world has continued to rise over time—crossing the threshold of 90 per cent for the first time in 2019 according to *World Bank Data* (World Bank, 2021a). The studies in this cell provide evidence on the labour market co-benefits of these interventions. Most of the remaining population who still lack access to electricity live in Sub-Saharan Africa, due to the significant increases in electrification in South Asia over the last decades (from around 45.5 per cent in 1994 to 94.4 per cent in 2019 (World Bank, 2021b). Despite the increase in the electrification rate in Sub-Saharan Africa from 26 per cent in 2000 to 46.8 per cent in 2019, this region is still where most future electrification will need to take place.

Due to the nature of the intervention, many of the papers focus on rural settings. While the labour markets in the locations the papers investigated could vary due to the geographical breadth of the papers included, they share similarities that follow from both setting and situation—rural and without access to electricity. The papers in this cell differ regarding how both the electrification and labour market effects are measured. Electrification measurements tend towards measuring either the village or municipal electrification rate or whether a household is connected to the grid (either directly, or indirectly, based on whether it uses the electric light). For measurements of labour market effects, two approaches were identified: measuring employment status and time spent in work activities. Of the 22 papers in the cell, 15 papers measure employment status, nine measure time spent in work activities on various scales, and two measure both. Due to the differences in interpretation, we did not combine these groups in the meta-analysis. Additionally, due to a lack of comparability and the difficulty of calculating a comparable estimate, we excluded studies that do

not present linear probability models unless they report their outcomes as marginal effects at the mean (of all variables measured). This meant excluding one study completely that only uses the time measure. For the reasons discussed above in sections III.C.2 and III.C.4, we decided to only further investigate employment status in our subsequent analysis.

Of the 15 remaining studies selected for the meta-analysis, five were conducted in Sub-Saharan Africa, five in South Asia, along with five in Latin America and the Caribbean. While none of the papers looked exclusively at urban settings, five papers investigated effects for both rural and urban settings.

The second cell collected papers that investigate market-based mechanisms to reduce CO_2 emissions, in particular, the emissions trading system (ETS). All papers in this cell attempt to measure the emission reductions from the introduction of the Chinese pilot ETS. Between August 2013 and February 2014, China implemented an ETS in Beijing, Chongqing, Guangdong, Hubei, Shanghai, Shenzhen, and Tianjin. In February 2021, China started expanding the ETS to cover the entire country (although still only in selected industries), which suggests that the government considered the pilot to be a success. Once fully rolled out, the Chinese ETS will be the largest ETS in the world by far, covering one-seventh of the global CO_2 emissions from fossil fuels (International Energy Agency, 2020a). How effective the Chinese ETS reduces CO_2 emissions is, therefore, one of the most important questions for climate change mitigation.

Because of the initial selective implementation in only some regions, instead of the whole country, the situation could serve as a quasi-experiment, allowing for papers investigating the effects to fall within our inclusion criteria. While all the papers investigate data based on the same events, their methods and choices regarding data and time periods differ. Although all papers use the DiD analysis, they vary regarding the unit of analysis—from whole regions to individual firms. The largest difference between the papers is the choice of what measure of CO_2 emission to use: while some papers report carbon intensity or carbon productivity, most papers rely on either tons of CO_2 emissions or their logarithms. To ensure the comparability of the measures as well, based on the number of available studies, we decided to exclude studies that only measure carbon intensity or carbon productivity.

Of the initial 10 studies from our original search, only five report the logarithm of carbon emissions, but after adding the papers found through snowballing, that number increased to 12. Only two of the original 10 studies measure GHG emissions in CO_2 equivalents, but even after adding the papers found through snowballing, this number only increased to seven. After deciding that these two measures should not be combined into one meta-analysis (given that absolute carbon emissions are likely to be highly dependent on the data sample, such as specific regions or industries), we decided to conduct the meta-analysis with only logarithmic CO_2 emission outcomes. As the Chinese ETS pilot was announced in 2011 and initiated in 2013–2014, long-term effects over multiple years would only be visible in recent data. This explains the large number of very recent studies that had not been published, and therefore, could not be retrieved at the time of our initial database search, but were later found through snowballing.

D. Risk-of-bias in included studies in the systematic review

1. Public health

Out of the total of 63 studies included in this systematic review, almost half obtained a high-quality rating (49 per cent), while at the same time, a significant number of studies (33 per cent) were rated as low-quality.

Table IV-9. Summary results of the risk-of-bias assessment

Overall confidence in study findings	Count	Percentage (%)
High	35	56
Medium	8	13
Low	20	32
Total	**63**	**100**

Source: Authors

2. Energy

Out of the total of 31 studies included in this systematic review, most obtained a medium-quality rating (42 per cent), while again, a significant number of studies (39 per cent) were rated as low-quality.

Table IV-10. Summary results of the risk-of-bias assessment

Overall confidence in study findings	Count	Percentage (%)
High	5	16
Medium	12	39
Low	14	45
Total	**31**	**100**

Source: Authors

The risk-of-bias data can be found in online appendices J and K.

E. Synthesis of results: meta-analysis

1. Public health

Given the breadth of the results that can be presented for the public health sector, we chose to focus on particular results, largely those driven by larger sample sizes (thereby implying larger confidence in those results) or larger effect sizes.

As per the aggregate map of the total 136 effect sizes presented in section C.1, we first discuss the effect sizes for a given intervention type. The intent of the presentation of the results is to assess which intervention function is likely to lead to transformational change within a particular abstract outcome category. However, it might still be relevant to determine if there are intervention functions that are likely to lead to transformational change, regardless of the abstract outcome type. Therefore, the first subsection of meta-results presents the intervention functions without disaggregating them by the four abstract outcomes, that is, we used the pool of all the abstract outcomes available to us. The following subsections then present the results by the four abstract outcomes.[20]

Unlike the EGM aggregate map on the **study level**, in section B.1, where consumption and purchasing behaviours (social or private) are much more frequent than health-seeking behaviour, the **treatment-arm-level outcome data** (Table IV-7 or Table IV-8) shows that the most populous

[20] In this section, the results for the intervention functions are not presented in the form of forest plots, given the low legibility of these figures. Instead, the results are presented in a tabular format, while the plots themselves can be found in Online Appendix L. We also include all the tables presented in this report in Online Appendix P.

outcomes are categorized within health-seeking behaviour (with private or social benefits). We present the results by decreasing sample sizes within outcome types, that is, after discussing the results from the intervention function samples on the pooled outcomes. The next subsection delves into the overall effects for the abstract outcomes of health-seeking behaviour with largely private benefits and the health-seeking behaviour with social benefits. Finally, we report on the consumption/purchasing decision outcomes, but only briefly, due to the small sample sizes, particularly when considering their intersection with the three prominent intervention functions of enablement, persuasion, and education.

It is important to note that each of the forest plots includes several effect sizes from the same study, but different treatment arms (as indicated by the first digit after the lead author's last name and year of publication), as well as various types of original outcomes (as indicated by the second digit in each effect identifier). Additionally, in order to show the possible publication biases that these cells might be facing (and thereby, qualify the interpretation of the results we present), we present the funnel plots for the results presented. All the funnel plots generated as part of the analysis, regardless of whether they were included within this report, can be found in online appendix M.

Finally, two other sets of results are presented in this report—one to account for the large variation in our included studies and the other to explore the practical magnitude of the effects, given the incongruousness of interpreting Cohen's d in practical terms. The first results discuss the meta-regressions performed on the row and column totals of the aggregate map. The final set of results shows two different samples, where the unstandardized coefficients from the odd's ratios and regressions, specifically for the utilization of the health services sector, are depicted.

a. Effects of intervention functions on all outcome types

From the theoretical causal chain, we are interested in the effects of the most commonly found intervention functions on the abstract outcomes. From the 136 effects included in the overall analysis, nearly half of the effects (66) relate to a treatment with an enablement aspect. Additionally, education (57) and persuasion (57) are also found to be the most popular intervention functions to influence the behaviour of individuals. In this section, we focus on the overall effect sizes of these interventions (across the different abstract outcomes) and present the results, disaggregated by abstract outcomes, in the following subsections.

Table IV-11 provides the overall effect size, the p-value, the I-squared and Tau-squared statistics, along with the number of observations, in the meta-analysis for all eight intervention functions in our study sample.

Table IV-11. Meta-analysis of the intervention function cells

Intervention function	Effect size (Cohen's d)	p-value	I sq.	Tau sq.	Tau Lower	Tau Upper	Obs.
Education	0.22	0.00	98.82	0.18	-0.60	1.05	57
Enablement	0.42	0.00	99.32	0.32	-0.70	1.54	66
Modelling	0.24	0.08	98.41	0.31	-0.86	1.34	18
Incentives	0.11	0.00	91.64	0.02	-0.15	0.36	38
Persuasion	0.24	0.00	98.79	0.19	-0.60	1.09	57
Training	0.24	0.00	97.66	0.11	-0.40	0.88	40
Physical restructuring	0.13	0.12	98.01	0.08	-0.42	0.69	13
Social restructuring	0.25	0.00	98.24	0.13	-0.45	0.96	41

Source: Authors

As can be seen in Table IV-11, besides physical restructuring, all other intervention functions show an effect size that is statistically significantly different from the null effect (modelling is significant only at the 10 per cent level). In terms of effect sizes, the largest effect is observed for the enablement category, where the treatment group mean lies above the control group mean by 0.42 standard deviations (SD). This effect is above our categorization of a small (relative) effect size (above 0.2 as per our effect-size definition given in section II) and nearly approaching a moderate transformational change effect of 0.5. Therefore, treatment arms that incorporate enablement as one of the intervention functions in their design affect small-to-moderately-large changes in behavioural health outcomes (that is, increased handwashing, reduced open defecation, rise in institutional delivery/decline in unassisted deliveries). The other intervention functions that report a small effect size on behaviour of between 0.22 and 0.25 are education, modelling, persuasion, training, social restructuring. Incentives and physical restructuring are the only interventions where the effect size lies below our defined threshold of a small effect size, at 0.11 and 0.13, respectively.

Given such a large diversity of designs and methodologies, substantial heterogeneity is to be expected. First, relative heterogeneity is high in all intervention function samples, as indicated by the high value in the I-squared (above 75 per cent), even approaching 100 per cent. This large value suggests that nearly all the heterogeneity in the results stems from between-study variability in true effects, rather than sampling error (or chance). The I-squared, in our case, is likely affected by the various types of effect sizes that were converted to Cohen's d and the large sample sizes that are part of the samples in our study.

To enable a more meaningful conclusion on heterogeneity, we focused on the absolute measure of heterogeneity—the variance of the true effect sizes—that is estimated by Tau-squared. Its standard deviation, Tau, enables us to predict the range that the true effect sizes will fall within. For the education, enablement, and persuasion samples, the true effect sizes fall, with 95 per cent certainty, within the ranges of -0.60 and 1.05, -0.70 and 1.54, along with -0.60 and 1.09, respectively. For instance, including the enablement intervention function within a treatment may lead to a medium negative effect of 0.7 SD, or a very large effect of 1.54 SD, implying a large absolute value of heterogeneity. Similarly, education and persuasion also suggest a large standard deviation of underlying true effect sizes.

In order to investigate heterogeneity, and in particular "small-study effects", such as due to publication bias, we used two types of tests—the regression-based Egger test and funnel plots.

The funnel plots for the three intervention functions - education, enablement, and persuasion (depicted in Figure IV-11 to Table IV-13, respectively)- plot the effect estimates against their

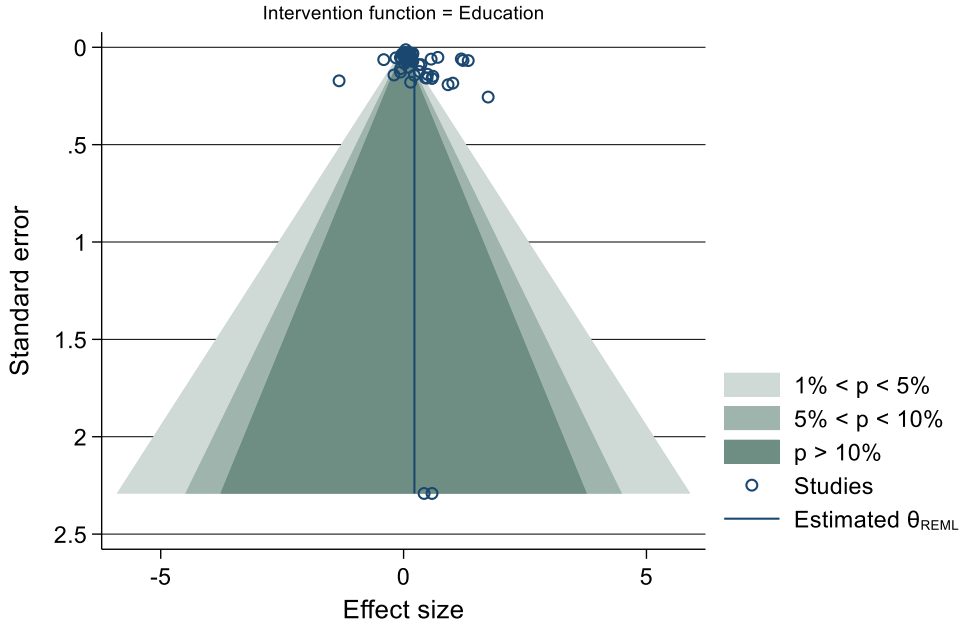
measure of precision (in our case, standard errors). The larger studies/effect sizes are placed at the top of these plots (because their estimates are more precise), while smaller sample studies are scattered in the lower part of the funnel. Each figure below indicates the significance boundaries between the one per cent and 10 per cent level, given the standard error, with the estimated effect size depicted by the vertical line. Non-biased reporting (no publication bias, no selective outcome reporting, sound methodological design) would imply that the estimates are symmetrically distributed around the overall effect-size line. More importantly, the effects would not be placed just inside the regions of significance and along their boundaries.

The plot for the enablement sample suggests evidence for some type of bias, where most estimates are bunched close to the significance boundaries (especially that of 1– 5 per cent). The effects can still be seen to lie outside the significance bounds in the direction favouring treatment (to the right of the effect size line). The samples for education and persuasion are not as clear, where the standard errors and effect sizes are not obviously displaying a downward correlation.

To test for publication bias statistically, we ran a regression-based Egger test (using the *meta bias* command in Stata) that regresses the effect size on its standard error to derive the correlation between the two (high correlation implying strong evidence for a small-study effect). The Egger test failed to confirm the evidence of a small-study effect for the education and persuasion samples. However, for three out of the eight intervention function samples, that is, enablement, incentives, training, and physical restructuring, the test does reveal a significant correlation between the magnitude of effect sizes and their standard errors ($p<0.011$), with the beta of the Egger test being the highest for the training sample (4.605). The tables in online appendix N provide the beta, standard errors, and p-value of the bias test of all intervention samples.

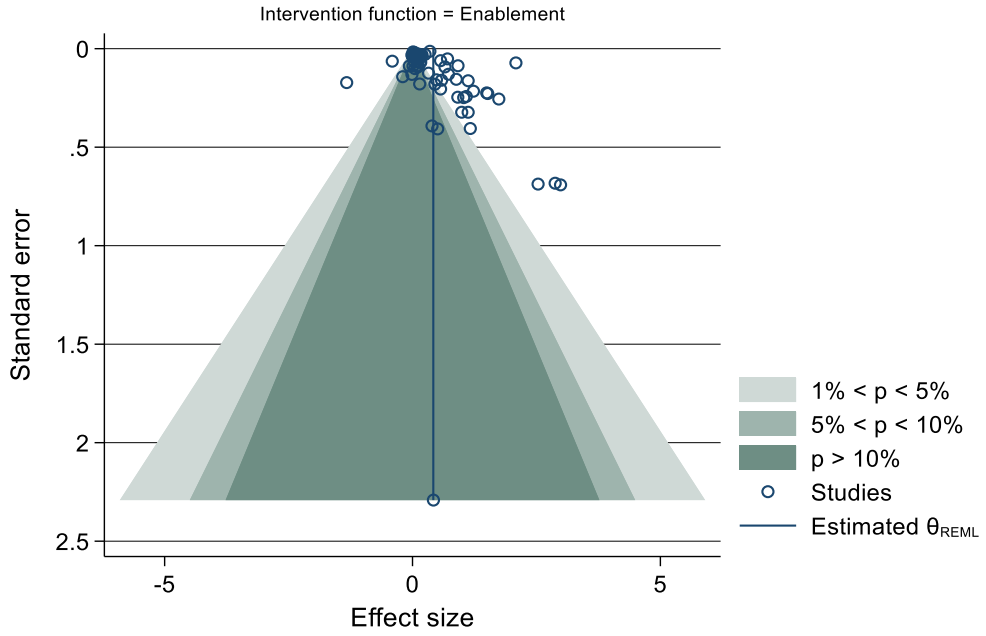
A limitation inherent in the inspections and tests of small-study effects is that they work under the assumption of homogeneity in effect sizes, which is clearly violated in our case. While we cannot think of a plausible reason for why smaller studies should have larger effects other than some form of publication bias or selective reporting, we cannot rule out other reasons linked to heterogeneity in effect sizes that could explain this correlation. As mentioned previously, the forest and funnel plots for the remaining intervention functions are provided in online appendices L and M.

Figure IV-11. Funnel plot for the Education intervention function sample



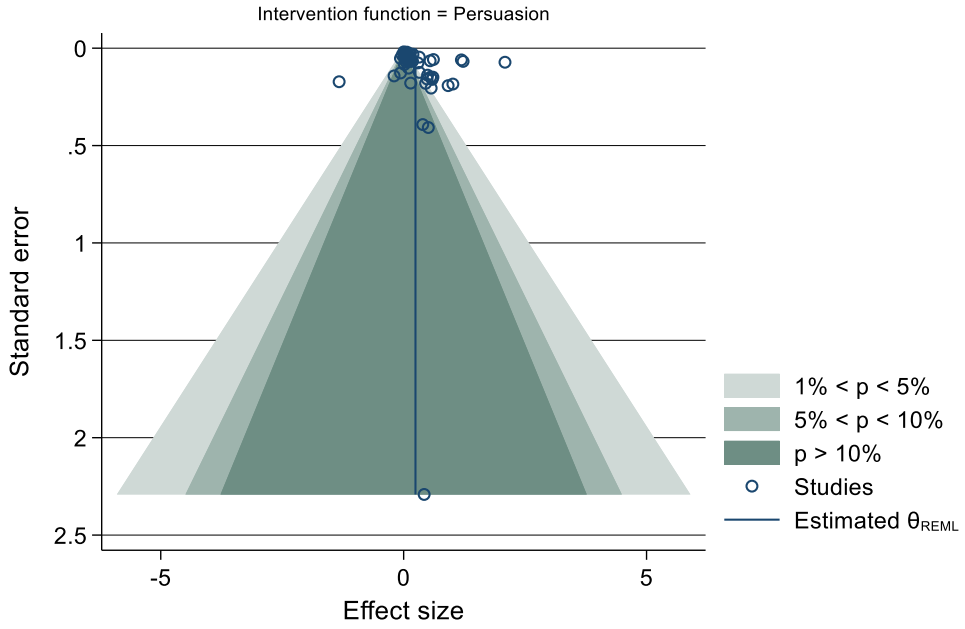
Source: Authors

Figure IV-12. Funnel plot for the Enablement intervention function sample



Source: Authors

Figure IV-13. Funnel plot for the Persuasion intervention function sample



Source: Authors

As evidenced by graphical representation and statistical tests of the relationship between the effect sizes and standard errors, there is a high likelihood of publication bias in studies from the enablement, training, and physical restructuring pool of studies, and to a lesser extent, the incentives sample.

b. Effects of intervention functions on health-seeking behaviour with largely private benefits

Given the large number of studies within the combined intervention function samples, this section focuses on the results for the subsamples related to the health-seeking behaviour with largely private benefits—the outcome with the largest overall and intervention function-specific sample types.

Table IV-12 tabulates the overall effect sizes of all eight intervention functions, where the samples for physical restructuring and modelling are quite low at six and eight observations. In the case of health-seeking behaviours with largely private benefits, all intervention functions report an effect size statistically significantly higher than the null effect, except in the case of physical restructuring. For the three largest samples of education, enablement, and persuasion, the largest overall effect size of 0.25 was found in the case of persuasion, followed by enablement (0.21), and education (0.19). The effect sizes are all considered small (or even lower than this threshold) as per our definition. The high I-squared value indicates large (relative) heterogeneity in effect sizes, that is, not driven by chance. For persuasion, the standard deviation of the true effect size, Tau, is quite large—95 per cent of true effect sizes is expected to range from -0.38 and 0.88, implying large absolute heterogeneity as well. Education and enablement report similarly high levels in Tau as well. Overall, all intervention functions (except for modelling) appear to have small or less-than-small effects on outcomes that fall under health-seeking with private benefits.

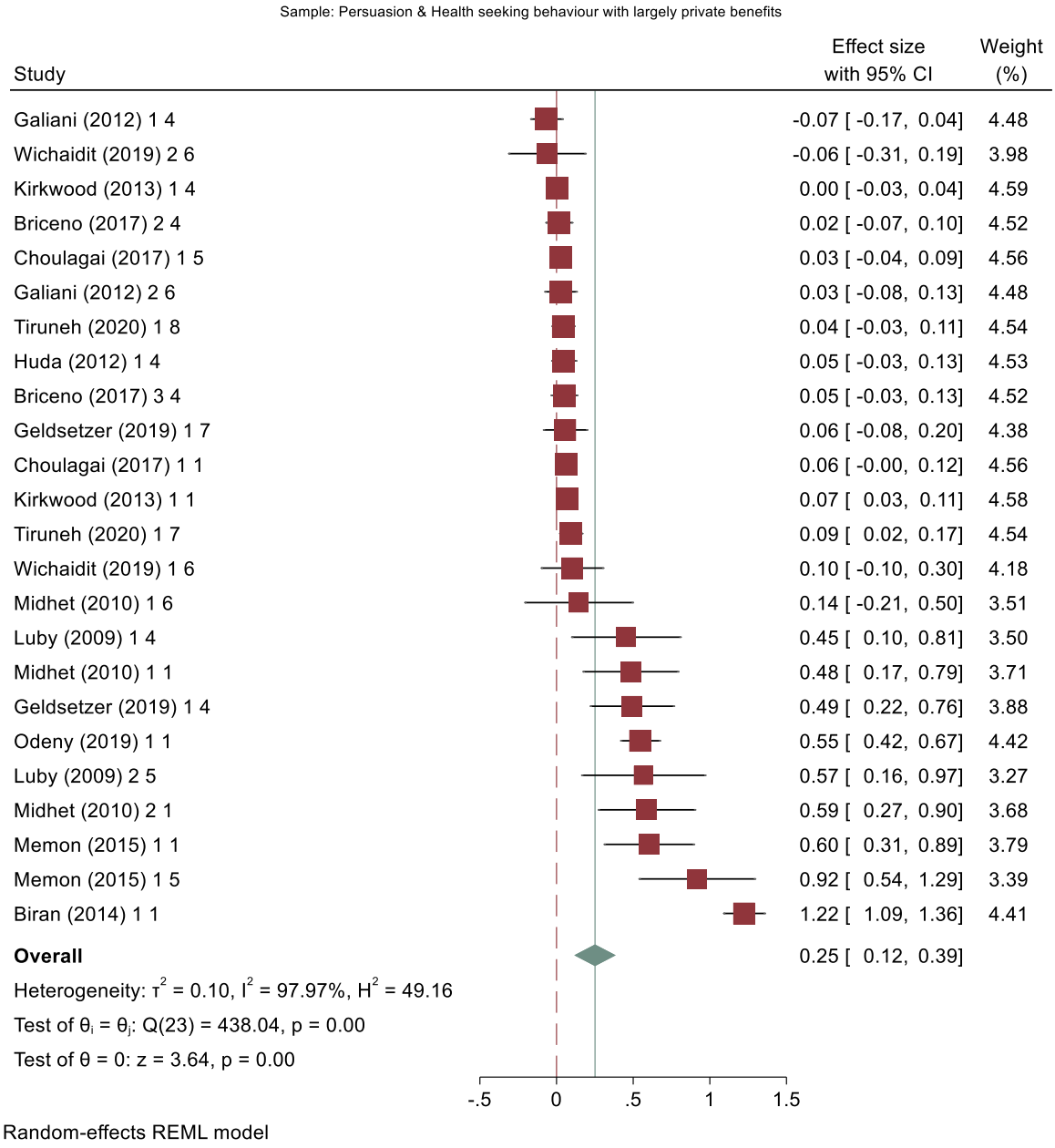
Table IV-12. Meta-analysis of the intervention function cells within health-seeking behaviour with largely private benefits outcomes

Intervention function	Effect size	p-value	I sq.	Tau sq.	Tau Lower	Tau Upper	Obs.
Education	0.19	0.01	97.32	0.11	-0.47	0.84	26
Enablement	0.21	0.00	97.29	0.08	-0.33	0.75	26
Modelling	0.34	0.05	97.58	0.24	-0.61	1.30	8
Incentives	0.08	0.00	7.73	0.00	0.06	0.11	14
Persuasion	0.25	0.00	97.97	0.10	-0.38	0.88	24
Training	0.14	0.01	94.67	0.05	-0.28	0.57	19
Physical restructuring	0.05	0.69	98.46	0.09	-0.54	0.64	6
Social restructuring	0.21	0.00	97.96	0.10	-0.43	0.84	21

Source: Authors

Due to the improved legibility, we also included the meta-plot for the persuasion intervention that has the largest sample and biggest effect size among all the intervention functions. The forest plot provides information already shown in Table IV-12, but includes all the effect sizes (and the corresponding studies) that form the overall sample. As reported above, the overall (statistically significant) effect size is 0.25 (teal diamond)—indicating that the treatment group mean lies above the control group mean by 0.25 SD. The lower bound of the 95 per cent confidence interval is 0.12 SD: this implies that the effect is significantly different from a null effect (also can be seen by the non-overlapping confidence intervals with the null effect—the dashed red line).

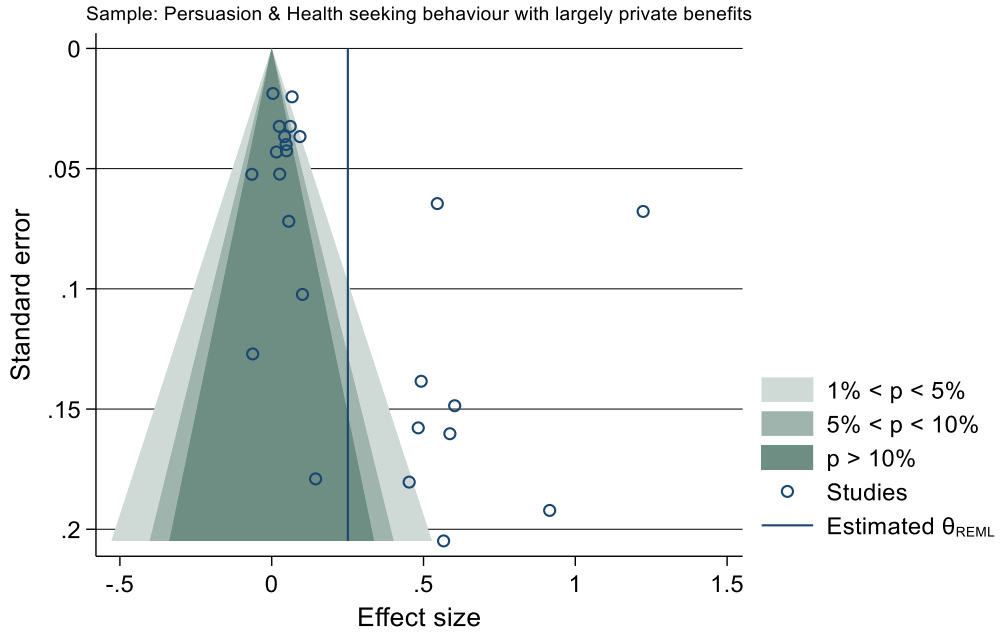
Figure IV-14. Overall effect sizes for the intervention type, Persuasion, on health-seeking behaviour with largely private benefits



Source: Authors

The funnel plot, provided in Figure IV-15, shows clear evidence of a small-study effect, such as publication bias. The regression-based Egger test also rejects the null hypothesis of small-study effects, with a beta slope (between the standard error and effect size) of 2.84 (p = 0.005).

Figure IV-15. Funnel plot for the intervention function, Persuasion, for outcomes related to health-seeking behaviour with largely private benefits



Source: Authors

c. Effects of intervention functions on health-seeking behaviour with social benefits

For the abstract outcome of "health-seeking behaviour with social benefits", the intervention function, enablement, is not only the most populous cell, but it also has the largest overall effect size of 0.37 (being larger than the conventional threshold of a small relative effect size of 0.2 by a considerable margin). The effect size shows that enablement (combined with other intervention functions) within a treatment arm leads to a higher mean of 0.37 SD in the treated sample (compared to the mean in the control sample distribution). The true effect in this sample lies, with 95 per cent confidence, between -0.74 and 1.47, implying high absolute variance as well. The funnel plot for the enablement sample suggests that there is publication bias, as seen by the number of observations lying outside the funnel, in the direction favouring treatment. The Egger test fails to reject the null hypothesis, however, indicating no clear evidence of publication bias (slope is 2.122, with a p-value of 0.11).

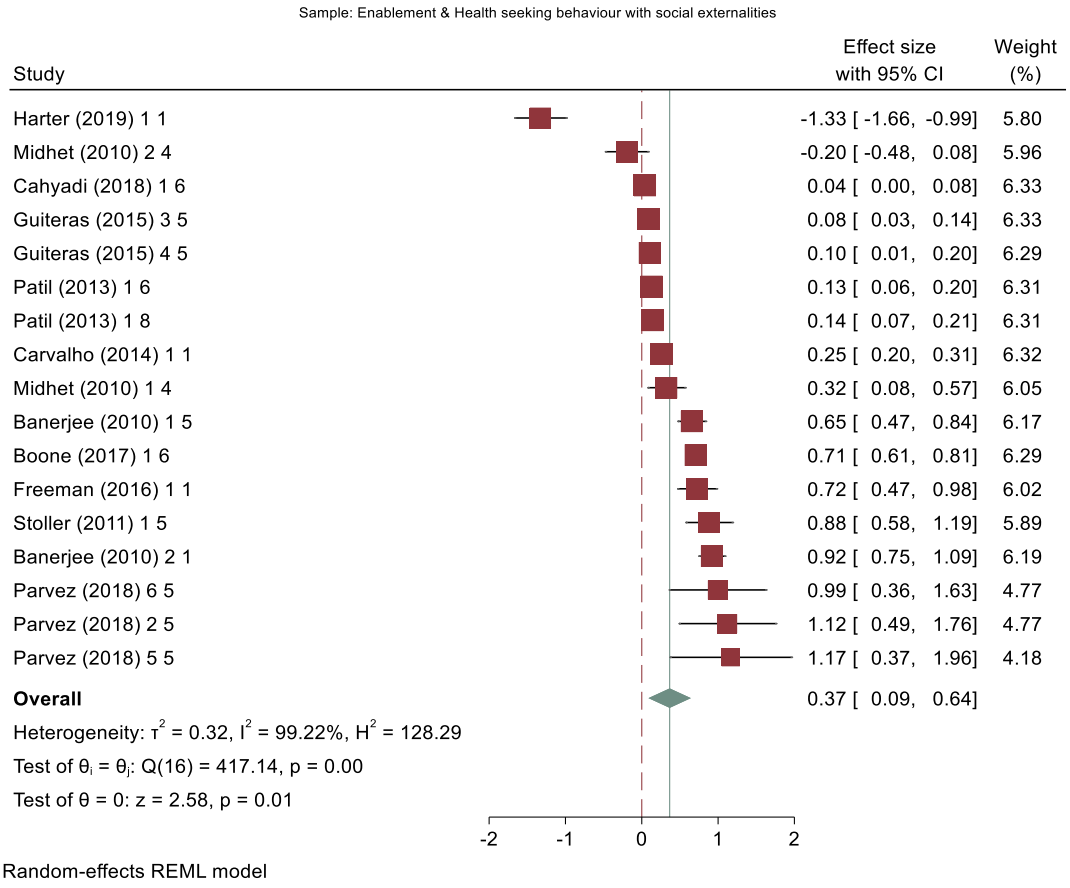
The effect sizes for two other intervention functions, that is, training and social restructuring, also lie above our defined threshold of the small effect size, at 0.33 and 0.25, respectively. All the other intervention functions (with at least more than one study) ranged in effect sizes from -0.09 to 0.30 that are not always significantly different from the null effect, and are therefore, not described.

Table IV-13. Meta-analysis of the intervention function cells within health-seeking behaviour with social benefits

INTERVENTION FUNCTION	EFFECT SIZE	P-VALUE	I SQ.	TAU SQ.	TAU LOWER	TAU UPPER	OBS.
Education	0.09	0.56	99.16	0.26	-0.92	1.10	12
Enablement	0.37	0.01	99.22	0.32	-0.74	1.47	17
Modelling	-0.09	0.80	98.31	0.56	-1.56	1.39	5
Incentives	0.16	0.12	98.08	0.08	-0.41	0.73	8
Persuasion	0.11	0.32	98.43	0.18	-0.72	0.93	16
Training	0.32	0.00	97.25	0.12	-0.35	1.00	11
Physical restructuring	0.30	0.14	98.57	0.12	-0.39	0.98	3
Social restructuring	0.25	0.01	96.79	0.08	-0.31	0.81	10

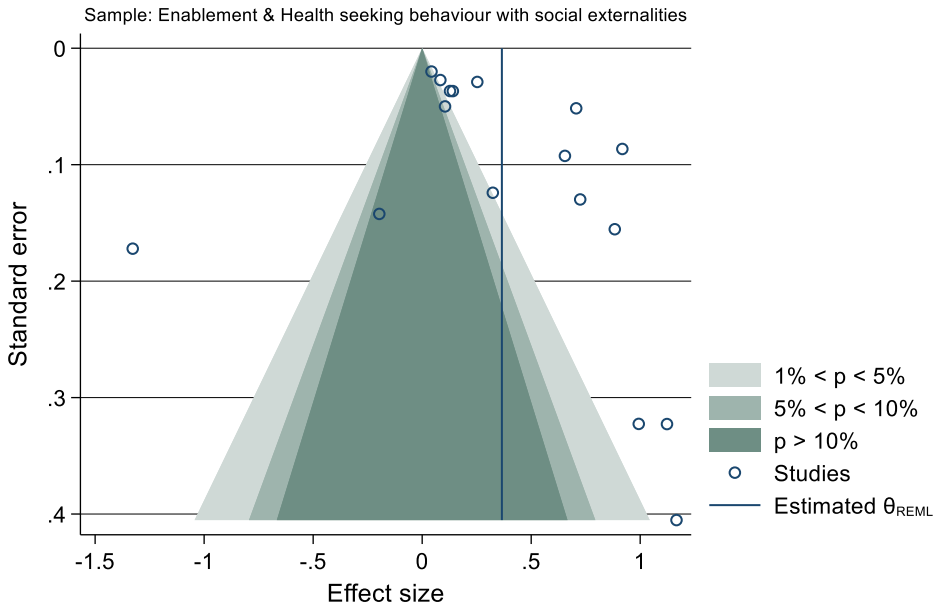
Source: Authors

Figure IV-16. Overall effect sizes of the intervention type, Enablement, on health-seeking behaviour with social benefits



Source: Authors

Figure IV-17. Funnel plot for the intervention function, Enablement, for outcomes related to health-seeking behaviour with social benefits



Source: Authors

d. Effects of intervention functions on consumption/purchasing behaviour with largely private benefits

For consumption/purchasing decisions with largely private benefits, the overall sample is rather small, and therefore, several intervention-specific samples are as small as three or four effect sizes. Estimating the effects of the eight intervention functions within treatments, the enablement cell provides the largest effect size. Statistically significant at the 99 per cent confidence level, the number of effects included is nine—just below the lower bound of the inclusion cut-off of 10. The effect size within the enablement sample is 1.12, well above the large-effect size threshold of our study (0.8). In fact, even without the largest effect size from the Parvez and others (2018) study, the effect is above the threshold of 0.8. This is observed in Figure IV-19: it estimates the overall effect size, when removing one effect size from the sample consecutively. When removing the entire Parvez and others (2018) study, however, the overall Cohen's d drops to 0.46 and it is found to be not statistically significantly different from a null effect (see online appendix O).

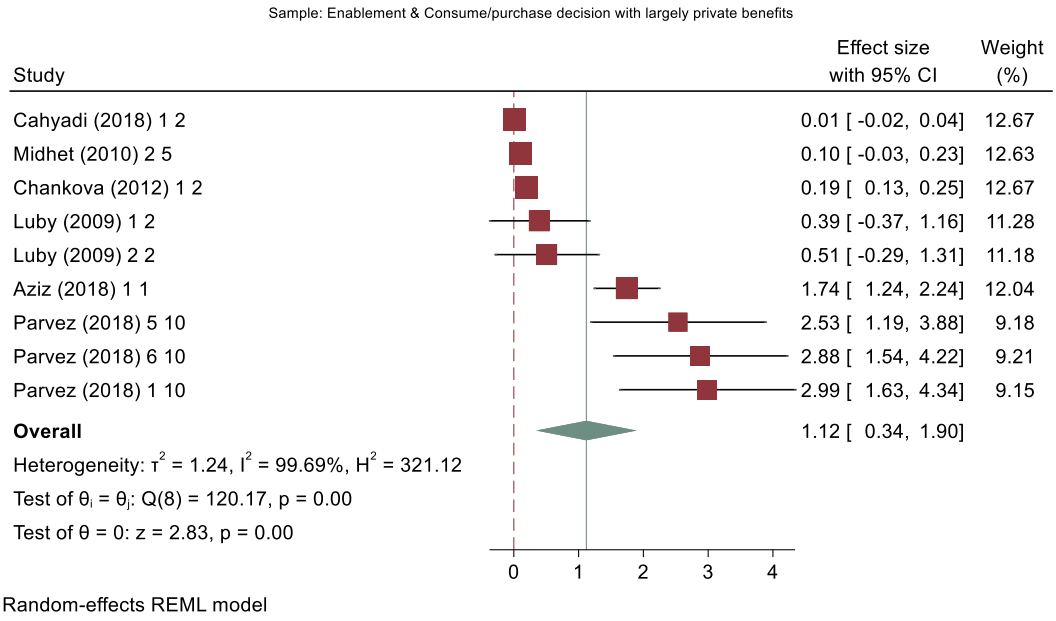
This sample also reports some of the lower values of I-squared (at above 66 per cent), indicating that the small sample size might have indeed played a role in the statistical heterogeneity observed in our sample. The funnel plot for enablement includes the "smaller" studies in our sample (large standard errors and quite small effect sizes as well). All other intervention function cells for consumption/purchasing behaviour outcomes with largely private benefits have much fewer studies (or small effect sizes), and are therefore, not investigated (see Table IV-14).

Table IV-14. Meta-analysis for the intervention function cells within consumption/purchasing decisions with largely private benefits

Intervention Function	Effect size	p-value	I sq.	Tau sq.	Tau Upper	Tau Lower	Obs.
Education	0.36	0.01	98.44	0.17	1.16	-0.44	9
Enablement	1.12	0.00	99.68	1.24	3.30	-1.06	9
Modelling	0.25	0.00	66.37	0.01	0.48	0.02	3
Incentives	0.10	0.00	76.86	0.00	0.22	-0.03	11
Persuasion	0.17	0.00	10.57	0.00	0.23	0.11	4
Training	0.52	0.16	99.19	0.54	1.97	-0.92	4
Physical restructuring							
Social restructuring	0.61	0.11	97.53	0.53	2.04	-0.82	4

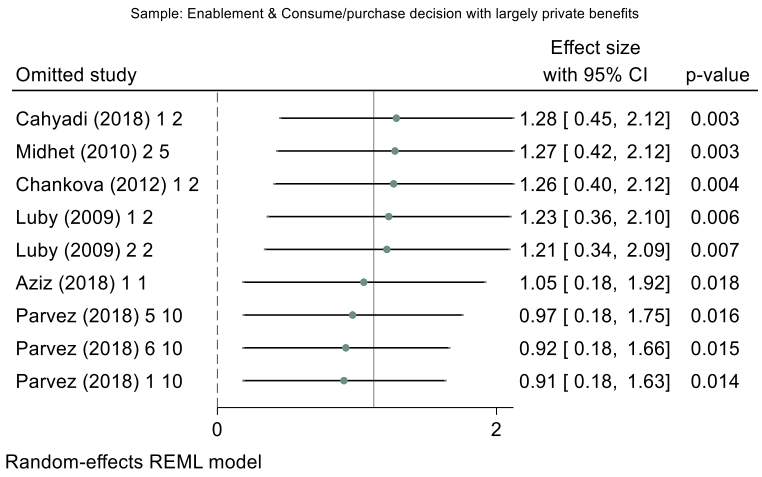
Source: Authors

Figure IV-18. Overall effect sizes for the intervention type, Enablement, on outcomes indicating consumption/purchasing decisions on largely private benefits



Source: Authors

Figure IV-19. Leaving one out analysis for Enablement, within consumption/purchasing decision with largely private benefits outcomes

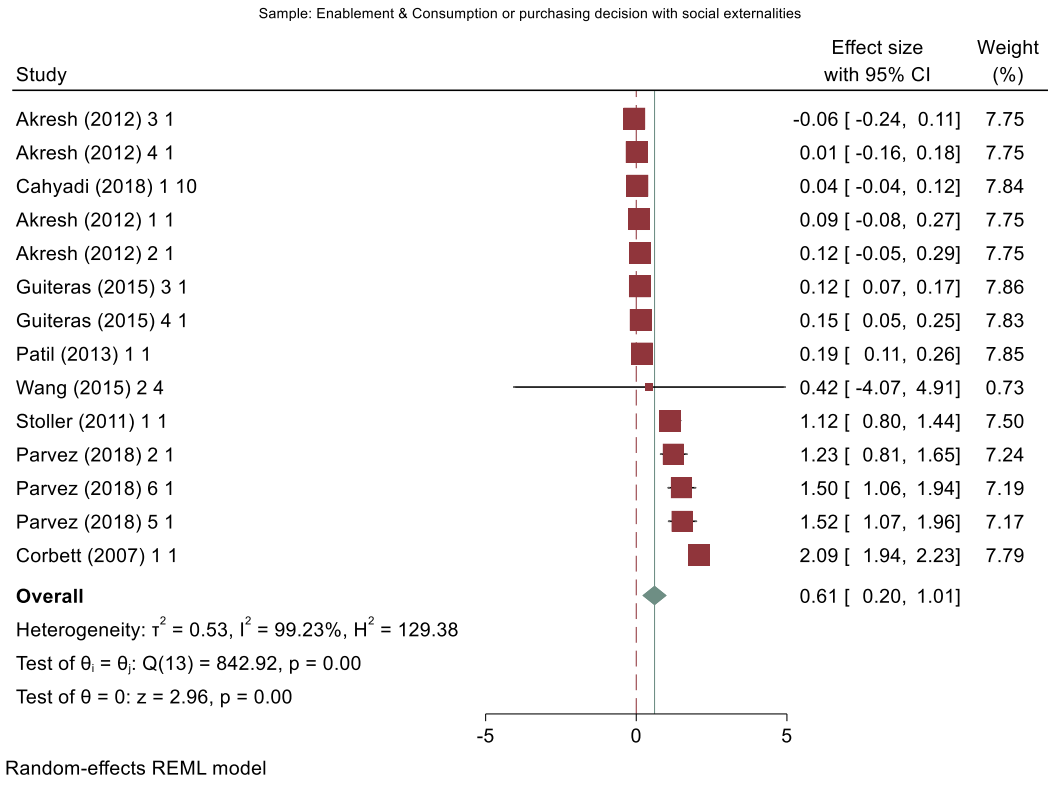


Source: Authors

e. Effects of intervention functions on consumption/purchasing behaviour with social benefits

For the last abstract outcome, we report the intervention type, enablement, again, as the most populous cell, with the largest overall effect size. With an effect size of 0.61, enablement falls within the range of medium-sized effect sizes according to our thresholds for outcomes in consumption/purchasing decisions with social benefits. As can be seen, this effect is statistically significant (95 per cent significance level) and does not appear to be driven by a few studies, although Corbett and others (2007) is driving the effect considerably (and removing it leads to a lower overall effect size of 0.47—see Figure IV-21). The heterogeneity is similar to the other forest plots, quite high in absolute and relative terms (Tau and I-squared, respectively).

Figure IV-20. Overall effect sizes for the intervention type, Enablement, on consumption/purchasing decisions with social benefits



Source: Authors

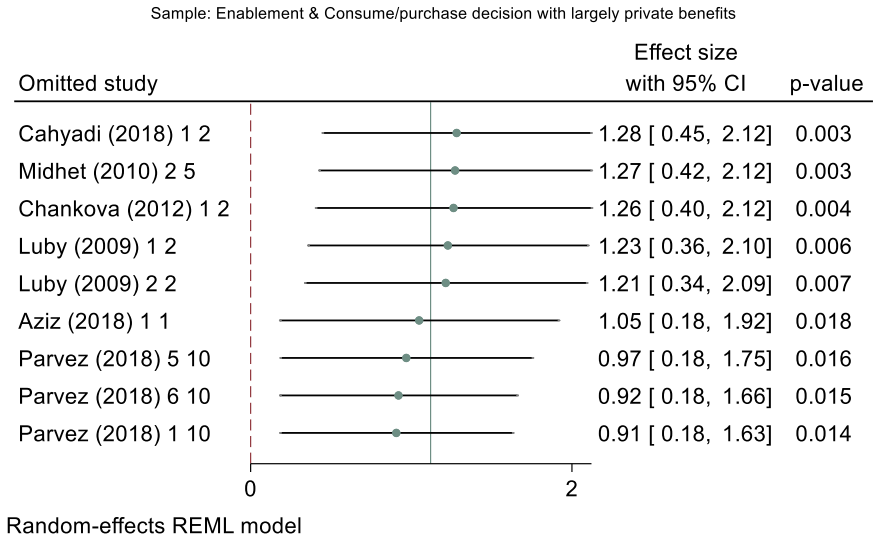
The funnel plot in Figure IV-22 does not clearly indicate evidence of publication bias either, as no clear correlation between effect sizes and standard error can be observed, although the effects are largely located to the right of the line of the effect size (favouring treatment). The regression-based Egger test, however, fails to reject the null hypothesis of small-study effects (p-value = 0.26).

Table IV-15. Meta-analysis for the intervention function cells within consumption/purchasing decisions with social benefits

Intervention function	Effect size	p-value	I sq.	Tau sq.	Tau Upper	Tau Lower	Obs.
Education	0.37	0.05	99.37	0.30	1.44	-0.70	10
Enablement	0.61	0.00	99.23	0.53	2.04	-0.83	14
Modelling	0.58	0.34	99.67	0.74	2.27	-1.11	2
Incentives	0.09	0.12	90.59	0.01	0.32	-0.15	5
Persuasion	0.39	0.03	99.42	0.39	1.62	-0.83	13
Training	0.24	0.13	98.67	0.15	0.99	-0.51	6
Physical restructuring	0.13	0.00	0.00	0.00	0.13	0.13	4
Social restructuring	0.26	0.17	99.16	0.21	1.16	-0.65	6

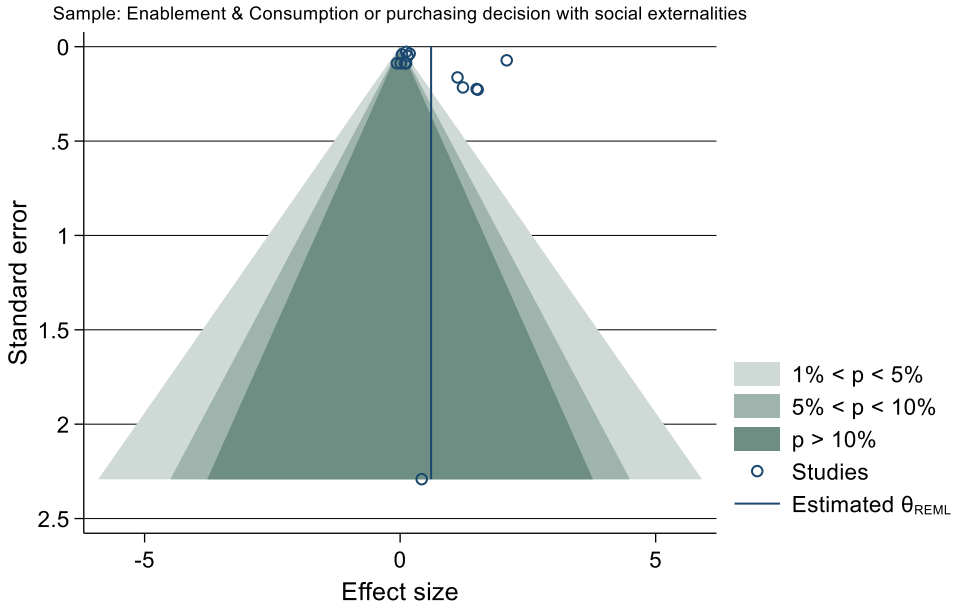
Source: Authors

Figure IV-21. Leaving one out analysis for enablement on consumption/purchasing decisions with social benefits



Source: Authors

Figure IV-22. Funnel plot for the intervention function, Enablement, for outcomes related to consumption/purchasing behaviour with social benefits



Source: Authors

We also present the results for the other intervention functions in Table IV-15. Besides enablement, two other intervention functions—persuasion and education—also report overall effect sizes of 0.39 and 0.37, respectively.

f. Meta-regression

Due to the large variations in methodology, region, and intervention function combinations, we used a meta-regression set-up to explore some of the sources of heterogeneity in effect sizes. Crucially, only through meta-regressions could we attempt to isolate the effects of individual intervention functions by controlling for other intervention functions that are also part of the intervention. Note, however, that the possibility to do so was limited: we could only study correlations based on the

combinations of intervention functions used in our sample of studies. These different combinations do not resemble an experimental (orthogonal) design, where only an intervention function would be changed at a time, keeping everything else constant, including moderators such as study region. Instead, we encountered a bunching of certain combinations of intervention functions, limiting the statistical power to differentiate between the effects of intervention functions. For instance, modelling was always found in combination with education, thereby rendering a clear and coherent effect size there unfeasible. It was, therefore, not included in the meta-analysis. Nevertheless, this analysis may provide a robustness check for the results of the preceding section, as well as suggestive evidence on the relative importance of the intervention functions.

Table IV-16 provides the first set of results from the meta-regression. In the table, the first column tabulates the effect of each intervention function on all four abstract outcomes (overall sample). The second and third specifications progressively include dummies for the methods used in the paper and regions where the studies are conducted, respectively. The last two specifications pertain to the abstract outcome of the subgroups' health-seeking behaviour and consumption/purchasing decision, where the private and social components were grouped together to achieve a large sample size (without which no regression on the subsample of consumption/purchasing decision would be feasible). These specifications also include both method and region dummies.

Table IV-16. Regression results for the overall sample, as well as by each abstract outcome subgroup

	Overall sample	Overall with method dummies	Overall with method and region dummies	Health-seeking behaviour	Consumption / Purchasing decisions	Largely private benefit outcomes	Social benefit outcomes
	b se (row below)	b se (row below)	b se (row below)	b se (row below)	b se (row below)	b se (row below)	b se (row below)
Education	0.047	0.039	0	-0.079	0.27	0.167	-0.245
	0.096	0.077	0.086	0.098	0.213	0.085	0.242
Enablement	0.408***	0.145*	0.037	0.265***	0.740***	0.264***	0.533***
	0.069	0.068	0.075	0.07	0.155	0.069	0.126
Incentives	0.076	-0.03	-0.11	0.052	0.086	0.089	0.154
	0.081	0.076	0.082	0.081	0.197	0.081	0.151
Persuasion	0.039	0.01	0.043	0.137	-0.257	-0.089	0.051
	0.112	0.087	0.086	0.109	0.31	0.107	0.223
Training	0.055	-0.004	-0.084	0.088	0.118	0.058	0.173
	0.109	0.091	0.092	0.106	0.311	0.101	0.243
Social restructuring	-0.425**	-0.269*	-0.281*	-0.265	-0.729	-0.353*	-0.307
	0.161	0.131	0.129	0.152	0.486	0.152	0.331
Physical restructuring	0.047	0.039	0	-0.079	0.27	0.167	-0.245
	0.096	0.077	0.086	0.098	0.213	0.085	0.242
Observations	136	136	136	86	50	78	58

	Overall sample	Overall with method dummies	Overall with method and region dummies	Health-seeking behaviour	Consumption / Purchasing decisions	Largely private benefit outcomes	Social benefit outcomes
	b se (row below)	b se (row below)	b se (row below)	b se (row below)	b se (row below)	b se (row below)	b se (row below)
R-squared	0.000	39.303	42.965	0.000	0.000	0.000	0.000
H squared	88.945	48.385	42.007	57.402	107.878	40.232	106.227
I squared	98.876	97.933	97.619	98.258	99.073	97.514	99.059
Tau-squared	0.195	0.108	0.102	0.126	0.345	0.084	0.32
p-value <	0.000	0.000	0.000	0.000	0.000	0.000	0.001
Chi squared	56.185	183.43	202.882	30.4	32.578	42.18	23.615

Source: Authors
Notes: Standard errors (se) are displayed under the beta coefficients (b). Significance is indicated as p<0.05, ** p<0.01, *** p<0.001.

The results of these regressions suggest a similar picture to those shown previously in the meta-analysis forest plots: enablement is the only intervention function that appears to affect abstract outcomes positively, across all specifications. The overall sample regression suggests that the effect size is about 0.41, that is, nearly a moderate-sized effect; it is reduced to 0.15 after the inclusion of dummies for the different methods used to obtain the raw coefficients in their original studies. The regressions, including regional dummies, do not have statistically significant results. Upon checking the subsamples for health-seeking behaviour versus consumption behaviour, the effect is nearly triple in the latter, implying that enablement affects consumption/purchasing outcomes in a much more effective manner. Examining the samples along the social benefit versus private benefits dimension, enablement also has double the effect on outcomes with social benefits, compared to outcomes with largely private benefits.

Social restructuring is another intervention function with a significant (at five per cent level) and a moderately large coefficient of -0.42, implying that its inclusion leads to a reduced change in behaviour in the desired direction. In the set-up of this regression, it basically implies that this intervention, in comparison to the other intervention functions, is the least effective in bringing about the desired behavioural change. This effect stems largely from the private benefits subsample, where the coefficient is 0.35 and weakly significantly different from zero. It is important to caveat the findings from the consumption/purchasing decision subsample that has only 29 total observations, with very few observations for some intervention functions. These regression results also suggest a high degree of variability in the study estimates (as shown by the very large I-squared and Q-value).

We additionally ran regressions that focused on the subsamples where the respective intervention function is part of the treatment (therefore, excluding papers where this intervention function is not included). We did so in order to disentangle the effects of the other intervention functions that are combined with it. We do not report the results for the modelling and physical restructuring subsamples that have very small samples, and therefore, provide insufficient degrees of freedom. From these regressions, we find consistently positive additional effects for the enablement and social restructuring intervention functions. Only enablement, however, reaches statistical significance, in the training and persuasion samples. As can be seen, in interventions with

persuasion, enablement has a small effect size of 0.22, while in the training sample, it has a moderate effect size of 0.52. Overall, the results, both from the meta-analyses and the meta-regression, single out enablement as the most effective intervention function within our set of studies.

Table IV-17. Regression results for each intervention function sample (with more than five degrees of freedom remaining)

	EDUCATION	ENABLEMENT	INCENTIVES	PERSUASION	TRAINING	SOCIAL RESTRUCTURING
	b se (row below)	b se	b se	b se	b se	b se
Education		-0.156	0.012	0.11	0.053	0.355**
		0.261	0.077	0.171	0.111	0.132
Enablement	0.045		0.220**	0.094	0.350*	0.024
	0.148		0.073	0.155	0.138	0.165
Incentives	-0.097	0.252		-0.303	-0.237	-0.385*
	0.131	0.283		0.212	0.17	0.173
Persuasion	0.042	-0.014	-0.083		0.096	0.155
	0.146	0.24	0.13		0.125	0.13
Training	0.044	0.336	0.005	0.07		-0.07
	0.143	0.287	0.123	0.185		0.124
Social restructuring	0.235	0.252	0.088	0.19	0.088	
	0.132	0.279	0.104	0.15	0.108	
Physical restructuring	-0.118	-0.177	0	-0.27	-0.454**	-0.183
	0.213	0.295		0.246	0.17	0.186
Observations	57	66	38	57	40	41
R-squared	0.000	0.000	0.000	0.000	14.971	0.000
H squared	83.566	231.13	11.617	91.308	40.569	55.45
I squared	98.803	99.567	91.392	98.905	97.535	98.197
Tau-squared	0.211	0.514	0.021	0.219	0.115	0.136
p-value	0.118	0.513	0.007	0.079	0.001	0.002
Chi squared	10.164	5.246	15.901	11.318	22.73	21.164

Source: Authors

g. Unconverted effect sizes in the utilization of the health services subsector

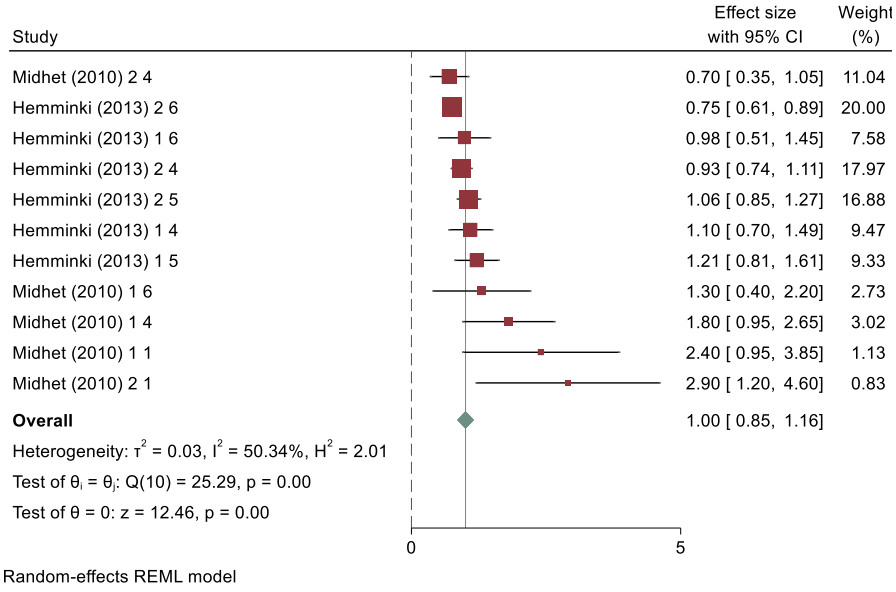
The descriptive statistics in Table IV-7 show that the most common methods found in our meta-analysis data are regressions, followed by risk difference and odds ratios. In this section, we focus on results from one subsector, namely, the utilization of health services, to provide evidence with non-standardized effect sizes. For this purpose, we provide two different non-standardized effects,

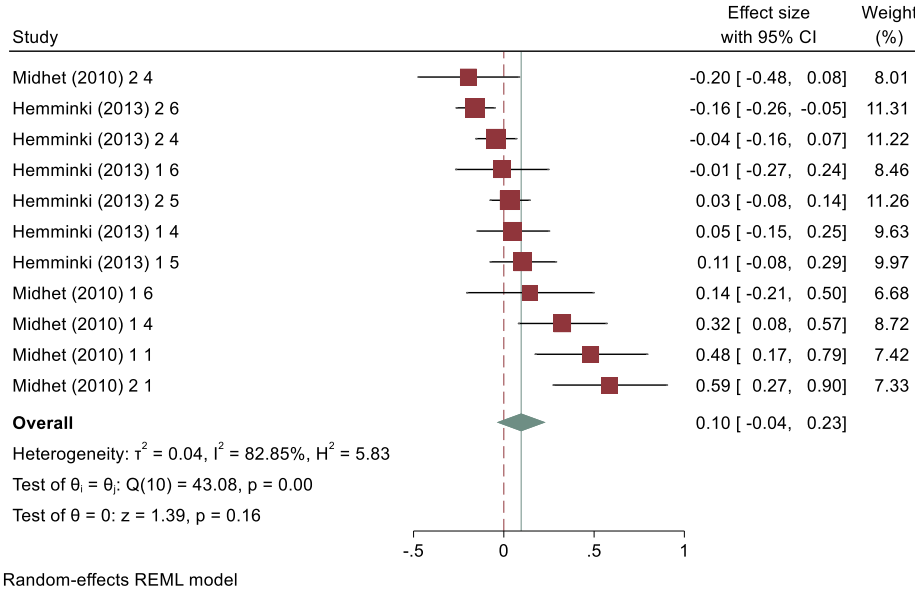
odd ratios, and beta coefficients from regressions, where we find the largest number of non-standardized effect sizes.

Figure IV-23 compares the standardized and non-standardized effect sizes within the odds ratio coefficients. This comparison finds that a statistically insignificant Cohen's d of 0.10 represents a statistically significantly higher likelihood of one per cent for the treated group to partake in favourable behaviour (health-seeking and consumption/purchasing). Therefore, the insignificant Cohen's d is comparable to the small effect found using the unstandardized coefficients. The funnel plot of the non-standardized effect size (Figure IV-25) also suggests the presence of publication bias, as evidenced by the asymmetry around the effect line, such that studies with larger standard errors tend to have larger effects.

Figure IV-24 compares the coefficients of the regression results, where the results are more comparable, and a 0.10 effect size in Cohen's d translates to a null effect size (in the treatment group), both being statistically insignificant. The funnel plot for the beta coefficients also indicates the presence of publication bias, with most of the effect sizes situated around the significance levels, larger effects with larger standard errors, and also in the preferred direction (indicated by the larger number of observations on the right-side favouring treatment).

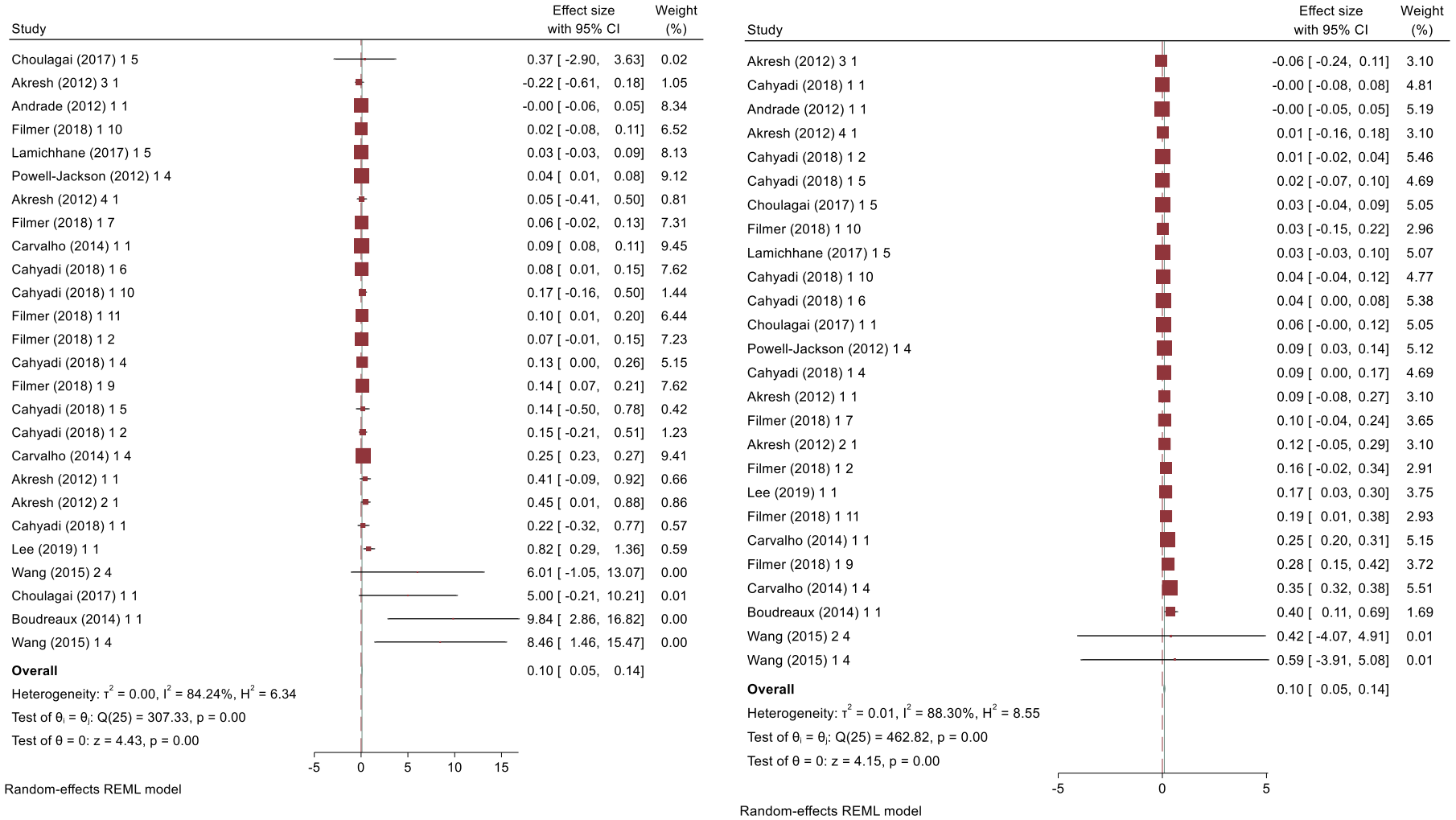
Figure IV-23. Overall effect size for non-standardized odds ratios (left) and standardized Cohen's d (right) for the utilization of the health services sector





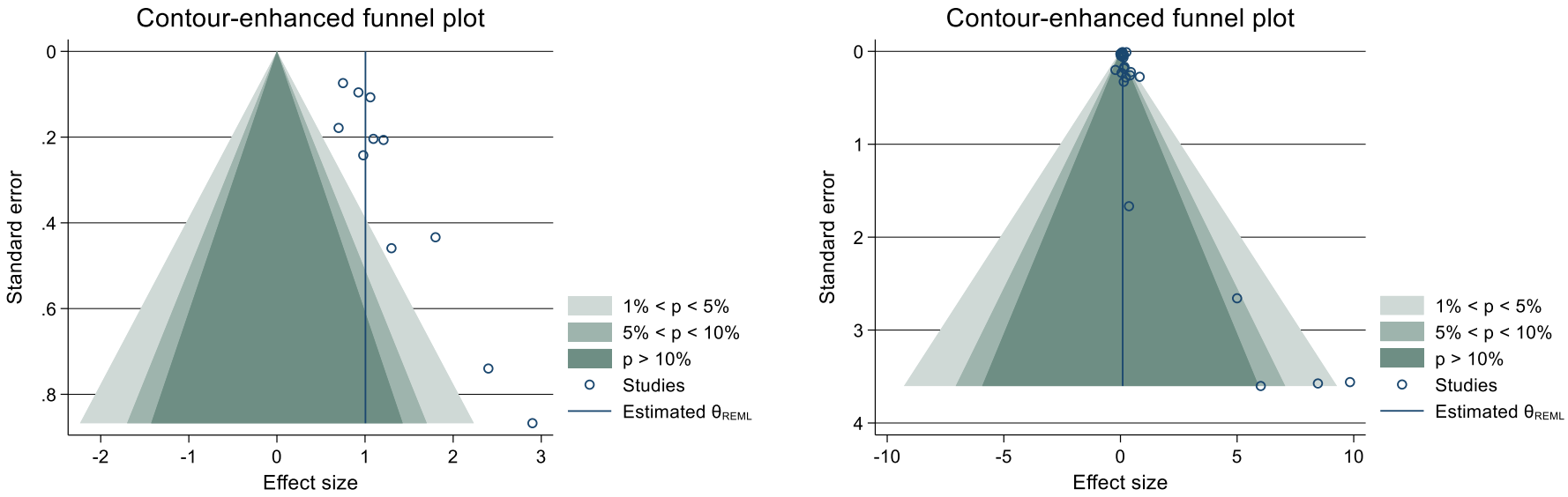
Source: Authors

Figure IV-24. Overall effect size for non-standardized regression beta coefficients (left) and standardized Cohen's d (right) for the utilization of the health services sector



Source: Authors

Figure IV-25. Funnel plot of the raw effect sizes for odds ratio (left) and regressions (right)



Source: Authors

2. ENERGY

We conducted systematic reviews in the form of a meta-analysis for two combinations of interventions and outcomes of high policy relevance: the effects of electrification on formal employment[21] and the effects of the Chinese pilot ETS on GHG emissions. We start with the results of the first topic.

a. Effects of electrification on employment

Most of the 15 studies in this meta-analysis do not report an employment outcome that is explicitly restricted to formal employment. In our main analysis, we use outcomes from all studies and test whether the type of employment outcome (formal or general employment including formal) matters for the overall effect.

A typical study in this category explores the effects of the availability of on-grid electricity on changes in time use, including productive activities, at the household level. Direct or indirect employment effects, due to the construction of infrastructure, were not the focus in our sample of studies. As all included studies use the same intuitive outcome scale—the probability of being employed, we ran the meta-analysis on the non-standardized coefficients first. Most studies report outcomes separately for men and women. Since men and women are separate samples and different effects may be expected for men and women, we entered both of them into the meta-analysis.

Furthermore, some of the studies report more than one model specification. We chose the specification that, according to the authors and general rules on the validity of impact evaluation methods, gave the most precise estimate of the true effect. For instance, specifications that combined DiD with PSM were favoured over specifications that only used DiD.

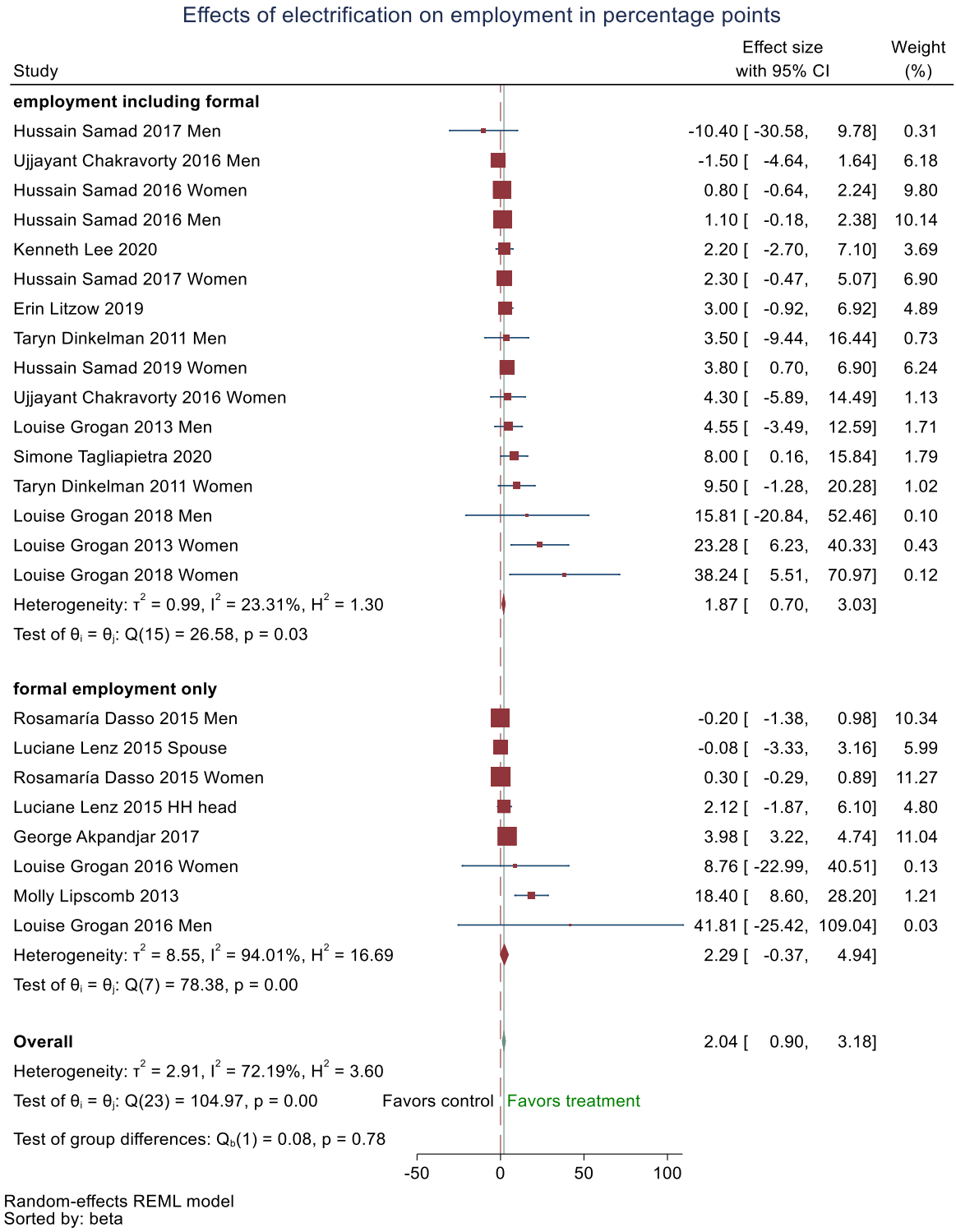
We present the results of the non-standardized effect sizes in the forest plot in Figure IV-26, based on a random-effects model. Note that we labelled the studies in the energy sector by the first author's full name and year in order to differentiate the studies. The forest plot shows a significant overall effect of around two percentage points. This implies that electrification leads, across all studies, to an increase in the employment rate by around two percentage points on average among households with access to electricity. The 95-per cent confidence interval for the average of the true effects ranges from 0.9 to 3.18 percentage points.

The forest plot shows no significant difference (p=0.78) between studies that measure formal employment explicitly and those measuring wider measures of employment (including formal employment). This is good news for further analysis, and we can, consequently, use the entire sample, rather than focus only on the (small) subset of studies measuring formal employment. There seems to be more heterogeneity within this subset, both in absolute and relative terms (I-squared and T-squared—the estimate of Tau-squared), than in the one using the broader measures of employment. Given the small number of studies (five) on formal employment, this may be purely coincidental. Overall, 72 per cent of the variance (I-squared) cannot be explained by the sampling error, and is therefore, accounted for by heterogeneity in true effect sizes. The large sample sizes, and therefore the small within-study variation, make this relative measure of heterogeneity uninformative. Instead, we focused our interpretation on the absolute measure of heterogeneity—the variance of the true effect sizes—estimated by T-squared. Its standard deviation, T, enabled us to predict that 95 per cent of the true effect sizes would fall within the range of -1.30 to +5.38

[21] We excluded "informal employment", as it was defined by the authors. Typically, they refer to the 15th International Conference of Labour Statisticians (International Labour Organization, 2003).

percentage points.[22] In other words, an electrification programme may lead, with a non-negligible probability, to a negative effect or also to a substantial positive effect of five percentage points.

Figure IV-26. Effects of electrification on employment (forest plot)



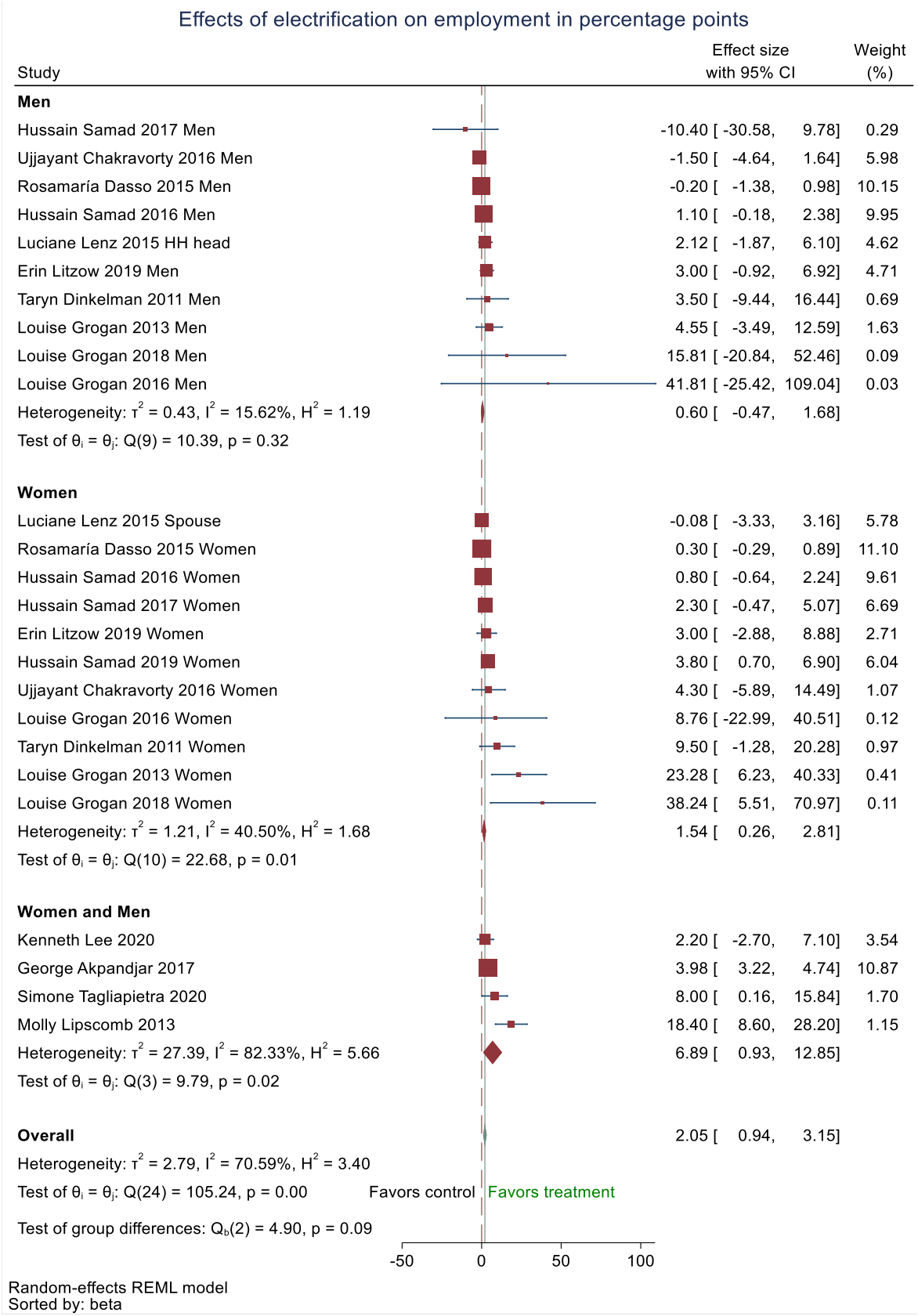
Source: Authors

[22] We assume a normal distribution of true effect sizes. In this case, 95 percent of true effect sizes lie within 1.96 SD (estimated by T) around the mean of 2.04.

As a first step towards exploring this substantial heterogeneity, we ran a subgroup analysis stratified by gender. The results suggest that the effects are stronger for women (1.54 percentage points) than for men (a mere 0.6 percentage points), though the difference between the subgroups of women and men fails to be significant (p=0.27). The overall estimated effect size for these two sets of studies drops to 0.86 percentage points because considerably larger effects are found in studies that only report their results pooled for both genders. As there are only four studies in total, of which two are—as we shall see below— influential outliers, we refrain from reading too much into this result. Furthermore, the male subsample does not provide any statistically significant evidence on positive employment effects for men (the 95 per cent confidence interval ranges from -0.47 to 1.68 percentage points). In order to keep one influential study that does not report effects separately for female and male subsamples (Lenz and others, 2015) in the pool of studies for this gender analysis, we attributed its results reported for household heads to men and those for spouses to women, even though this would not be correct in approximately 28 per cent of the cases.[23] Excluding this study would have led to even larger differences between the genders: 2.05 percentage points for women versus 0.5 percentage points for men. It should also be noted that splitting the analysis by gender reduced the between-study heterogeneity substantially, thus supporting the conjecture that the employment effects of electrification might be different for women and men.

[23] In Habimana and Pasqua (2017), this is the percentage of female-headed households in a sample of around 6,900 households in Rwanda.

Figure IV-27. Effects of electrification on employment by gender (forest plot)

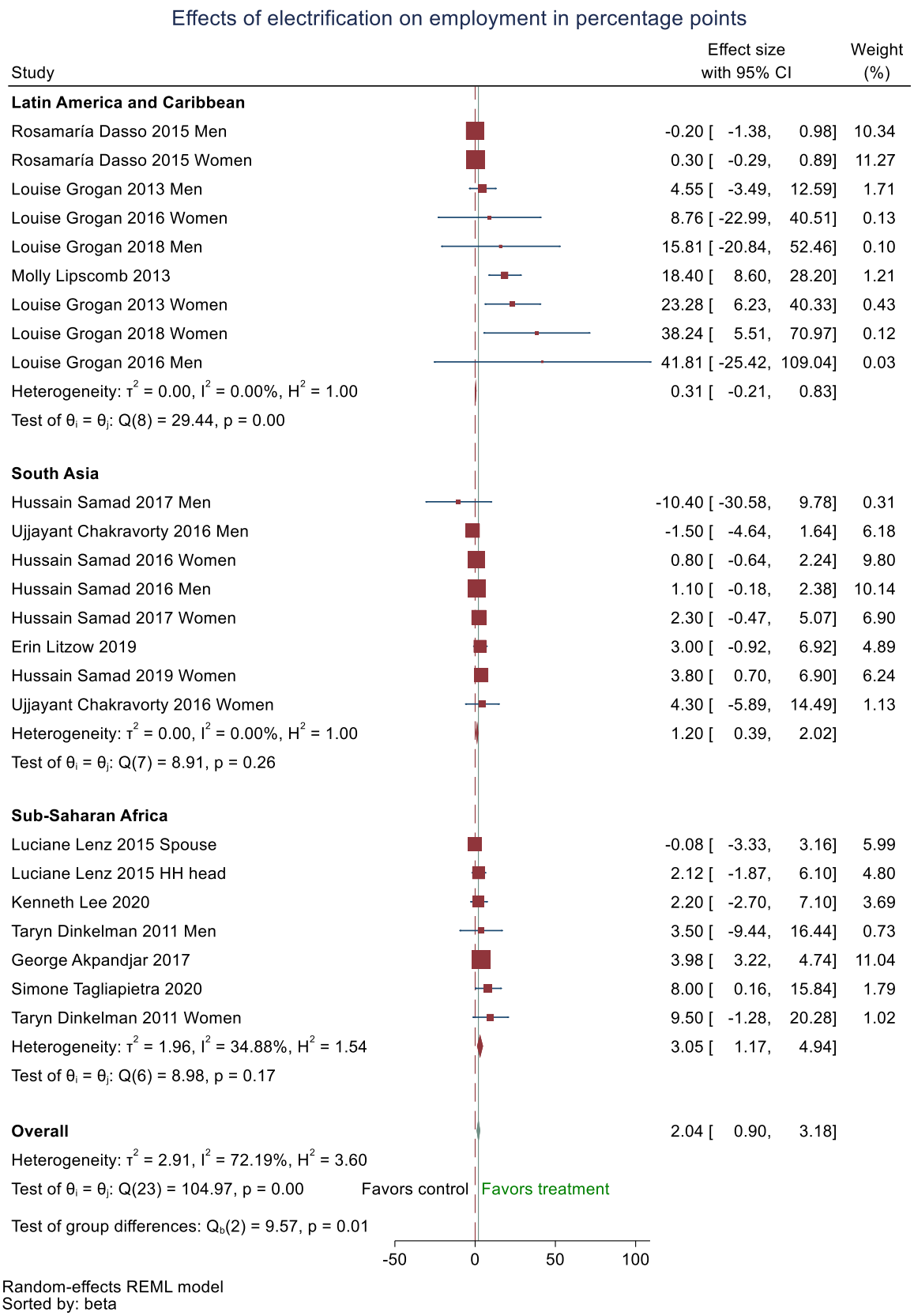


Source: Authors

In the next step towards understanding heterogeneity, we looked at the regional variation in effect sizes. As it turns out, studies conducted in Sub-Saharan Africa show substantially larger effects than

those in Latin America and the Caribbean as well as those in South Asia (only at a weak significance level in the latter case). As a note of caution on this result, the number of observations in each region is relatively small and the number of different interventions is even lower (five in each region).

Figure IV-28. Subgroup analysis based on region

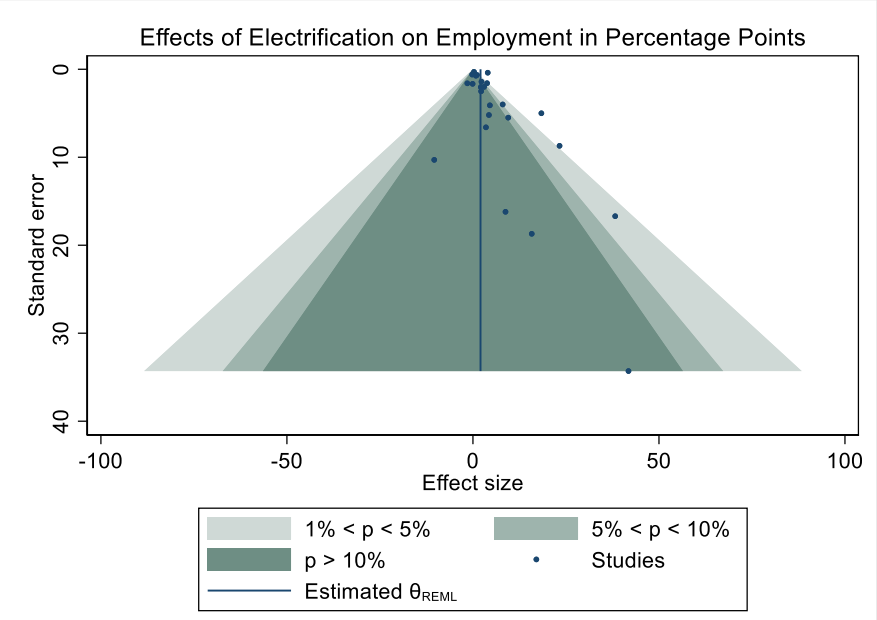


Source: Authors

Next, we looked into the role of sample sizes, that is, we tested for the "small-study effect". The visual inspection of the funnel plot is not clear-cut, as most studies have effect size estimates with

small standard errors. They seem to be symmetrically distributed around the estimated overall effect size. However, a few studies with larger standard errors are suspiciously close to the significance boundaries and mostly on the side of positive effects. To test more formally for the presence of a small-study effect, we ran an Egger test to regress the effect sizes on their standard errors, which reveals a significant correlation between the magnitude of effect sizes and their standard errors ($p<0.001$). We do not see a potential reason for this correlation other than publication bias, though we cannot rule out other sources of heterogeneity that could explain this pattern.

Figure IV-29. Funnel plot for the effects of electrification on employment



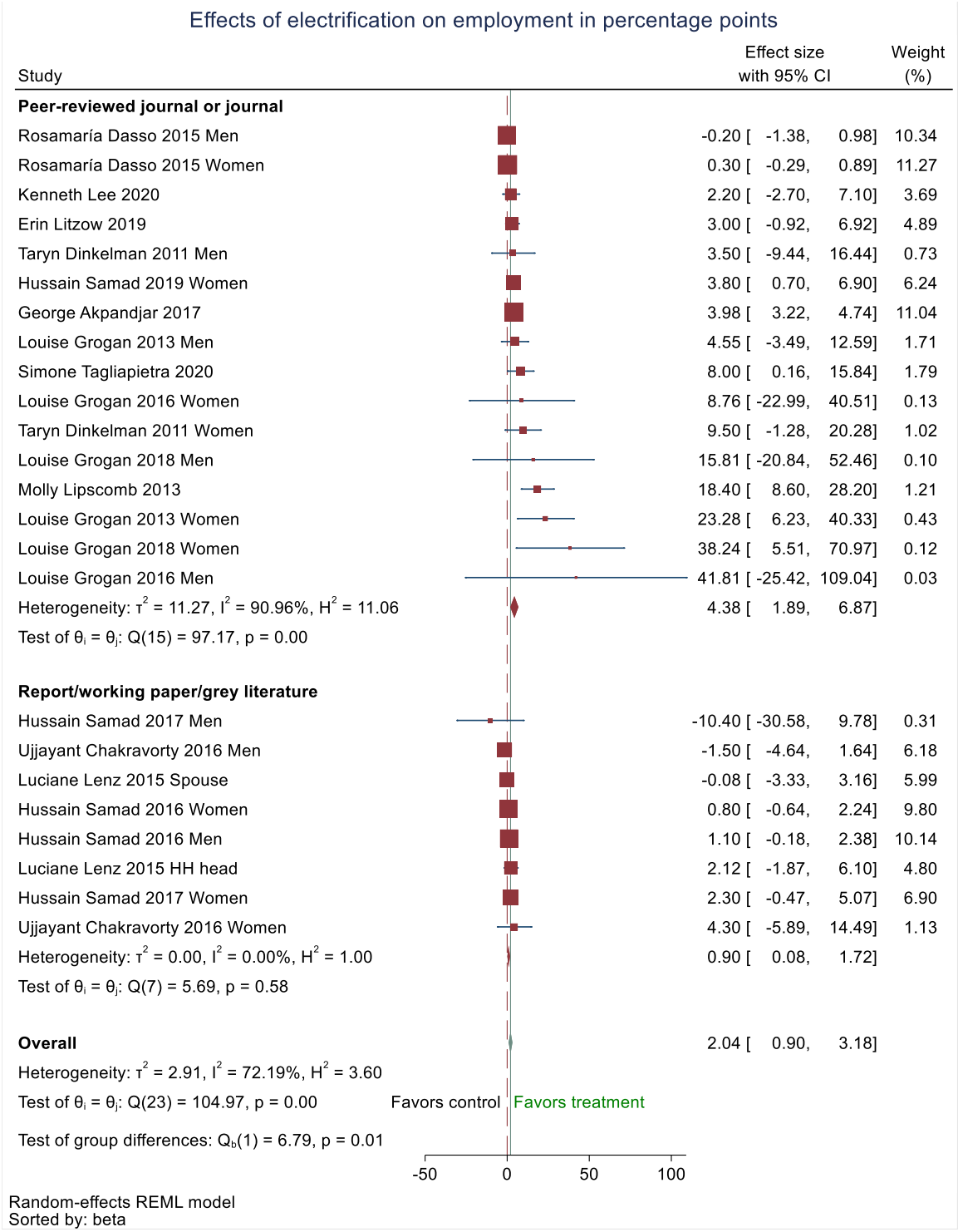
Source: Authors

In order to further explore the likelihood of publication bias, we ran a subgroup analysis, based on whether the study has been published in a journal or in grey literature. In fact, strong differences emerge (see Figure IV-30), with published articles having almost five times the overall effect size than grey literature. Interestingly, all the papers within grey literature show insignificant effects. However, when combining them through a meta-analysis, the overall effect is, nevertheless, significantly different from zero ($p=0.03$, not reported in the forest plot), though modest in economic terms—an increase in employment of 0.9 percentage points. This hints at the individual studies being underpowered to detect the small overall effect size. It shows, at the same time, the well-known benefits of the meta-analysis in overcoming this problem. Further evidence that the initial overall mean of two percentage points may be inflated comes from the subgroup analysis, based on the risk-of-bias assessment (see Figure IV-31). Studies rated as high confidence in their study findings have a lower overall mean (1.08 percentage points) than studies rated as medium- or low-confidence (3.15 and 2.74 percentage points, respectively), though the subgroup differences are only weakly significant ($p=0.09$).

What may be the true mean of the effects of electrification? Based on our piecing together of the pieces of evidence (notably the asymmetry in the funnel plot driven by published studies with comparatively small samples, which does not exist for the grey literature), we should give more weight to the estimates within the grey literature than to the published studies. This would thus lower our estimate of the average of the effect sizes below the two percentage points estimated in the first forest plot. A subgroup analysis, based on the risk-of-bias assessment and an inspection of

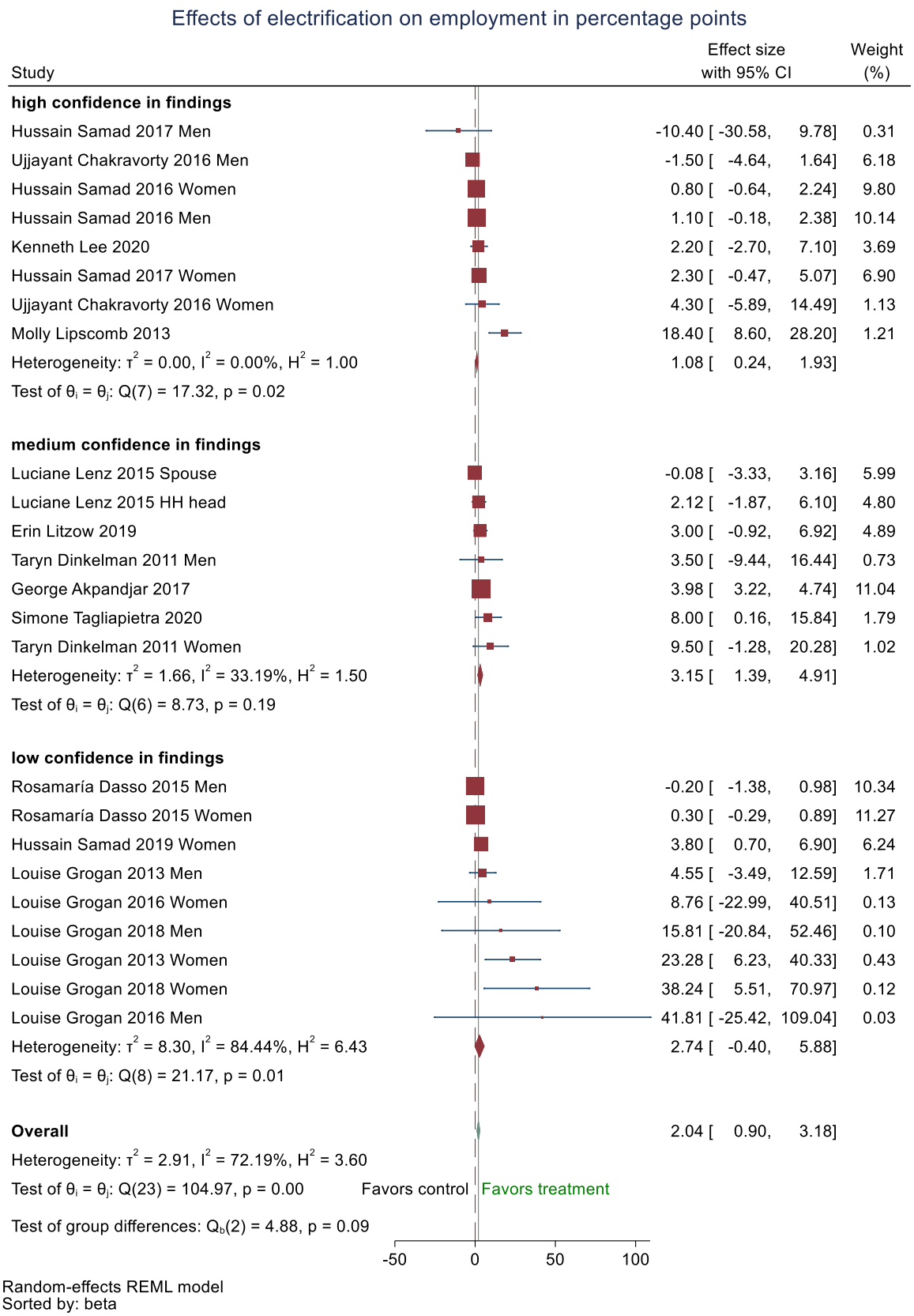
outliers (Figure IV-32), provides further evidence that the initially detected two percentage points are likely to be an overestimation of the true mean. The most influential study is Akpandjar and Kitchens' (2017) published study. Merely excluding this one study from the meta-analysis would lower the overall effect to 1.40 percentage points (see the effect size estimate of 1.40 in the respective row in Figure IV-32). On a more positive note, the subgroup analyses tentatively suggest that effects are stronger for women and when the intervention is conducted in Sub-Saharan Africa.

Figure IV-30. Subgroup analysis for testing differences between grey and published literature



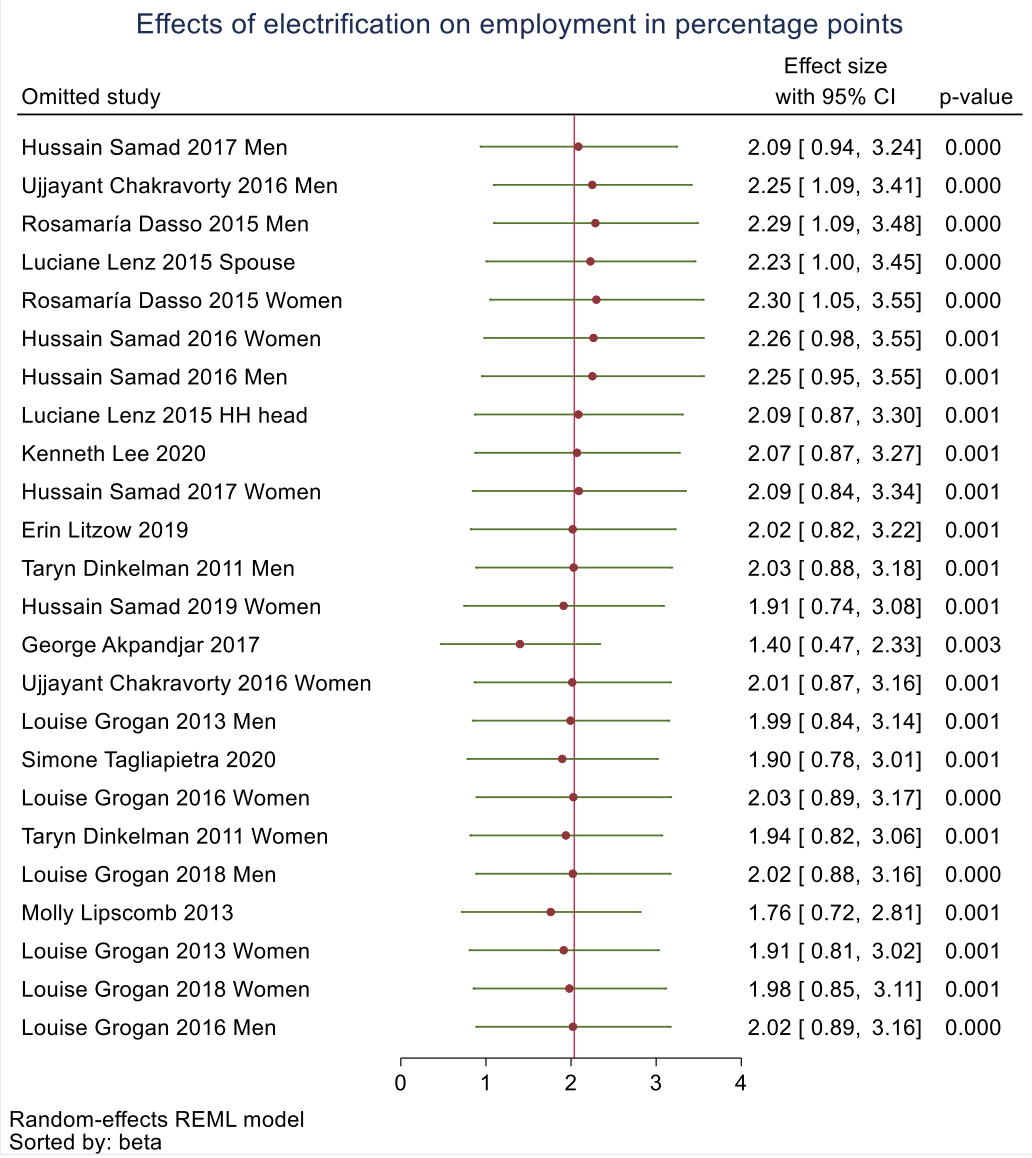
Source: Authors

Figure IV-31. Subgroup analysis based on the risk-of-bias assessment



Source: Authors

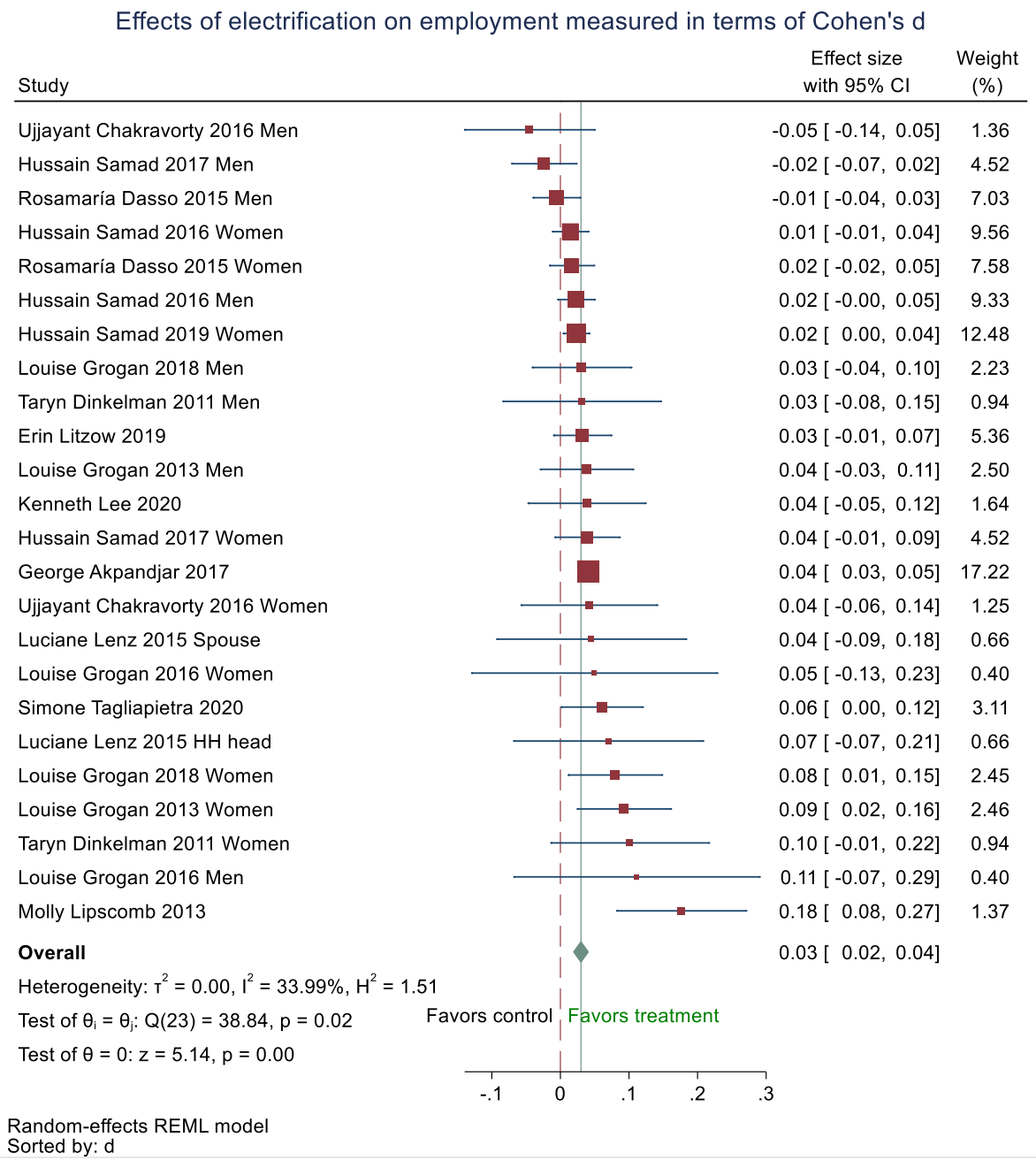
Figure IV-32. Outlier inspection: estimating the overall effect size of electrification while excluding one study at a time



Source: Authors

How does the estimated overall effect of at most two percentage points, possibly only to be expected for women, relate to transformational change? Is this a large or a small effect? One way to answer this is by converting the effect sizes into standardized effect sizes for which conventions exist that distinguish large and small. To this goal, we repeated the meta-analysis, this time based on Cohen's d. As it turns out, the estimated Cohen's ds are far below the threshold of a small effect (0.2). The overall Cohen's d is just 0.03—hardly what we aimed for in our quest for transformational change. Does this imply that the increase in employment of roughly two percentage points is a negligible effect? Not necessarily. This example may show the limits of statistical conventions on effect-size magnitudes. To judge whether two percentage points are large or small, one needs to compare this intervention to other interventions that aim to increase (female) employment, including their cost and other benefits. It is beyond the scope of this evidence synthesis to do so. Therefore, we can only conclude here that electrification, based on the magnitude of its effects on formal employment, as measured in experimental and quasi-experimental literature, falls short of our goal of finding transformational change, as defined by the three criteria used in this report.

Figure IV-33. Effects of electrification on employment measured, in terms of Cohen's d (forest plot)



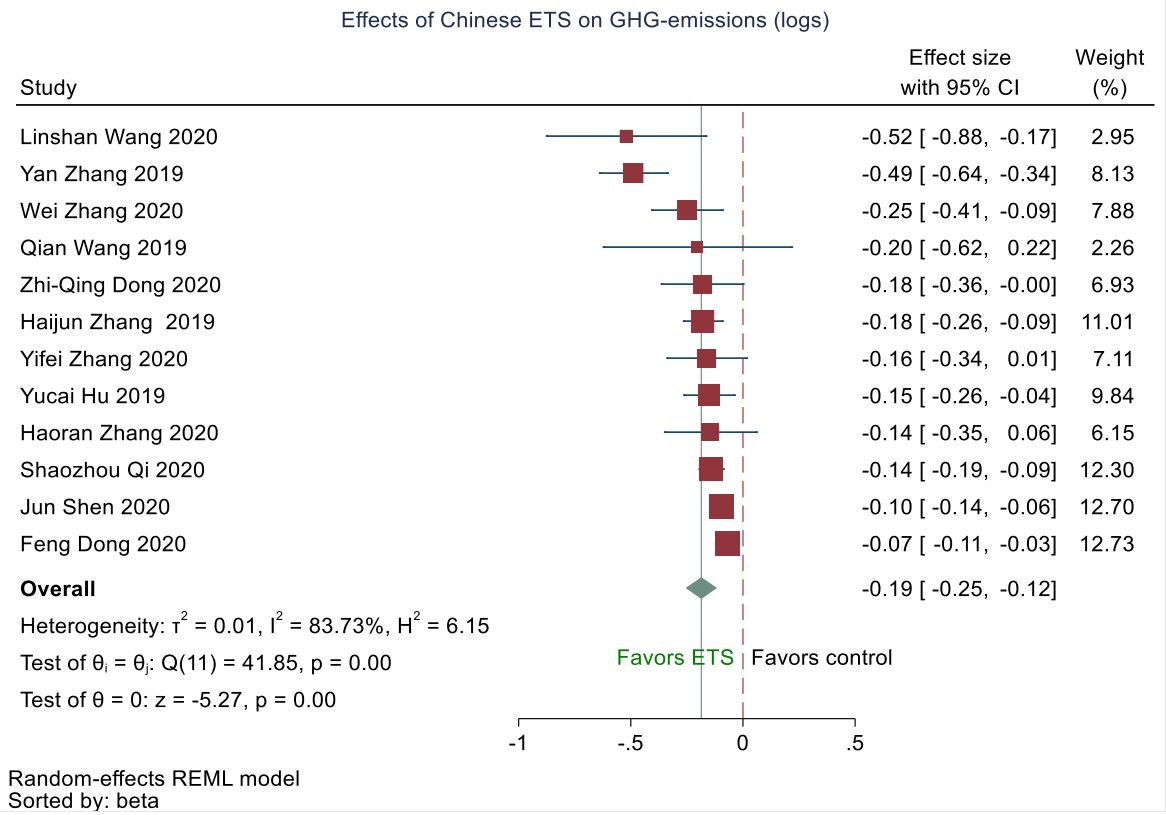
Source: Authors

b. Effects of the Chinese pilot ETS on GHG emissions

As in the section on electrification and employment, we started this analysis by synthesizing studies based on their non-standardized regression coefficients. We selected all models that measure GHG emissions in logarithmic units (logs). This has two advantages. First, most studies do report logarithmic models. Second, our coefficient of interest—a binary ETS dummy, in this case—shows a relative (percentage) change compared to the control group. Given that the included studies use different control groups with consequently different baseline levels of GHG emissions, the interpretation of the overall effect size becomes more meaningful. Furthermore, between-study heterogeneity in effect sizes should thereby also be reduced. As the studies try to estimate the effects of the same pilot ETS, one may argue that a common-effect meta-analysis may be appropriate.

However, the studies are not random samples from the same population. Instead, the data underlying the studies have been chosen by the researchers and the most important variable (GHG emissions) has to be calculated, based on assumptions. Therefore, we used the more cautious random-effects approach, accepting the possibility that the true effect sizes of what the studies actually infer may differ.

Figure IV-34. Effects of the Chinese pilot ETS on logarithmic GHG emissions (forest plot)

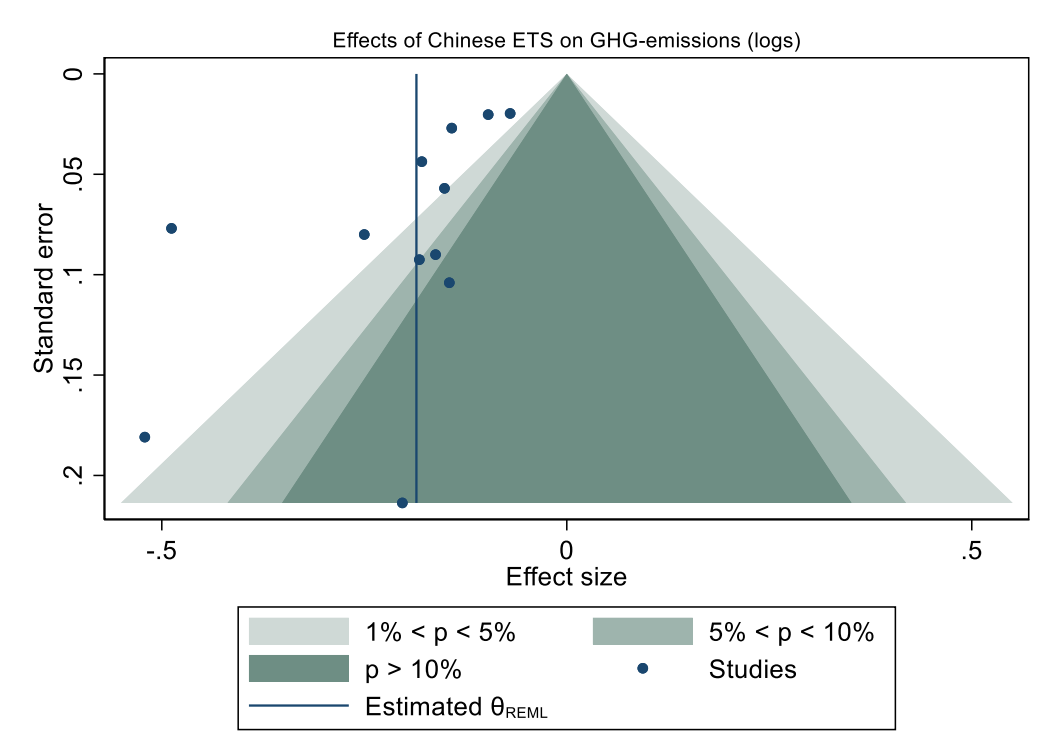


Source: Authors

Overall, the meta-analysis (see Figure IV-34) shows a significant and large overall effect. An overall effect of -0.19 equates to a reduction of around 17 per cent from the level of GHG emissions of the control group. There is substantial heterogeneity in both relative (I-squared) and absolute (T-squared) terms, which vindicates the decision to use a random-effects model. Eighty-four per cent (I-squared) of the variation can be attributed to between-study differences. By taking advantage of the estimated Tau, we can calculate the range in which 95 per cent of the true effect sizes lies. It ranges from a reduction of around 31 per cent to a basically non-existing change (an increase of 0.4 per cent).

Before converting the effect sizes into Cohen's d, we first tested for the presence of a small-study effect. The visual inspection of the funnel plot (Figure IV-35) is startling. It looks like a textbook example of publication bias.

Figure IV-35. Test for "small-study effects": effects of the Chinese pilot ETS on logarithmic GHG emissions (funnel plot)



Source: Authors

Many studies cluster around and outside the boundaries of statistical significance. There is, furthermore, a clear asymmetry, where studies with large standard errors only have above-average effect sizes. A statistical test confirms the visual impression of the asymmetry ($p<0.01$, using the regression-based Egger test).

Of course, a funnel plot cannot reveal the source of the correlation between standard errors and effect size. We first have to ask whether there could be valid theoretic reasons as to why smaller studies may have larger effects, other than publication bias. Since the studies all evaluate the effects of the ETS, possible reasons related to the size of the intervention do not hold water. Instead, we may speculate that some regions or sectors in China may have shown stronger reduction effects than others: this could explain why *some* smaller studies (those focusing on these regions or sectors) may show larger effects than larger and more representative studies based on data from more areas covered by the pilot ETS. However, this explanation does not account for the absence of small effects for studies with large standard errors. For this outcome, there must, furthermore, be a mechanism that makes the data on samples, where reduction effects are larger, more likely to be chosen. While this possible mechanism may not be publication bias directly, it follows the logic that somewhere in the process of publishing results, significant results are favoured—be it by journals or by researchers cherry-picking the data they work with.

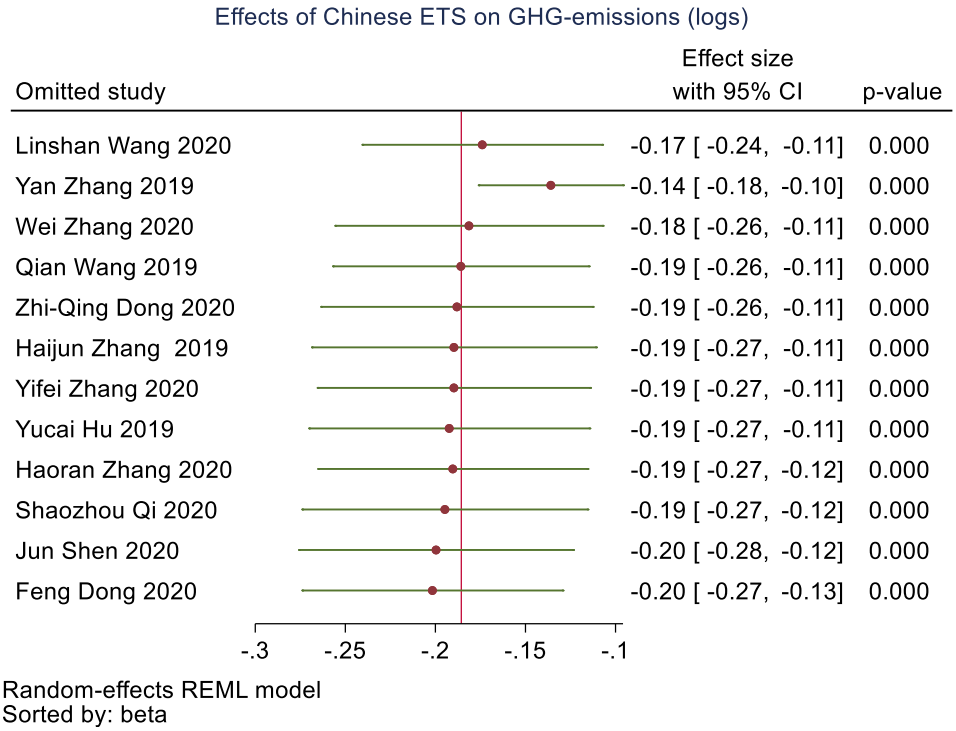
Unfortunately, all studies in this meta-analysis are published articles. We, therefore, have no way of using grey literature as a robustness check. Furthermore, according to the risk-of-bias assessment, no study received a rating of high confidence in its findings. There is no meaningful difference in the estimated effect sizes between studies of medium- and low-confidence. This subgroup analysis is, therefore, not shown here.

We can, however, use the estimated intercept of the regression on which the Egger test rests to make a prediction of what a non-biased effect size might be. Before presenting the results, it needs to be

noted that this prediction is based, as the Egger test itself, on the assumption of homogeneity between studies, on a small sample of 12 studies (making it highly dependent on possible outliers), as well as on a linear interpolation towards studies with infinitely large sample sizes (the intercept of the regression).

Nevertheless, it is insightful to see that the correlation between the effect sizes and their standard errors is so strong that the predicted unbiased effect size would be a reduction of only around 0.077 in log units or of around 7.5 per cent. While this is still a meaningful reduction, it is less than half of what the possibly biased overall effect size estimate suggests. Furthermore, the outlier inspection shows that merely excluding the most influential study (Zhang and Zhang, 2019) could result in a decline of the estimated overall reduction in GHG emissions by four percentage points to around 14 per cent.

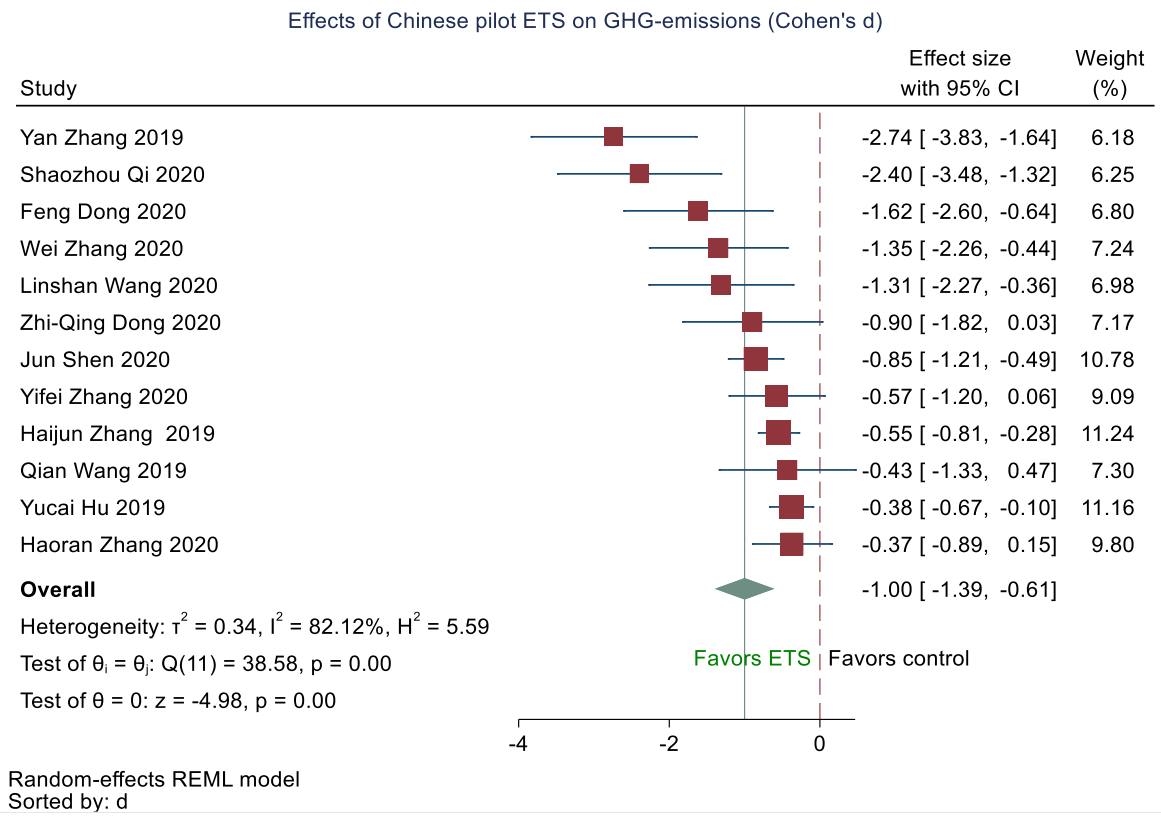
Figure IV-36. Outlier inspection: estimating the overall effect size of electrification while excluding one study at a time



Source: Authors

The likelihood of some form of publication bias introduces a cautionary tone when assessing whether the Chinese pilot ETS may have been transformational. We again converted the effect sizes into Cohen's d and compared the results. As shown by the forest plot (Figure IV-37), the effect sizes, in terms of Cohen's d, are large, with an overall effect of -1. This is, in absolute terms, between the conventional thresholds for large (0.8) and huge (1.2) effects. Therefore, in the absence of publication bias, ETSs, as implemented in China, would be a candidate for transformational change, as defined by this study. However, it is unclear how large the effects truly are without publication bias. It is likely that they are considerably smaller.

Figure IV-37. Effects of the Chinese pilot ETS on GHG emissions, in terms of Cohen's d (forest plot)



Source: Authors

V. CONCLUSIONS AND IMPLICATIONS

a. SUMMARY

Climate change is one of the most pressing global priorities of the 21st century. To achieve the necessary mitigation and adaptation activities, transformational changes—significant and long-lasting effects on a large scale—are needed across systems and individual behaviour. This joint evidence synthesis by GCF/IEU and CIF E&L Initiative, carried out by C4ED, with the assistance of the Africa Centre for Systematic Reviews & Knowledge Translation and advice from the Campbell Collaboration, mapped out the landscape of evidence on the causal drivers of transformational change in two sectors-energy and public health. This evidence synthesis applied a novel approach and combined two different syntheses into one learning exercise on transformational change. The goal was not only to learn for each sector separately, but also to see whether climate interventions can draw lessons from effective interventions in the public health sector.

The report discussed the various steps the authors undertook—from searching and screening studies, extracting and cleaning data, and setting up the EGMs for both sectors to the final stages where the evidence was analysed through a meta-analysis and meta-regressions. At each stage, the synthesis made explicit the decisions that were made in order to allow other research teams to extend, expand, and improve upon the work described therein. Needless to say, many different routes could have been taken at different junctures. The eventual choices carried out were made on the basis of defensible pragmatism: to deliver a high-quality and easily replicable evidence synthesis within the resource and time constraints of the study. This work relied on the expert advice of a range of academics and practitioners to guide the study and make it as relevant as possible.

In public health, a broad scope of papers covering behavioural outcomes within the sectors of health services, hygiene practices, nutrition, physical activity, and substance abuse were reviewed. The lack of papers, fitting the inclusion criteria pertaining to the sectors of physical activity and substance abuse, resulted in the exclusion of these sectors from the meta-analysis stage. The biggest reason for their exclusion could be attributed to the fact that the evidence is lacking in terms of the transformational nature: this is because most studies focus either on small samples or cover a short time frame. For the three remaining sectors, namely, the utilization of health services, hygiene practices, and nutrition, the meta-analysis focused on the results from eight out of the 10 intervention types adapted from Michie and others (2011) behaviour wheel. All eight intervention functions, except for social restructuring (that has the smallest sample size), appear to have small (approaching moderate) effect sizes. Of these eight interventions, three intervention types featured prominently in the treatment designs in our study sample: education, persuasion, and enablement. Of these three, regardless of which subsample was analysed, only enablement consistently approached Cohen's d values of between 0.2 and 1.12, thus passing the thresholds for small and large-effect sizes, using the conventions in the literature. The larger effect sizes for enablement were observed for the consumption/purchasing decision samples, which have smaller sample sizes compared to those in the health-seeking behaviour samples. The test for bias suggests that publication bias does affect the results in these samples, although there is some unexplained source of heterogeneity in our study sample as well. The results observed for the enablement intervention function within the forest plots were corroborated by the meta-regression coefficients. The effect sizes of the regression results, however, hint at smaller effects, not even crossing the threshold of the small effect size of 0.2. These results became statistically insignificant when method and region dummies were included.

The findings from the quantitative synthesis on behavioural change in public health indicate that enablement is potentially the intervention function that can—across the different types of outcomes, concrete interventions, and contexts—produce changes of a moderate size, as judged by the conventional thresholds on Cohen's d. It should, at this stage, be noted that Cohen's d is only one way to evaluate the magnitudes of effect sizes; we revisit this point in the following subsection on the limitations of our review.

The findings of the quantitative synthesis on energy offer some interesting, yet ultimately, anticlimactic results. First, with regard to the effects of electrification on formal employment, the review offers estimates of a percentage point increase of two per cent, driven mainly by the greater employment of women. Such a change is precisely what economic theory and the development trajectories of countries show. As women are released from repetitive domestic responsibilities due to the use of electric domestic appliances and mechanical solutions, they have greater time and scope to enter the labour market. In our review, however, the effect of the two per cent needs to be revised downward, when considering heterogeneity and publication biases, leading to a Cohen's d figure of 0.03—considerably below a small effect size of 0.2 and far removed from the transformational change threshold of 0.8.

Does this mean that the employment effects are irrelevant? Not at all. Two areas merit attention. First, these employment effects are employment co-benefits from electrification. They are not the priority outcome areas targeted by the interventions, but additional benefits. Second, we need to consider the context: social norms, in terms of women's employment, will differ across the countries considered here, as participation will be more circumscribed in some countries, especially in more traditional rural areas. In this respect, further work could consider the baseline levels of comparison groups to assess whether the 1.5-percentage point increase in women's formal employment is derived from a low base in certain countries.

The second cell that was ripe for a meta-analysis also offers findings that are indicative of potential. The finding that the Chinese pilot ETS offers causal evidence of transformational change, as defined in this paper, is substantial. A reduction of emissions of 17 per cent is considerable and a Cohen's d figure of -1 would certainly imply transformational change. However, the possible influence of publication bias casts doubt on the true effect size. The studies forming this cell belong to very recently published literature, and we are hopeful to see more variability in evidence from published and grey literature in the years ahead. Nevertheless, that an intervention from a market-based mechanism offers the promise of transformational change remains a relevant finding. It resurrects the promise of a market-based ETS and foregrounds the two flexibility mechanisms with the Kyoto Protocol designed to reduce the costs of compliance. In addition to the Joint Implementation channel for Annex I Parties, Kyoto led to the creation and operation of the Clean Development Mechanism (CDM), where Certified Emission Reduction credits from non-Annex I Parties could be traded with offset credits sold to Annex I Parties. That the pilot ETS in China is leading to substantial emission reductions may increase the wider adoption of such market-based mechanisms.[24]

B. QUALITY OF THE EVIDENCE AND LIMITATIONS OF THE REVIEW PROCESS

Based on the risk-of-bias assessment, carried out as part of the evidence synthesis, the public health sector, also a long-standing sector in terms of causal evidence, mostly comprises high-quality evidence, with very few studies of low-quality evidence. In comparison, the energy sector mostly

[24] For a general overview of the design and implementation of the Chinese Pilot ETS, see Duan and others (2014); Partnership for Market Readiness (2014); Zhao and others (2019); and Zhou and others (2020).

includes evidence of medium quality. Tests for small-study effects show that both sectors may suffer from the problem of publication bias, although this occurrence is not as prevalent in the public health sector. Our own study has several potential limitations.

First, we excluded studies published earlier than 1990 (energy sector) and 2000 (public health sector).

Second, the WB country ranking by income status for 2020, used to group countries in this report, may not take into consideration the transitional nature of the previous income status when the literature was published or the data collected. Furthermore, the focus on developing countries means missing out on evaluating potentially transformative interventions, such as the Clean Air Acts in the United Kingdom and the United States. This is, nevertheless, not a very large limitation, given that the focus of the work of CIF and GCF lies in developing countries, where the evidence was comprehensively harvested.

Third, there are limitations by language: excluding the study literature in French, Spanish, German, and Mandarin may have limited the generalizability of our findings, particularly in francophone Africa. Furthermore, the likely publication bias found in the meta-analysis of the ETS in China may be exacerbated by the absence of working papers or reports that may exist in Mandarin, but they were not included in our review because of the language restriction.

Fourth, some studies included in this evidence synthesis have missing statistical information, for example, intervention and control samples. However, such data is required to aid the meta-analysis. The review team obtained some of the missing information by conducting searches for online annexes published in different databases. Furthermore, as described previously, we imputed missing data whenever feasible. Nevertheless, some studies had to be dropped from the meta-analysis for this reason. The quality of the reporting, however, is often highly correlated with the quality of the paper, implying that this limitation might even have benefited the quality of the results synthesized; as such, this is not really a large limitation.

Fifth, we did not consider attrition for the studies in the energy sector, since a substantial number of them are ex-post studies.

Sixth, in order to judge whether an effect is big, Cohen's d only provides one perspective and one that is necessarily limited. It measures an effect size, in terms of the variability of the outcome. Its advantage—and hence the reason that we adopted it for our review—is its capacity to make effect sizes comparable across a wide range of different studies (as long as the necessary statistical information is provided).

However, Cohen's d certainly has its weaknesses. If the variability is large, a small Cohen's d can still represent a large change in absolute terms. Furthermore, despite the existence of conversion formulas in widely used textbooks (Borenstein and others, 2011), the conversions of binary data into Cohen's d faces conceptual challenges, as there is no additional information available on the variability of the outcome not already reflected in the mean. Cohen's d is, in this case, not as reliable for continuous data.

A limitation of the quantitative synthesis in the health sector is that studies differ in their effect-size metrics to such a degree that we largely had to rely on Cohen's alone to evaluate the magnitude of effects. Evaluating whether the magnitude of the effect is large in absolute terms can be very context-specific, and hinges, for instance, on the effectiveness of alternative interventions, the importance of the outcome, and the targets set by policymakers. As Howard White argues,

answering this question, therefore, requires detailed sectoral knowledge, which is beyond the scope of this broad review.[25]

Seventh, the results in public health show substantial heterogeneity, both in relative and absolute terms. This raises the question of how comparable the studies we had pooled together, based on intervention functions, are. Part of the heterogeneity may also have stemmed from the necessary conversions into a common metric—Cohen's d—in case outcomes differ substantially in their variability.

Lastly, our evidence synthesis is subject to limitations inherent in our focus on searching for causal evidence on the drivers of transformational change. Transformational change can be advanced in multiple ways, including through incremental changes that eventually combine with other factors to reach a tipping point as well as through more rapid or sudden shifts. It is, thus, not always clear to what extent an intervention has been, or is currently, on a pathway to being transformational. Relatedly, experimental and quasi-experimental methods, seeking to explain causality in such complex intervention contexts and multiple outcome areas, are not always well-suited to capturing such changes in the breadth, depth, and level of nuance required. Transformation is dynamic and non-linear, often requiring sequential, multi-stage, or parallel interventions that causal experimental studies do not capture. Further, despite the availability of high-quality evaluations, finding evidence of transformational change—as defined here—puts additional requirements on the data, such as the scope and time frame that only a limited number of studies satisfy. Therefore, as in any systematic review, the gaps in evidence this report have revealed need to be seen in light of its inclusion criteria. This is also visible in the limitation related to the timing of the intervention—also a criterion in our definition. The insufficient number of studies, with the timing of the data collection going beyond a few years from the full roll-out of the intervention, imposes a further constraint on the extent and direction of the observed long-term effects in our sample of studies.

C. AREAS FOR FURTHER INVESTIGATION

Section B on limitations already provides several avenues where the work in this report can be extended. We now highlight a few additional areas whereby further investigation could be most fruitful.

First, as has been described in the limitations section, we set a restricted definition of "transformational change", with the goal of being able to quantitatively synthesize evidence on its causal drivers. Possibly, a narrative synthesis, based on other types of evidence than just experimental and quasi-experimental studies, could reveal lessons on how to achieve transformational change, as has been done through several studies by CIF and GCF before. By broadening the scope, it is possible to find more insights on the climate change adaptation of energy systems and how legislative interventions may produce transformational change. Such interventions are usually implemented over large administrative areas (such as nation-states), which precludes studying their impacts through quantitative methods. The Chinese ETS is a welcome exception of a large-scale policy experiment that could be studied through quasi-experimental methods. Further work beyond impact evaluations may also require searching in different databases, such as in the fields of engineering and social sciences.

Second, given that enablement is the best candidate for producing transformational change, it would be worthwhile to focus specifically on this intervention function, while at the same time, broadening the study of outcomes to other behaviours. This may lead to a larger mass of studies on enablement

[25] See https://www.3ieimpact.org/blogs/how-big-big-need-sector-knowledge-judging-effect-sizes-and-performing-power-calculations (accessed July 09, 2021).

interventions, which may allow for a better understanding of the types of conditions in which enablement works best.

Third, and relatedly, it would be important to learn whether enablement can be equally effective in sectors relevant to climate change adaptation and mitigation. As discussed in this concluding chapter, the possible enabling effects of electrification and public transport systems on mitigation behaviour were not studied as part of this synthesis but would certainly be highly relevant.

D. CROSS-SECTORAL LEARNINGS AND AUTHORS' CONCLUSIONS

We conclude this evidence synthesis by looking at the potential lessons to be learned across sectors, and in particular from the health sector, for climate change adaptation and mitigation behaviour.

In the energy sector, of the two combinations of interventions and outcomes where we found a mass of studies that passed our inclusion criteria, the ETS in China shows the greatest promise for transformational change, though with the caveat of a strong risk of publication bias in the overall results. The intervention is categorized as a structural intervention that does not target behaviour change at the household level. It, therefore, does not directly coincide with any of the categories in the public health sector interventions, as defined by the intervention functions of the behaviour change wheel.

The second intervention-outcome combination in the energy sector—the effects of electrification on employment—was also coded as structural, without directly targeting behavioural change with regards to employment. Nevertheless, we may link it to two closely related intervention functions from the health sector: enablement and physical restructuring. The effects of electrification are significant, albeit small. However, the impact of access to electricity on the household may be much broader and accumulate through its continuing and nearly irreversible nature. Therefore, the actual longer-term behavioural change effects of enablement and physical restructuring may be difficult to isolate from other intervention functions coming into play indirectly over time (through greater access to knowledge and media, for example).[26]

In the health sector, none of the intervention functions we looked at shows large effects on health behaviour over time, as judged by the conventional threshold of Cohen's d. Nevertheless, the intervention function, enablement, holds forth the greatest potential for behavioural change across all priority areas, in particular, for outcomes with social benefits. We, nonetheless, suggest that the reader accept this evaluative judgment cautiously, given the risk of publication bias as well as the limitations presented earlier. This finding, nevertheless, hints at the types of interventions that could work for mitigation or adaptation interventions. Recall that mitigation interventions provide a global public good, while adaptation interventions provide a range of public goods, toll goods, private goods, or the better management of common-pool resources (across sectors and at different scales). So, in summary, mitigation interventions provide social benefits, while adaptation interventions provide either private or social benefits. Following this logic, enablement interventions could be particularly well-suited for mitigation interventions, as the effects are the largest for outcomes with largely social benefits. Enablement also works, albeit overall with a smaller effect size, for outcomes with private benefits. In general, the findings suggest that enablement holds particular potential for mitigation interventions, and to a lesser extent, for adaptation.

Given that mitigation is a global public good, a key concern is whether individuals are sufficiently motivated to engage in mitigation when they cannot profit from their behaviour individually. Internalizing the positive externalities of mitigation behaviour is either costly or subject to strong

[26] This may suggest that interventions should be categorized by continuity and inherent irreversibility in future searches for interventions with transformational change potential.

public resistance (for example, against carbon taxes). Therefore, it would be good news, if merely enabling individuals proves to be effective also for mitigation behaviour. In the energy sector, electrification may enable households to reduce their consumption of firewood or fossil fuels. This possible effect of electrification has not been synthesized as part of this report. Other examples of interventions that can enable mitigation behaviour include public transport systems and electric cars if they are convenient and low-cost. In addition, the opportunity to work from home, instead of commuting or travelling to business meetings, may be a positive long-term side effect of the COVID-19 pandemic.

This report emphasizes how difficult it is to find and synthesize evidence on the causal drivers of transformational change. Searching for causality restricts the available evidence solely to those coming from quantitative experimental and quasi-experimental studies. The nature of these methods means that such studies usually focus on singular interventions on outcomes, at the individual or household level, within a relatively short time frame. Therefore, long-term changes over decades, rather than years, occurring at the system level, as a result of a complex interplay between many incremental changes from various interventions, are not captured. Ergo, as a possible lesson from our review, transformational change may be difficult to achieve with interventions that are studied through experimental and quasi-experimental study designs. As this does not imply that transformational change is elusive, the search for causal drivers may need to be complemented by a narrative synthesis of case studies and other mixed-method approaches.

REFERENCES

References to studies included in public health sector EGM (144 studies)

Abdu, Zeidan, and others (2004). The impact of user fee exemption on service utilization and treatment seeking behaviour: The case of malaria in Sudan. *International Journal of Health Planning and Management*, vol. 19, issue 1, S95.

Abebaw, Degnet, Yibeltal Fentie and Belay Kassa (2010). The impact of a food security program on household food consumption in Northwestern Ethiopia: A matching estimator approach. *Food Policy*, vol. 35, issue 4, pp. 286–293.

Adams, Katherine P, and others (2018). Unintended effects of a targeted maternal and child nutrition intervention on household expenditures, labor income, and the nutritional status of non-targeted siblings in Ghana. *World Development*, vol. 107, pp. 138–150.

Ahmed, Syed Masud, and others (2006). Targeted intervention for the ultra poor in rural Bangladesh: Does it make any difference in their health-seeking behaviour? *Social Science & Medicine*, vol. 63, issue 11, pp. 2899–2911.

Akanbi, Maxwell Oluwole, and others (2019). The efficacy of smoking cessation interventions in low- and middle-income countries: A systematic review and meta-analysis. *Addiction*, vol. 114, pp. 620–635.

Akresh Richard, Damien De Walque, and Harounan Kazianga (2012). Alternative cash transfer delivery mechanisms: Impacts on routine preventative health clinic visits in Burkina Faso. National Bureau of Economic Research, Inc, NBER Working Paper No. 17785.

Althabe Fernando, and others (2019). A multifaceted intervention to improve syphilis screening and treatment in pregnant women in Kinshasa, Democratic Republic of the Congo and in Lusaka, Zambia: A cluster randomised controlled trial. *The Lancet Global Health*, vol. 7, e655–e663.

Andrade, Monica V, and others (2012). Income transfer policies and the impacts on the immunization of children: The Bolsa Familia Program. *Cadernos de Saude Publica*, vol. 28, pp. 1347–1358.

Andres, Luis Alberto, and others (2020). A Multiple-Arm, Cluster-Randomized Impact Evaluation of the Clean India (Swachh Bharat) Mission Program in Rural Punjab, India. Policy research working paper No. 9249. Washington D.C.: World Bank.

Arnold, Benjamin, and others (2009). Evaluation of a pre-existing, 3-year household water treatment and handwashing intervention in rural Guatemala. *International Journal of Epidemiology*, vol. 38, pp. 1651–1661.

Atnafu, Asfaw, and others (2015). Mobile health (mHealth) intervention in maternal and child health care: Evidence from resource-constrained settings - A review. *Ethiopian Journal of Health Development*, vol. 29, issue 3, pp. 140–153.

Aziz, Zahra, and others (2018). A group-based lifestyle intervention for diabetes prevention in low- and middle-income country: Implementation evaluation of the Kerala Diabetes Prevention Program. *Implementation Science*, vol. 13, article No. 97. Available at https://doi.org/10.1186/s13012-018-0791-0.

Banerjee, Abhijit Vinayak, and others (2010). Improving immunisation coverage in rural India: clustered randomised cControlled evaluation of immunisation campaigns with and without incentives. *BMJ*, vol. 340.

Biran, Adam, and others (2014). Effect of a behaviour-change intervention on handwashing with soap in India (SuperAmma): A cluster-randomised trial. *The Lancet Global Health*, vol. 2, E145–E154.

Boisson, Sophie, and others (2010). Field assessment of a novel household-based water filtration device: A randomised, placebo-controlled trial in the Democratic Republic of Congo. *PLoS ONE [Electronic Resource]*, vol. 5, E12613.

Boone, Peter, and others (2017). Community health promotion and medical provision for neonatal health—CHAMPION cluster randomised trial in Nagarkurnool district, Telangana (formerly Andhra Pradesh), India. *PLoS Medicine / Public Library of Science*, vol. 14, E1002324.

Boruett, Patrick, and others (2013). Facility-level intervention to improve attendance and adherence among pPatients on anti-retroviral treatment in Kenya—A quasi-experimental study using time series analysis. *BMC Health Services Research*, vol. 13. Available at https://doi.org/10.1186/1472-6963-13-242.

Boudreaux, Chantelle, Phetdara Chanthala, and Magnus Lindelow (2014). Assessing the elimination of user fees for delivery services in Laos. *Public Library of Science ONE*, vol 9, issue 3, pp. 1–8. Available at https://doi.org/10.1371/journal.pone.0089784.

Briceño, Bertha, Aidan Coville, and Sebastian Martinez (2015). Promoting handwashing and sanitation: evidence from a large-scale randomized trial in rural Tanzania. Policy research working paper No. 7164. Washington D.C.: World Bank.

Briceño, Bertha, and others (2017). Are there synergies from combining hygiene and sanitation promotion campaigns: Evidence from a large-scale cluster-randomized trial in rural Tanzania. *PLoS ONE*, vol. 12.

Brown, Joe, Sorya Proum, and Sobsey Mark D. (2009). Sustained use of a household-scale water filtration device in rural Cambodia. *Journal of Water & Health*, vol. 7, pp. 404–412.

Brugh, Kristen, and others (2018). Impacts of the Malawi social cash transfer program on household food and nutrition security. *Food Policy*, vol. 76, pp. 19–32.

Brunie, Aurelie, and others (2014). Can village savings and loan groups be a potential tool in the malnutrition fight? Mixed method findings from Mozambique. *Children and Youth Services Review*, vol. 47, part 2, pp. 113–120 (December). Available at https://doi.org/10.1016/j.childyouth.2014.07.010.

Busza, Joanna, and others (2019). Enhancing national prevention and treatment services for sex workers in Zimbabwe: a process evaluation of the SAPPH-IRe trial. *Health Policy and Planning*, vol. 34, issue 5, pp. 337–345.

Cahyadi, Nur, and others (2018). Cumulative Impacts of Conditional Cash Transfer Programs: Experimental Evidence from Indonesia. National Bureau of Economic Research, Inc, NBER working paper No. 24670. Available at https://www.nber.org/papers/w24670.

Cameron, Lisa, Susan Olivia and Manisha Shah (2013). Impact evaluation of a large-scale rural sanitation project in Indonesia. Policy research working paper No. 6360. Washington D.C.: World Bank. Available at http://documents.worldbank.org/curated/en/783781468044120966/Impact-evaluation-of-a-large-scale-rural-sanitation-project-in-Indonesia.

Carvalho, Natalie, and others (2014). More evidence on the impact of India's conditional cash transfer program, Janani Suraksha Yojana: quasi-experimental evaluation of the effects on childhood immunization and other reproductive and child health outcomes. *PLoS ONE [Electronic Resource]*, vol. 9, issue 10, e109311.

Chang, Annie, and others (2017). The effect of BRT implementation and streetscape redesign on physical activity: A case study of Mexico City. *Transportation Research: Part A: Policy and Practice*, vol. 100, pp. 337–347.

Chankova, Slavea, Laurel Hatt and Sabine Musange (2012). A community-based approach to promote household water treatment in Rwanda. *Journal of Water and Health*, vol. 10, issue 1, pp. 116–129.

Chao, Jianqian, and others (2012). The effect of community-based health management on the health of the elderly: a randomized controlled trial from China. *BMC Health Services Research*, vol. 12, issue 449. Available at https://bmchealthservres.biomedcentral.com/articles/10.1186/1472-6963-12-449.

Choulagai Bishnu P, and others (2017). A cluster-randomized evaluation of an intervention to increase skilled birth attendant utilization in mid- and far-western Nepal. *Health Policy and Planning*, vol. 32, issue 8, pp. 1092–1101.

Clasen, Thomas F., and others (2010). Interventions to improve disposal of human excreta for preventing diarrhoea. *Cochrane Database of Systematic Reviews*.

Clasen, Thomas, and others (2014). Effectiveness of a rural sanitation programme on diarrhoea, soil-transmitted helminth infection, and child malnutrition in Odisha, India: a cluster-randomised trial. *The Lancet Global Health*, vol. 2, E645–E653.

Coates, Thomas J., and others (2014). Effect of community-based voluntary counselling and testing on HIV incidence and social and behavioural outcomes (NIMH Project Accept; HPTN 043): a cluster-randomised trial. *Lancet Global Health*, vol. 2, issue 5, e267–e277.

Corbett, Elizabeth L., and others (2006). Uptake of workplace HIV counselling and testing: a cluster-randomised trial in Zimbabwe. *PLoS Medicine / Public Library of Science*, vol. 3, issue 7, e238.

Corbett, Elizabeth L., and others (2007). HIV incidence during a cluster-randomized trial of two strategies providing voluntary counselling and testing at the workplace, Zimbabwe. *AIDS*, vol. 21, issue 4, pp. 483–489.

Crocker, Jonny, and others (2016). Impact evaluation of training natural leaders during a community-led total sanitation intervention: a cluster-randomized field trial in Ghana. *Environmental Science & Technology*, vol. 50, pp. 8867–8875.

Crocker, Jonny, and others (2016). Teachers and sanitation promotion: an assessment of community-led total sanitation in Ethiopia. *Environmental Science & Technology*, vol. 50, issue 12, pp. 6517–6525.

Daniels, Joseph, and others (2014). Effects of a community intervention on HIV prevention behaviours among men who experienced childhood sexual or physical abuse in four African settings: Findings from NIMH Project Accept (HPTN 043). *PLoS ONE*, vol. 9, issue 6, e99643.

Doocy, Shannon, and others (2018). Improving Hhousehold food security in Eastern Democratic Republic of the Congo: a comparative analysis of four interventions. *Food Security*, vol. 10, issue 3, pp. 649–660.

Dzakpasu, Susie, Timothy Powell-Jackson and Oona M R.Campbell (2014). Impact of user fees on maternal health service utilization and related health outcomes: a systematic review. *Health Policy and Planning*, vol. 29, issue 2, pp. 137–150.

Edmond, Karen M., and others (2019). Conditional Cash Transfers to Improve Use of Health Facilities by Mothers and Newborns in Conflict Affected Countries, a Prospective Population Based Intervention Study from Afghanistan. *BMC Pregnancy & Childbirth*, vol. 19, issue 1, pp. 193–211.

Ercumen, Ayse, and others (2015). Effects of source versus household contamination of tubewell water on child diarrhea in rural Bangladesh: a randomized controlled trial. *PLoS ONE [Electronic Resource]*, vol. 10, issue 3, E0121907.

Ercumen, Ayse, and others (2018). Effects of single and combined water, sanitation and handwashing interventions on fecal contamination in the domestic environment: a cluster-randomized controlled trial in rural Bangladesh. *Environmental Science & Technology*, vol. 52, issue 21, pp. 12078–12088.

Ezeanolue Echezona E., and others (2015). Effect of a congregation-based intervention on uptake of HIV testing and linkage to care in pregnant women in Nigeria (Baby shower): a cluster randomised trial. *Lancet Global Health*, vol. 3, issue 11, E692–E700.

Ezeanolue, Echezona E., and others (2017). What do you need to get male partners of pregnant women tested for HIV in resource limited settings? The baby shower cluster randomized trial. *AIDS & Behaviour*, vol. 21, issue 2, pp. 587–596.

Fabbri, Camilla, and others (2019). The effect of report cards on the coverage of maternal and neonatal health care: a factorial, cluster-randomised controlled trial in Uttar Pradesh, India. *Lancet Global Health*, vol. 7, issue 8, e1097–e1108.

Filmer, Deon P., and others (2018). General equilibrium effects of targeted cash transfers: Nutrition impacts on non-beneficiary children. Policy research working paper series No. 8377. Washington D.C.: World Bank.

Fonner, Virginia A., and others (2012). Voluntary counselling and testing (VCT) for changing HIV-related risk behaviour in developing countries. *Cochrane Database of Systematic Reviews,* vol. 9, issue 9, CD001224.

Freeman, Matthew C., and others (2016). The impact of a rural sanitation programme on safe disposal of child faeces: a cluster randomised trial in Odisha, India. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, vol. 110, issue 7, pp. 386–392.

Galiani, Sebastian, Paul Gertler and Alexandra Orsola-Vidal (2012). Promoting handwashing behaviour in Peru: the effect of large-scale mass-media and community level interventions. Policy research working paper No. 6257. Washington D.C.: World Bank.

Garn, Joshua V., and others (2017). The impact of sanitation interventions on latrine coverage and latrine use: a systematic review and meta-analysis. *International Journal of Hygiene and Environmental Health*, vol. 220, issue 2, pp. 329–340.

Geldsetzer, Pascal, and others (2019). Community health workers to improve uptake of maternal healthcare services: a cluster-randomized pragmatic trial in Dar es Salaam, Tanzania. *PLoS Medicine*, vol. 16, issue 3, e1002768.

Gibson, Dustin G., and others (2017). Mobile phone-delivered reminders and incentives to improve childhood immunisation coverage and timeliness in Kenya (M-SIMU): a cluster randomised controlled trial. *Lancet Global Health*, vol. 5, issue 4, E428–E438.

Guiteras, Raymond, James Levinsohn and Ahmed Mushfiq Mobarak (2015). Encouraging sanitation investment in the developing world: a cluster-randomized trial. *Science*, vol. 348, issue 6237, pp. 903–906.

Gullo, Sara, and others (2017). Effects of a social accountability approach, CARE's Community Score Card, on reproductive health-related outcomes in Malawi: a cluster-randomized controlled evaluation. *PLOS ONE*, vol. 12, issue 2, e0171316.

Harati, Hadi, and othes (2010). Reduction in incidence of type 2 diabetes by lifestyle intervention in a Middle Eastern Community. *American Journal of Preventive Medicine*, vol. 38, issue 6, pp. 628–636:e1.

Harter, Miriam, Nadja Contzen, and Jennifer Inauen (2019). The role of social identification for achieving an open-defecation free environment: a cluster-randomized, controlled trial of community-led Total Sanitation in Ghana. *Journal of Environmental Psychology*, vol. 66, 101360.

Hemminki, Elina, and others (2013). Impact of financial and educational interventions on maternity care: Results of cluster randomized trials in rural China, CHIMACA. *Maternal and Child Health Journal*, vol. 17, issue 2, pp. 208–221.

Higgs, Elizabeth S., and others (2014). Understanding the role of mHealth and other media interventions for behaviour change to enhance child survival and development in low- and middle-income countries: an evidence review. *Journal of Health Communication*, vol. 19, issue sup1, pp. 164–189.

Hoddinott, John and Emmanuel Skoufias (2004). The impact of PROGRESA on food consumption. *Economic Development and Cultural Change*, vol. 53, issue 1, pp. 37–61.

Hotz, Christine, and others (2012). A large-scale intervention to introduce orange sweet potato in rural Mozambique increases vitamin A intakes among children and women. *British Journal of Nutrition*, vol. 108, issue 1, pp. 163–176.

Huda, Tarique Md Nurul, and others (2012). Interim evaluation of a large scale sanitation, hygiene and water improvement programme on childhood diarrhea and respiratory disease in rural Bangladesh. *Social Science & Medicine*, vol. 75, issue 4, pp. 604–611.

Hunter, Benjamin M., and others (2017). The effects of cash transfers and vouchers on the use and quality of maternity care services: a systematic review. *PLOS ONE*, vol. 12, issue 3.

Jafar, Tazeen H., and others (2009). Community-based interventions to promote blood pressure control in a developing country. *Annals of Internal Medicine*, vol. 151, issue 9, pp. 1–11.

Jemmott III, John B., and others (2015). Mediation of effects of a theory-based behavioural intervention on self-reported physical activity in South African men. *Preventive Medicine*, vol. 72, pp. 1–7.

Jürgensen Marte, and others (2013). Effects of Home-based Voluntary Counselling and Testing on HIV-related Stigma: Findings from a Cluster-randomized Trial in Zambia. *Social Science & Medicine*, vol. 81, pp. 18–25. Available at https://doi.org/10.1016/j.socscimed.2013.01.011.

Khetan, Aditya, and others (2019). Effect of a community health worker-based approach to integrated cardiovascular risk factor control in India: a cluster randomized controlled trial. *Science Direct*, vol. 14, issue 4, pp. 355–365. Available at https://doi.org/10.1016/j.gheart.2019.08.003.

Kim Sunny S., and others (2018). Large-scale social and behaviour change communication interventions have sustained impacts on infant and young child feeding knowledge and practices: Results of a 2-year follow-up study in Bangladesh. *Journal of Nutrition*, vol. 148, issue 10, pp. 1605–1614.

Kirby, Miles A., and others (2019). Effects of a large-scale distribution of water filters and natural draft rocket-style cookstoves on diarrhoea and acute respiratory infection: a cluster-randomized controlled trial in Western Province, Rwanda. *PLoS Medicine*, vol. 16, issue 6, e1002812.

Kirkwood, Betty R., and others (2013). Effect of the Newhints home-visits intervention on neonatal mortality rate and care practices in Ghana: a cluster randomised controlled trial. *Lancet*, vol. 381, issue 9884, pp. 2184–2192.

Kusuma, Dian, and others (2016). Can cash transfers improve determinants of maternal mortality? Evidence from the household and community programs in Indonesia. *Social Science & Medicine*, vol. 163, pp. 10–20.

Kusuma, Dian, and others (2017). The impact of household and community cash transfers on children's food consumption in Indonesia. *Preventive Medicine*, vol. 100, pp. 152–158.

Lagarde, Mylene and Natasha Palmer (2011). The impact of user fees on access to health services in low- and middle-income countries. *Cochrane Database of Systematic Reviews*. Available at https://doi.org/10.1002/14651858.CD009094.

Lagarde, Mylene, and Natasha Palmer (2008). The impact of user fees on health service utilization in low- and middle-income countries: How strong is the evidence? *Bulletin of the World Health Organization*, vol. 86, issue 11, pp. 839–848. Available at https://doi.org/10.2471/BLT.07.049197.

Lagarde, Mylene, Andy Haines, and Natasha Palmer (2007). Conditional cash transfers for improving uptake of health interventions in low- and middle-income countries: a systematic review. *JAMA, Journal of the American Medical Association*, vol. 298, issue 16, pp. 1900–1910.

Lagarde, Mylene, Andy Haines, and Natasha Palmer (2009). The impact of conditional cash transfers on health outcomes and use of health services in low- and middle-income countries. *Cochrane Database of Systematic Reviews*. Available at https://doi.org/10.1002/14651858.CD008137.

Lamichhane, Prabhat, Anurag Sharma, and Ajay Mahal (2017). Impact evaluation of free delivery care on maternal health service utilisation and neonatal health in Nepal. *Health Policy and Planning*, vol. 32, issue 10, pp. 1427–1436.

Lee, Hocheol, and others (2019). Community-based maternal and child health project on 4+ antenatal care in the Democratic Republic of Congo: a difference-in-differences analysis. *Reproductive Health*, vol. 16, issue 157.

Leroy, Jef L., and others (2010). Cash and in-kind transfers in poor rural communities in Mexico increase household fruit, vegetable, and micronutrient consumption but also lead to excess energy consumption. *Journal of Nutrition*, vol. 140, issue 3, pp. 612–617.

Lindegren, Mary Lou, and others (2012). Integration of HIV/AIDS services with maternal, neonatal and child health, nutrition, and family planning services. *Cochrane Database of Systematic Reviews*, vol. 9.

Lippman, Sheri A., and others (2017). Community mobilization for HIV testing uptake: Results from a community randomized trial of a theory-based intervention in rural South Africa. *JAIDS—Journal of Acquired Immune Deficiency Syndromes*, vol. 74, issue sup 1, S44–S51.

Liu, Xiaoning, Hong Yan, and Duolao Wang (2010). The evaluation of "Safe Motherhood" program on maternal care utilization in rural Western China: a difference in difference approach. *BMC Public Health*, vol. 10, issue 556.

Luby, Stephen P., and others (2009). Difficulties in maintaining improved handwashing behaviour, Karachi, Pakistan. *American Journal of Tropical Medicine & Hygiene*, vol. 81, issue 1, pp. 140–145.

Luby, Stephen P., and others (2018). Effects of water quality, sanitation, handwashing, and nutritional interventions on diarrhoea and child growth in rural Bangladesh: a cluster randomised controlled trial. *The Lancet Global Health*, vol. 6, issue 3, E302–E315.

Lv, Jun, and others (2014). A community-based multilevel intervention for smoking, physical activity and diet: short-term findings from the community interventions for health programme in Hangzhou, China. *Journal of Epidemiology & Community Health*, vol. 68, issue 4, pp. 333–339.

Lyatuu, Goodluck Willey, and others (2018). Engaging community leaders to improve male partner participation in the prevention of mother-to-child transmission of HIV in Dar es Salaam, Tanzania. *PLoS ONE [Electronic Resource]*, vol. 13, E0207986.

Makhema, Joseph, and others (2019). Universal testing, expanded treatment, and incidence of HIV infection in Botswana. *New England Journal of Medicine*, vol. 381, issue 3, pp. 230–242.

Marquis, Grace S., and others (2018). An agriculture-nutrition intervention improved children's diet and growth in a randomized trial in Ghana. *Maternal & Child Nutrition*, vol. 14, issues sup 3, E12677.

Matovu, Joseph K.B., and others (2016). Evaluation of a demand-creation intervention for couples' HIV testing services among married or cohabiting individuals in Rakai, Uganda: a cluster-randomized intervention trial. *BMC Infectious Diseases*, vol. 16, issue 379.

Mayén, Ana-Lucia, and others (2016). Interventions Promoting Healthy Eating as a Tool for Reducing Social Inequalities in Diet in Low- and Middle-income Countries: A Systematic Review. *International Journal for Equity in Health*, vol. 15, issue 1.

McNairy, Margaret L., and others (2017). Effectiveness of a Combination Strategy for Linkage and Retention in Adult HIV Care in Swaziland: The Link4Health Cluster Randomized Trial. *PLOS MEDICINE*, vol. 14, issue 11.

Memon, Zahid A., and others (2015). Impact of a community-based perinatal and new-born preventive care package on perinatal and neonatal mortality in a remote mountainous district in Northern Pakistan. *BMC Pregnancy & Childbirth*, vol. 15, pp. 106–106.

Midhet, Farid and Stan Becker (2010). Impact of community-based interventions on maternal and neonatal health indicators: Results from a community randomized trial in rural Balochistan, Pakistan. *Reproductive Health*, vol. 7, issue 30.

Miller Candace M., Maxton Tsoka and Kathryn Reichert (2011). The impact of the social cash transfer scheme on food security in Malawi. *Food Policy*, vol. 36, issue 2, pp. 230–238.

Mills, Edward J., and others (2014). Interventions to promote adherence to antiretroviral therapy in Africa: a network meta-analysis. *Lancet HIV*, vol. 1, issue 3, E104.

Mohammadifard, Noushin, and others (2009). Effect of a community-based intervention on nutritional behaviour in a developing country setting: the Isfahan Healthy Heart Programme. *Public Health Nutrition*, vol. 12, issue 9, pp. 1422–1430.

More, Neena Shah, and others (2012). Community Mobilization in Mumbai Slums to Improve Perinatal Care and Outcomes: A Cluster Randomized Controlled Trial. *PLoS Medicine / Public Library of Science*, vol. 9, issue 7, E1001257.

Morris, Saul S., and others (2004). Monetary incentives in primary health care and effects on use and coverage of preventive health care interventions in rural Honduras: cluster randomised trial. *Lancet*, vol. 364, issue 9450, pp. 2030–2037.

Mosler, H-J., S M Kraemer, and R B Johnston (2013). Achieving long-term use of solar water disinfection in Zimbabwe. *Public Health*, vol. 127, issue 1, pp. 92–98.

Murray, Joanna, and others (2018). Modelling the effect of a mass radio campaign on child mortality using facility utilisation data and the Lives Saved Tool (LiST): findings from a cluster randomised trial in Burkina Faso. *BMJ Global Health*, vol. 3, issue 4, e000828.

Murthy, Nirmala, and others (2019). The Impact of an mHealth Voice Message Service (mMitra) on Infant Care Knowledge, and Practices Among Low-Income Women in India: Findings from a Pseudo-Randomized Controlled Trial. *Maternal and Child Health Journal*, vol. 23, issue 12, pp. 1658–1669.

Najnin, Nusrat, and others (2017). Impact of adding hand-washing and water disinfection promotion to oral cholera vaccination on diarrhoea-associated hospitalization in Dhaka, Bangladesh: evidence from a cluster randomized control trial. *International Journal of Epidemiology*, vol. 46, issue 6, pp. 2056–2066.

Nguyen, Quang Ngoc, and others (2012). Effectiveness of community-based comprehensive healthy lifestyle promotion on cardiovascular disease risk factors in a rural Vietnamese population: a quasi-experimental study. *BMC Cardiovascular Disorders*, vol. 12, issue 56.

Njuguna, John (2019). Impact of a conditional cash transfer program aimed at promoting maternal and child health services in Kakamega County, Kenya. *Journal of Health Care for the Poor and Underserved*, vol. 30, issue 1, pp. 329–338.

Null, Clair, and others (2018). Effects of Water Quality, Sanitation, Handwashing, and Nutritional Interventions on Diarrhoea and Child Growth in Rural Kenya: A Cluster-randomised Controlled Trial. *Lancet Global Health*, vol. 6, issue 3, E316–E329.

O'Meara, Wendy Prudhomme, and others (2018). Improving rational use of ACTs through diagnosis-dependent subsidies: Evidence from a cluster-randomized controlled trial in western Kenya. *PLOS Medicine*, vol. 15, issue 7. Available at https://doi.org/10.1371/journal.pmed.1002607.

Odeny, Thomas A., and others (2019). Text messaging for maternal and infant retention in prevention of mother-to-child HIV transmission services: A pragmatic stepped-wedge cluster-randomized trial in Kenya. *PLoS Medicine / Public Library of Science*, vol. 16, E1002924.

Oyo-Ita Angela, and others (2016). Interventions for improving coverage of child immunization in low- and middle-income Countries. *Cochrane Database of Systematic Reviews*, CD008145.

Pandey, Priyanka, and others (2007). Informing resource-poor populations and the delivery of entitled health and social services in rural India: a cluster randomized controlled trial *JAMA, Journal of the American Medical Association*, vol. 298, issue 16, pp. 1867–1875.

Parvez, Sarker Masud, and others (2018). Achieving optimal technology and behavioural uptake of single and combined interventions of water, sanitation hygiene and nutrition, in an efficacy trial (WASH benefits) in rural Bangladesh. *Trials [Electronic Resource]*, vol. 19, issue 1, 358.

Patil, Sumeet R., and others (2013). A Randomized, Controlled Study of a Rural Sanitation Behaviour Change Program in Madhya Pradesh, India. Policy research working paper No. 6702. Washington D.C.: World Bank.

Pega, Frank, and others (2015). Unconditional cash transfers for assistance in humanitarian disasters: effect on use of health services and health outcomes in low- and middle-income countries. *Cochrane Database of Systematic Reviews*, vol. 9.

Pega, Frank, and others (2017). Unconditional cash transfers for reducing poverty and vulnerabilities: effect on use of health services and health outcomes in low- and middle-income countries. *Cochrane Database of Systematic Reviews*, vol. 11, issue 11.

Peletz, Rachel, and others (2013). Follow-up study to assess the use and performance of household filters in Zambia. *American Journal of Tropical Medicine & Hygiene*, vol. 89, issue 6, pp. 1190–1194.

Pickering, Amy J., and others (2015). Effect of a community-led sanitation intervention on child diarrhoea and child growth in rural Mali: a cluster-randomised controlled trial. *Lancet Global Health*, vol. 3, issue 11, E701–E711.

Ponce, Juan, and Jesus Ramos-Martin (2017). Impact of two policy interventions on dietary diversity in Ecuador. *Public Health Nutrition*, vol. 20, issue 8, pp. 1473–1480.

Porter, Catherine, and Radhika Goyal (2016). Social protection for all ages? Impacts of Ethiopia's Productive Safety Net Program on child nutrition. *Social Science & Medicine*, vol. 159, pp. 92–99.

Powell-Jackson, Timothy, and Kara Hanson (2012). Financial incentives for maternal health: impact of a national programme in Nepal. *Journal of Health Economics*, vol. 31, issue 1, pp. 271–284.

Quayyum, Zahidul, and others (2013). “Can community level interventions have an impact on equity and utilization of maternal health care" - evidence from rural Bangladesh. *International Journal for Equity in Health*, vol. 2, pp. 12-22.

Ranganathan, Meghna, and Mylene Lagarde (2012). Promoting healthy behaviours and improving health outcomes in low- and middle-income countries: a review of the impact of conditional cash transfer programmes. *Preventive Medicine*, vol. 55, S95–S105.

Reygadas, Fermín, and others (2018). Measuring User Compliance and Cost Effectiveness of Safe Drinking Water Programs: A Cluster-Randomized Study of Household Ultraviolet Disinfection in Rural Mexico. *American Journal of Tropical Medicine and Hygiene*, vol. 98, pp. 824–834.

Ruzagira, Eugene, and others (2017). Linkage to HIV care after home-based HIV counselling and testing in sub-Saharan Africa: A systematic review. *Tropical Medicine & International Health*, vol. 22, issue 7, pp. 807–821.

Saeterdal, Ingvil, and others (2014). Interventions aimed at communities to inform and/or educate about early childhood vaccination. *Cochrane Database of Systematic Reviews*, vol. 11.

Sarrafzadegan, Nizal, and others (2009). Do lifestyle interventions work in developing countries? Findings from the Isfahan Healthy Heart Program in the Islamic Republic of Iran. *Bulletin of the World Health Organization*, vol. 87, issue 1, pp. 39–50.

Solomon Sunil S., and others (2019). Integrated HIV testing, prevention, and treatment intervention for key populations in India: a cluster-randomised trial. *Lancet HIV*, vol. 6, issue 5, E283–E296.

Stoller, Nicole E., and others (2011). Efficacy of latrine promotion on emergence of infection with ocular Chlamydia trachomatis after mass antibiotic treatment: a cluster-randomized trial. *International Health*, vol. 3, issue 2, pp. 75–84.

Sturm, Roland, and others (2013). A cash-back rebate program for healthy food purchases in South Africa: results from scanner data. *American Journal of Preventive Medicine*, vol. 44, issue 6, pp. 567–572.

Suthar, Amitabh B., and others (2013). Towards universal voluntary HIV testing and counselling: a systematic review and meta-analysis of community-based approaches. *PLoS Medicine*, vol. 10, issue 8, E1001496.

Tian, Xiangyang, and others (2019). Evaluation of a multi-layered health promotion approach in rural China. *Global Health Promotion*, vol. 26, issue 4, pp. 14–24.

Tiruneh, Gizachew Tadele, and others (2020). Effectiveness of participatory community solutions strategy on improving household and provider health care behaviours and practices: A mixed-method evaluation. *PLoS One*, vol. 15, issue 2.

Tougher, Sarah, and others (2018). Effect of a multifaceted social franchising model on quality and coverage of maternal, newborn, and reproductive health-care services in Uttar Pradesh, India: a quasi-experimental study. *Lancet Global Health*, vol. 6, issue 2, E211.

Triyana, Margaret, and Anuraj H. Shankar (2017). The effects of a household conditional cash transfer programme on coverage and quality of antenatal care: a secondary analysis of Indonesia's pilot programme. *BMJ Open*, vol. 7, issue 10.

Tutor, Melba V. (2014). The impact of the Philippines’ conditional cash transfer program on consumption. *Philippine Review of Economics*, vol. 51, issue 1, pp. 117–161.

Wang, Paul C., and others (2015). A Cluster Randomised Trial on the Impact of Integrating Early Infant HIV Diagnosis with the Expanded Programme on Immunization on Immunization and HIV Testing Rates in Rural Health Facilities in Southern Zambia. *PLoS One*, vol. 10, issue 10.

Wang, Shengnan, and others (2016). Increasing Coverage of Hepatitis B Vaccination in China: A Systematic Review of Interventions and Implementation Experiences. *Medicine (Baltimore)*, vol. 95, issue 19, E3693.

Watson, S.I., and others (2016). The impact of user fees on health services utilization and infectious disease diagnoses in Neno District, Malawi: a longitudinal, quasi-experimental study. *BMC Health Services Research*, vol. 16, issue 16, article No. 595.

Wichaidit, Wit, and others (2019). Effectiveness of a large-scale handwashing promotion intervention on handwashing behaviour in Dhaka, Bangladesh. *Tropical Medicine & International Health*, vol. 24, issue 8, pp. 972–986.

Wu, D, and others (2018). Assessment of tobacco control advocacy behavioural capacity among public health faculty and students: a cluster intervention study. *Public health*, vol. 165, pp. 95–105.

Yeboah-Antwi, Kojo, and others (2019). Improving Sanitation and Hygiene through Community-led Total Sanitation: The Zambian Experience. *American Journal of Tropical Medicine and Hygiene*, vol. 100, issue 4, pp. 1005–1012.

Zhang, Zili, and Rong Zheng (2020). The Impact of Cigarette Excise Tax Increases on Regular Drinking Behaviour: Evidence from China. *International Journal of Environmental Research and Public Health*, vol. 17, issue 9.

Zhou, Huan, and others (2020). Conditional cash transfers, uptake of maternal and child health services, and health outcomes in western rural China. *BMC Public Health*, vol. 20, issue 1.

Zombré, David, Manuela De Allegri, and Valéry Ridde (2020). No effects of pilot performance-based intervention implementation and withdrawal on the coverage of maternal and child health services in the Koulikoro region, Mali: an interrupted time series analysis. *Health Policy and Planning*, vol. 35, issue 4, pp. 379–387.

References to studies included in energy sector EGM (96 studies)

Aklin, Michaël, and others (2017). Does basic energy access generate socioeconomic benefits? A field experiment with off-grid solar power in India. *Science Advances*, vol. 3, issue 5.

Akpandjar, George, and Carl Kitchens (2017). From Darkness to Light: The Effect of Electrification in Ghana, 2000–2010. *Economic Development and Cultural Change*, vol. 66, issue 1, pp. 31–54.

Ali, Akhter, and others (2019). Alternate energy sources for lighting among rural households in the Himalayan region of Pakistan: Access and impact. *Energy & Environment*, vol. 30, issue 7, pp. 1291–1312.

Barnes, Douglas F., Raluca Golumbeanu, and Issa Diaw (2016). Beyond Electricity Access: Output-Based Aid and Rural Electrification in Ethiopia. Washington, DC: World Bank.

Bedi, Arjun Singh, Lorenzo Pellegrini, and Lucia Tasciotti (2015). The effects of Rwanda's biogas program on energy expenditure and fuel use. *World Development*, vol. 67, pp. 461–474.

Bedi, Arjun Singh, Robert Sparrow, and Lucia Tasciotti (2017). The impact of a household biogas programme on energy use and expenditure in East Java. *Energy Economics*, vol. 68, pp. 66–76.

Bensch, Gunther, and Jörg Peters (2015). The intensive margin of technology adoption – Experimental evidence on improved cooking stoves in rural Senegal. *Journal of Health Economics*, vol. 42, issue C, pp. 44–63.

Bensch, Gunther, Jochen Kluve, and Jörg Peters (2011). Impacts of rural electrification in Rwanda. *Journal of Development Effectiveness*, vol. 3, issue 4, pp. 567–588.

Bensch, Gunther. (2019). The effects of market-based reforms on access to electricity in developing countries: a systematic review. *Journal of Development Effectiveness*, vol. 11, issue 2, pp. 165–188.

Beyene, Abebe, and others (2015). Do Improved Biomass Cookstoves Reduce Fuelwood Consumption and Carbon Emissions? Evidence from Rural Ethiopia Using a Randomized Treatment Trial with Electronic Monitoring (Availability Note: Information provided in

collaboration with the RePEc Project: http://repec.org). Policy research working paper No. 7324. Washington D.C.: World Bank.

Bluffstone, Randall, and others (2019). Do improved biomass cookstoves reduce PM2.5 concentrations? If so, for whom? Empirical evidence from rural Ethiopia. Washington, D.C.: World Bank.

Burke, Paul J., and Sandra Kurniawati (2018). Electricity subsidy reform in Indonesia: Demand-side effects on electricity use. *Energy Policy*, vol. 116, pp. 410–421.

Burrough-Boenisch, Joy, eds. (2014). *Access To Energy In Rwanda: Impact Evaluation Of Activities Supported By The Dutch Promoting Renewable Energy Programme*. IOB Evaluation, No.396 (August). Ministry of Foreign Affairs of the Netherlands. Available at https://www.government.nl/documents/reports/2015/04/28/iob-access-to-energy-in-rwanda-impact-evaluation-of-activities-supported-by-the-dutch-promoting-renewable-energy-programme.

Chakravorty, Ujjayant, Kyle Emerick, and Majah-Leah Ravago (2016). Lighting up the Last Mile: The Benefits and Costs of Extending Electricity to the Rural Poor. Discussion Paper, pp. 16–22 (July). Resources for the Future (RFF). Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2851907.

Cui, Jingbo, Junjie Zhang, and Yang Zheng (2018). Carbon Pricing Induces Innovation: Evidence from China's Regional Carbon Market Pilots. *AEA Papers and Proceedings*, vol. 108, pp. 453–457.

Dasso, Rosamaría, and Fernando Fernandez (2015). The effects of electrification on employment in rural Peru. *IZA Journal of Labor & Development*, vol. 4, issue 6.

Davis, Lucas W., Alan Fuchs, and Paul Gertler (2014). Cash For Coolers: Evaluating A Large-Scale Appliance Replacement Program in Mexico. *American Economic Journal: Economic Policy*, vol. 6, issue 4, pp. 207–238.

Dinkelman, Taryn (2011). The Effects of Rural Electrification on Employment: New Evidence from South Africa. Working Paper No. 1255. Princeton University, Woodrow Wilson School of Public and International Affairs, Research Program in Development Studies.

Dong, Feng, and others (2019). Can a carbon emission trading scheme generate the Porter effect? Evidence from pilot areas in China. *Science of the Total Environment*, vol. 653, pp. 565–577.

Dong, Zhi-Qing, and others (2020). The validity of carbon emission trading policies: Evidence from a quasi-natural experiment in China. *Advances in Climate Change Research*, vol. 11, issue 2, pp. 102–109.

Du, Weijian, and Mengjie Li. (2020). Assessing the impact of environmental regulation on pollution abatement and collaborative emissions reduction: Micro-evidence from Chinese industrial enterprises. *Environmental Impact Assessment Review*, vol. 82.

Du, Yimeng, and Kenji Takeuchi (2019). Can climate mitigation help the poor? Measuring impacts of the CDM in rural China. *Journal of Environmental Economics and Management*, vol. 95, pp. 178–197.

Du, Yimeng, and Kenji Takeuchi (2020). Does a small difference make a difference? Impact of feed-in tariff on renewable power generation in China. *Energy Economics*, vol. 87.

Filippini, Massimo, and Lin Zhang (2019). Impacts of heat metering and efficiency retrofit policy on residential energy consumption in China. *Environmental Economics and Policy Studies*, vol. 21, issue 2, pp. 203–216. Available at https://doi.org/10.1007/s10018-018-0227-8.

Gao, Hang, and Johannes van Biesebroeck. (2014). Effects of Deregulation and Vertical Unbundling on the Performance of China's Electricity Generation Sector. *The Journal of Industrial Economics*, vol. 62, issue 1, pp. 41–76.

Gao, Yuning, and others (2020). Evaluation of effectiveness of China's carbon emissions trading scheme in carbon mitigation. *Energy Economics*, vol. 90, 104872.

Gerard, Francois. (2013). What Changes Energy Consumption, and for How Long? New Evidence from the 2001 Brazilian Electricity Crisis. Discussion Paper No. 13-6. Resources for the Future.

Grogan, Louise (2016). Household Electrification, Fertility, and Employment: Evidence from Hydroelectric Dam Construction in Colombia. *Journal of Human Capital*, vol. 10, issue 1, pp. 109–158. University of Chicago Press.

Grogan, Louise (2018). Time use impacts of rural electrification: Longitudinal evidence from Guatemala. *Journal of Development Economics*, vol. 135, pp. 304–317.

Grogan, Louise, and Asha Sadanand (2013). Rural Electrification and Employment in Poor Countries: Evidence from Nicaragua. *World Development*, vol. 43, pp. 252–265.

Hanna, Rema, Esther Duflo, and Michael Greenstone (2014). Up in Smoke: The Influence of Household Behaviour on the Long-Run Impact of Improved Cooking Stoves. *American Economic Journal: Economic Policy*, vol. 8, issue 1, pp. 80–114.

He, Xiaoping (2019). China's electrification and rural labor: Analysis with fuzzy regression discontinuity. *Energy Economics*, vol. 81, pp. 650–660.

Hu, Yuanan, and Hefa Cheng (2017). Displacement efficiency of alternative energy and trans-provincial imported electricity in China. *nature communications*, vol. 8, article No. 14590.

Hu, Yucai, and others (2020). Can carbon emission trading scheme achieve energy conservation and emission reduction? Evidence from the industrial sector in China. *Energy Economics*, vol. 85.

International Finance Corporation (2012). Estimating Employment Effects of Powerlinks Transmission Limited Project in India and Bhutan. Washington, D.C.

Jagger, Pamela, and Ipsita Das (2018). Implementation and scale-up of a biomass pellet and improved cookstove enterprise in Rwanda. *Energy for Sustainable Development*, vol. 46, pp. 32–41.

John, Prachitha, and others (2015). Impact of Private Sector Participation on access and quality of services: systematic review of evidence from the electricity, telecommunications and water supply sectors. *Journal of Development Effectiveness*, vol. 7, issue 1, pp. 64–89.

Khandker, Shahidur R., and others (2009). Welfare Impacts of Rural Electrification: Evidence from Vietnam. Policy research working paper No. 5057. Washington D.C.: World Bank.

Khandker, Shahidur R., and others (2012). Who Benefits Most from Rural Electrification? Evidence in India. Policy research working paper No. 6095. Washington D.C.: World Bank.

Kim, Sun S., and others (2012). A systematic review of smoking cessation intervention studies in China. *Nicotine & Tobacco Research*, vol. 14, issue 8, pp. 891–899.

Lee, Chui Ying, and others (2019). The Impact of an Energy Efficiency Improvement Policy on the Economic Performance of Electricity-Intensive Firms in Ghana. *energies*, vol. 12, issue 19, 3684.

Lee, Kenneth, Edward Miguel and Catherine Wolfram (2020). Experimental Evidence on the Economics of Rural Electrification. *Journal of Political Economy*, vol. 128, issue 4, pp. 1523–1565.

Lenz, Luciane, and others (2017). Does Large-Scale Infrastructure Investment Alleviate Poverty? Impacts of Rwanda's Electricity Access Roll-Out Program. *World Development*, vol. 89, pp. 88–110.

Li, Jianglong, and Boqiang Lin (2017). Environmental impact of electricity relocation: A quasi-natural experiment from interregional electricity transmission. *Environmental Impact Assessment Review*, vol. 66, pp. 151–161.

Limi, Atsushi, and others (2019). Energy-Saving Effects of Progressive Pricing and Free CFL Bulb Distribution Program: Evidence from Ethiopia. *World Bank Economic Review*, vol. 33, issue 2, pp. 461–478.

Lin, Boqiang, and Xing Chen (2018). Is the implementation of the Increasing Block Electricity Prices policy really effective? Evidence based on the analysis of synthetic control method. *Energy*, vol. 163, pp. 734–750.

Lipscomb, Molly, A. Mushfiq Mobarak, and Tania Barham (2013). Development Effects of Electrification: Evidence from the Topographic Placement of Hydropower Plants in Brazil. *American Economic Journal: Applied Economics*, vol. 5, issue 2, pp. 200–231.

Litzow, Erin L, Subhrendu K. Pattanayak, and Tshering Thinley (2019). Returns to rural electrification: Evidence from Bhutan. *World Development*, vol. 121, pp. 75–96.

Liu, Zuankuo, Yufang Jiang, and Dennis Bolayog (2019). Does "replacing business tax with value-added tax" promote the energy efficiency of the logistics industry in China? *Environmental Science and Pollution Research*, vol. 26, issue 32, pp. 33169–33180.

Malik, Kabir, and others (2011). Estimating the Impact of Restructuring on Electricity Generation Efficiency: The Case of the Indian Thermal Power Sector. NBER Working Paper No. 17383.

Mensah, Justice Tei (2018). Jobs! Electricity shortages and unemployment in Africa. Policy Research Working Paper No. 8415. Washington D.C.: World Bank.

Nawaz, Saima, and Nasir Iqbal (2020). The impact of unconditional cash transfer on fuel choices among ultra-poor in Pakistan: Quasi-experimental evidence from the Benazir Income Support Program. *Energy Policy*, vol. 142.

Pan, Xiongfeng, and others (2019). Effects of China's environmental policy on carbon emission efficiency. *International Journal of Climate Change Strategies and Management*, vol. 11, issue 3, pp. 326–340.

Pope, Daniel, and others (2017). Real-life effectiveness of 'improved' stoves and clean fuels in reducing PM 2.5 and CO: Systematic review and meta-analysis. *Environment International*, vol. 101, pp. 7–18.

Qi, Shaozhou, Shihan Cheng, and Jingbo Cui (2021). Environmental and economic effects of China's carbon market pilots: Empirical evidence based on a DID model. *Journal of Cleaner Production*, vol. 279, 123720.

Quansah, Reginald, and others (2017). Effectiveness of interventions to reduce household air pollution and/or improve health in homes using solid fuel in low-and-middle income countries: A systematic review and meta-analysis. *Environment International*, vol. 103, pp. 73–90.

Rains, Eily, and Ronald J. Abraham (2018). Rethinking barriers to electrification: Does government collection failure stunt public service provision? *Energy Policy*, vol. 114, pp. 288–300.

Rathi, Sambhu Singh, and Claire Vermaak (2018). Rural electrification, gender and the labor market: A cross-country study of India and South Africa. *World Development*, vol. 109, pp. 346–359.

Salmon, Claire, and Jeremy Tanguy (2016). Rural Electrification and Household Labor Supply: Evidence from Nigeria. *World Development*, vol. 82, pp. 48–68.

Samad, Hussain A., and Fan Zhang (2016). Benefits of Electrification and the Role of Reliability: Evidence from India. Policy Research Working Paper No. 7889. Washington D.C.: World Bank.

Samad, Hussain A., and Fan Zhang (2017). Heterogeneous Effects of Rural Electrification: Evidence from Bangladesh. Policy Research Working Paper No. 8102. Washington D.C.: World Bank.

Samad, Hussain A., and Fan Zhang (2019). Electrification and Women's Empowerment: Evidence from Rural India. Policy Research Working Paper No. 8796. Washington D.C.: World Bank.

Schmid, Gisèle (2012). The development of renewable energy power in India: Which policies have been effective? *Energy Policy*, vol. 45, pp. 317–326.

Sen, Anupama, Rabindra Nepal, and Tooraj Jamasb (2018). Have Model, Will Reform: Assessing the Outcomes of Electricity Reforms in Non-OECD Asia. *The Energy Journal*, vol. 39, issue 4, pp. 181–210.

Shen, Jun, Pengcheng Tang, and Hao Zeng (2020). Does China's carbon emission trading reduce carbon emissions? Evidence from listed firms. *Energy for Sustainable Development*, vol. 59, pp. 120–129.

Tagliapietra, Simone, and others (2020). The impact of electrification on labour market outcomes in Nigeria. *Economia Politica*, vol. 37, issue 3, pp. 737–779.

Tan-Soo, Jie-Sheng, and others (2019). Using electricity prices to curb industrial pollution. *Journal of Environmental Management*, vol. 248, 109252.

Tang, Kai, and others (2020a). The effectiveness and heterogeneity of carbon emissions trading scheme in China. *Environmental Science and Pollution Research*, pp. 1–13.

Tang, Kai, and others (2020b). Urban carbon emission intensity under emission trading system in a developing economy: evidence from 273 Chinese cities. *Environmental Science and Pollution Research*, vol. 28, pp. 1–12.

Tang, Keyi, and Yingjiao Shen (2020). Do China-financed dams in Sub-Saharan Africa improve the region's social welfare? A case study of the impacts of Ghana's Bui Dam. *Energy Policy*, vol. 136.

Urpelainen, Johannes, and Semee Yoon (2017). Can product demonstrations create markets for sustainable energy technology? A randomized controlled trial in rural India. *Energy Policy*, vol. 109, pp. 666–675.

Van de Walle, Dominique, and others (2017). Long-term Gains from Electrification in Rural India. *World Bank Economic Review*, vol. 31, issue 2, pp. 385–411.

Wang, Han, and others (2019). Can a carbon trading system promote the transformation of a low-carbon economy under the framework of the porter hypothesis? —Empirical analysis based on the PSM-DID method. *Energy Policy*, vol. 129, pp. 930–938.

Wang, Limin, and others (2011). Quantifying Carbon and Distributional Benefits of Solar Home System Programs in Bangladesh. Policy Research Working Paper No. 5545. Washington, D.C.: World Bank.

Wang, Linshan, Chuanming Liu, and Xi Yang (2020). Research on Carbon Emission Reduction Effect of China's Carbon Trading Pilot - Based on Different in Different Method. *Advances in Social Sciences Research Journal*, vol. 7, issue 5, pp. 240–250.

Wang, Qian, Cuiyun Gao, and Dai Shuanping. (2019). Effect of the Emissions Trading Scheme on CO_2 Abatement in China. *Sustainability*, vol. 11, issue 4, 1055.

Wang, Yongpei, and others (2018). Does grid-connected clean power promote regional energy efficiency? An empirical analysis based on the upgrading grid infrastructure across China. *Journal of Cleaner Production*, vol. 186, pp. 736–747.

Wang, Zhaohua, Yefei Sun, and Bo Wang (2020). Policy cognition is more effective than step tariff in promoting electricity saving behaviour of residents. *Energy Policy*, vol. 139.

Wen, Yang, and others (2020). Does China's Carbon Emissions Trading Scheme Really Work? A Case Study of the Hubei Pilot. *Journal of Cleaner Production*, vol. 277, 124151.

Xu, Shang, and H. Allen Klaiber (2019). The impact of new natural gas pipelines on emissions and fuel consumption in China. *Resource and Energy Economics*, vol. 55, pp. 49–62.

Yan, Yaxue, and others (2020). Emissions trading system (ETS) implementation and its collaborative governance effects on air pollution: The China story. *Energy Policy*, vol. 138.

Yi, Lan, and others (2020). Evaluation on the effectiveness of China's pilot carbon market policy. *Journal of Cleaner Production*, vol. 246, No. 119039. Available at https://www.sciencedirect.com/science/article/abs/pii/S0959652619339095?via%3Dihub.

Yu, Yantuan, and Ning Zhang (2019). Does smart city policy improve energy efficiency? Evidence from a quasi-natural experiment in China. *Journal of Cleaner Production*, vol. 229, pp. 501–512.

Zhang, Haijun, Maosheng Duan, and Zhe Deng (2019). Have China's pilot emissions trading schemes promoted carbon emission reductions? the evidence from industrial sub-sectors at the provincial level. *Journal of Cleaner Production*, vol. 234, pp. 912–924.

Zhang, Haoran, and others (2020). Has China's Emission Trading System Achieved the Development of a Low-Carbon Economy in High-Emission Industrial Subsectors? *Sustainability*, vol. 12, issue 13, 5370.

Zhang, Hui, and Li-Wei Fan (2019). Can emission trading help to improve energy efficiency in China? *Energy Efficiency*, vol. 12, issue 4, pp. 979–991.

Zhang, Ning, Keke Zhao, and Yuantuan Yu (2020). The effect of environmental regulation on air pollution, productivity, and factor structure: a quasi-natural experiment evidence from China. *Environmental Science and Pollution Research*, vol. 27, issue 16, pp. 20392-20409.

Zhang, Pan. (2019). Do energy intensity targets matter for wind energy development? Identifying their heterogeneous effects in Chinese provinces with different wind resources. *Renewable Energy*, vol. 139, pp. 968–975.

Zhang, Wei, and others (2020). Emission reduction effect and carbon market efficiency of carbon emissions trading policy in China. *Energy*, vol. 196.

Zhang, Yan, and Jiekuan Zhang (2019). Estimating the impacts of emissions trading scheme on low-carbon development. *Journal of Cleaner Production*, vol. 238, issue 1.

Zhang, Yifei, and others (2020). The effect of emission trading policy on carbon emission reduction: Evidence from an integrated study of pilot regions in China. *Journal of Cleaner Production*, vol 265, issue 2, 121843.

Zhang, Yue-Jun, and others (2018). The impact of China's Central Rise Policy on carbon emissions at the stage of operation in road sector. *Economic Modelling*, vol. 71, pp. 159–173.

Zhang, Zibin, Wenxin Cai, and Xiangzhao Feng. (2017). How do urban households in China respond to increasing block pricing in electricity? Evidence from a fuzzy regression discontinuity approach. *Energy Policy*, vol. 105, pp. 161–172.

Zhao, Xiaoli, Haitao Yin, and Yue Zhao (2015). Impact of environmental regulations on the efficiency and CO_2 emissions of power plants in China. *Applied Energy*, vol. 149, pp. 238-247.

Zhou, Bo, and others (2019). How does emission trading reduce China's carbon intensity? An exploration using a decomposition and difference-in-differences approach. *Science of the Total Environment*, vol. 676, pp. 514–523.

Zhou, Di, and others (2020). Does Emission Trading Boost Carbon Productivity? Evidence from China's Pilot Emission Trading Scheme. *International Journal of Environmental Research and Public Health*, vol. 17, issue 15, 5522.

Zou, Hong, and others (2020). The effect of Increasing Block Pricing on Urban Households' Electricity Consumption: Evidence from difference-in-differences models. *Journal of Cleaner Production*, vol. 257, issue 3, 120498.

Additional references

Ajibade, Idowu, and Ellis A. Adams (2019). Planning principles and assessment of transformational adaptation: towards a refined ethical approach. *Climate and Development*, vol. 11, issue 10, pp. 850–862.

Akpandjar, George, and Carl Kitchens (2017). From Darkness to Light: The Effect of Electrification in Ghana, 2000–2010. *Economic Development and Cultural Change*, vol. 66 issue 1, pp. 31-54.

Arnott, Bronia, and others (2014). Efficacy of behavioural interventions for transport behaviour change: systematic review, meta-analysis and intervention coding. *International Journal of Behavioural Nutrition and Physical Activity*, vol. 11, issue 1, 133.

Banerjee, Abhijit, and others (2017). From Proof of Concept to Scalable Policies: Challenges and Solutions, with an Application. *Journal of Economic Perspectives*, vol. 31, issue 4, pp. 73–102.

Bird, Neil, Yue Cao, and Andriana Quevedo (2019). *Transformational change in the Climate Investment Funds: A synthesis of the evidence*, pp. 80. Available at https://www.odi.org/publications/11273-transformational-change-climate-investment-funds-synthesis-evidence.

Borenstein, Michael, and others (2011). *Introduction to Meta-analysis*. Hoboken, New Jersey, USA: John Wiley & Sons.

Bouyé, Mathilde, Sven Harmeling, and Nils-Sjard Schulz (2018). *Connecting the Dots: Elements for a Joined-Up Implementation of the 2030 Agenda and Paris Agreement*. Bonn, Germany: Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH; Washington, D.C.: World Resources Institute.

Buchner, Barbara, and others (2019). *Global Landscape of Climate Finance 2019*. London: Climate Policy Initiative (CPI).

Climate Investment Funds (2015). Climate Investment Funds: accomplishments, transformational impact, and additionality in the climate finance architecture. Presented at the Joint Meeting of the CIF and SCF Trust Fund Committees. Available at https://www.climateinvestmentfunds.org/sites/cif_enc/files/meeting-documents/ctf_scf_15_3_cif_strategic_paper_final_posting.pdf.

_____(2019). *Evaluation of Transformational Change in the Climate Investment Funds*. Final evaluation report (January).

_____(2020). *Signals of Transformational Change: Insights from the Evaluation of Transformational Change in the CIF*.

_____(2021). *Transformational Change Concepts*. Updated May 2021. Available at https://www.climateinvestmentfunds.org/sites/cif_enc/files/knowledge-documents/tclp_workshop_updated_tc_concepts_may2021.pdf.

Cohen, Jacob (1988). *Statistical Power Analysis for the Behavioural Sciences* (2nd ed.) New York: Routledge.

Cole, Daniel H. (2008). Climate Change, Adaptation, and Development. *UCLA Journal of Environmental Law and Policy*, vol. 26, issue 1, pp. 1–19.

Cooper, Harris (ed.) (2010). *Research Synthesis and Meta-analysis: A Step-by-step Approach* (5th ed.) Thousand Oaks, California: Sage Publications.

Corbett, Elizabeth L., and others (2007). HIV incidence during a cluster-randomized trial of two strategies providing voluntary counselling and testing at the workplace, Zimbabwe. *AIDS*, vol. 21, issue 4, pp. 483–489.

Deaton, Angus (2010). Instruments, Randomization, and Learning about Development. *Journal of Economic Literature*, vol. 48, issue 2, pp. 424–455.

Duan, Maosheng, Tao Pang, and Xiliang Zhang (2014). Review of Carbon Emissions Trading Pilots in China. *Energy & Environment*, vol. 25, issue 3–4, pp. 527–549.

Ebinger, Jane O., and Walter Vergara (2011). *Climate Impacts on Energy Systems: Key Issues for Energy Sector Adaptation.* A World Bank Study No. 60051. Washington, D.C.

Edenhofer, Ottmar, and others (eds.) (2014). Summary for Policymakers. In *Climate Change 2014: Mitigation of Climate Change. IPCC Working Group III Contribution to AR5*. Cambridge, UK and New York, USA: Cambridge University Press.

Egger, Matthias, and others (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ (Clinical Research Ed.)*, vol. 315, article No. 7109, pp. 629–634.

Falconí, Fander, and others (2019). Toxic Income as a Trigger of Climate Change. *Sustainability*, vol. 11, issue 8, 2448.

Fazey, Ioan, and others (2017). Transformation in a changing climate: A research agenda. *Climate and Development*, 10(3), 197–217. Available at https://www.tandfonline.com/doi/full/10.1080/17565529.2017.1301864.

Fedele, Giacomo, and others (2019). Transformative adaptation to climate change for sustainable social-ecological systems. *Environmental Science & Policy*, vol. 101, pp. 116–125.

Feinstein, Oswaldo (2019). Dynamic evaluation for transformational change. In *Evaluation for Transformational Change: Opportunities and Challenges for the Sustainable Development Goals*, Rob D. van den Berg, Cristina Magro, and Silvia S Mulder (eds.) (pp. 17–31). Exeter, UK: International Development Evaluation Association (IDEAS). Available at https://ideas-global.org/wp-content/uploads//2019/11/2019-11-05-Final_IDEAS_EvaluationForTransformationalChange.pdf.

Feola, Giuseppe (2015). Societal transformation in response to global environmental change: A review of emerging concepts. *Ambio*, vol. 44, issue 5, pp. 376–390.

Few, Roger, and others (2017). Transformation, adaptation and development: relating concepts to practice. *Palgrave Communications*, vol. 3, issue 1, pp. 1–9.

Global Commission on Adaptation (2019). *Adapt now: a global call for leadership on climate resilience.* Flagship reports. Available at https://gca.org/reports/adapt-now-a-global-call-for-leadership-on-climate-resilience/.

Gurevitch, Jessica, and others (2018). Meta-analysis and the science of research synthesis. *Nature*, vol. 555, article No. 7695, pp. 175–182.

Guyatt, Gordon H., and others (2011). GRADE Guidelines: 9. rating up the quality of evidence. *Journal of Clinical Epidemiology*, vol. 64, issue 12, pp. 1311–1316. https://doi.org/10.1016/j.jclinepi.2011.06.004

Habimana, Olivier (2017). Gender Differences in Time Allocation: Evidence from Rwanda. Working Paper No. 8. Turin School of Development, International Training Centre of the International Labour Organization.

Independent Evaluation Group (2016). Supporting transformational change for poverty reduction and shared prosperity: Lessons from World Bank Group experience. Washington D.C.: World Bank. Available at http://ieg.worldbankgroup.org/sites/default/files/Data/reports/WBGSupportTransformationalEngagements.pdf.

Initiative for Climate Action Transparency (2017). Assessing the transformational impacts of policies and actions. Available at https://climateactiontransparency.org/wp-content/uploads/2020/04/Transformational-Change-Assessment-Guide.pdf.

Intergovernmental Panel on Climate Change (2014). *Climate Change 2014: Synthesis Report.* Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. [Core Writing Team, Rajendra K.Pachauri, and Leo A. Meyer (eds.)].

_____(2018). *IPCC Sixth Assessment Report*.

International Climate Fund (2014). Independent mid-term evaluation of the International Climate Fund. Available at http://iati.dfid.gov.uk/iati_documents/4773708.pdf.

International Energy Agency (2019a). *World Energy Outlook 2019*.

_____(2019b). *Securing Investments in Low-Carbon Power Generation Sources*. OECD.

_____(2020a). *China's Emissions Trading Scheme*.

_____(2020b). *World Energy Outlook 2020*.

International Finance Corporation (2019). *AIMM Sector Framework Brief: Power* (AIMM Sector Frameworks Consultation). Available at https://www.ifc.org/wps/wcm/connect/9f89e8f3-62fd-4af0-9c23-3c59f74a2630/AIMM-Power-Consultation.pdf?MOD=AJPERES&CVID=nmTfdIr

International Labour Organization (2003). Guidelines concerning a statistical definition of informal employment. Normative instrument (November). Available at http://www.ilo.org/global/statistics-and-databases/standards-and-guidelines/guidelines-adopted-by-international-conferences-of-labour-statisticians/WCMS_087622/lang--en/index.htm.

Kates, Robert W., William R. Travis, and Thomas J. Wilbanks (2012). Transformational adaptation when incremental adaptations to climate change are insufficient. *Proceedings of the National Academy of Sciences*, vol. 109, issue 19, pp. 7156–7161.

Lally, Phillippa, and others (2010). How are habits formed: Modelling habit formation in the real world. *European Journal of Social Psychology*, vol. 40, issue 6, pp. 998–1009.

Lee, Shoou-Yih Daniel, and others (2013). Organizational transformation: a systematic review of empirical research in health care and other industries. *Medical Care Research and Review*, vol. 70, issue 2, pp. 115-142. Available at https://pubmed.ncbi.nlm.nih.gov/22955697.

Lenz, Luciane, and others (2015). Does Large Scale Infrastructure Investment Alleviate Poverty? Impacts of Rwanda's Electricity Access Roll-Out Program. *Ruhr Economic Papers*, vol. 555, pp. 1–55. Available at https://doi.org/10.4419/86788636.

Madrian, Brigitte C. (2014). Applying Insights from Behavioural Economics to Policy Design. *Annual Review of Economics*, vol. 6, issue 1, pp. 663–688.

Mapfumo, Paul, and others (2017). Pathways to transformational change in the face of climate impacts: an analytical framework. *Climate and Development*, vol. 9, issue 5, pp. 439–451.

Memon, Zahid A., and others (2015). Impact of a community-based perinatal and newborn preventive care package on perinatal and neonatal mortality in a remote mountainous district in Northern Pakistan. *BMC Pregnancy & Childbirth*, vol. 15, No. 106.

Michie, Susan, Maartje M van Stralen, and Robert West (2011). The behaviour change wheel: a new method for characterising and designing behaviour change interventions. *Implementation Science*, vol. 6, No. 42.

Muralidharan, Karthik, and Paul Niehaus (2017). Experimentation at Scale. *Journal of Economic Perspectives*, vol. 31, issue 4, pp. 103–124.

Neufeldt, Henry, and others (2018). *The Adaptation Gap Report 2018.* Nairobi, Kenya: United Nations Environment Programme (UNEP).

O'Brien, Karen, and Linda Sygna (2013). Responding to climate change: The three spheres of transformation. Transformation in Changing Climate Proceedings, 19-21 June. Oslo, Norway: University of Oslo. Available at https://www.sv.uio.no/iss/english/research/projects/adaptation/publications/1-responding-to-climate-change---three-spheres-of-transformation_obrien-and-sygna_webversion_final.pdf.

Olsen, Karen Holm, and Jørgen Villy Fenhann (2016). Transformational change taxonomy: Methodological framework for the assessment of transformational change in NAMAs. Version 1. NAMA partnership, International Partnership on Mitigation and MRV, UNFCCC. Available at https://www.transparency-partnership.net/sites/default/files/tc-taxonomy---nama-partnership-web.pdf.

Partnership for Market Readiness (2014). A Survey of the MRV Systems for China's ETS Pilots. World Bank Technical Note No. 8. Available at https://openknowledge.worldbank.org/handle/10986/21816.

Parvez Sarker Masud, and others (2018). Achieving optimal technology and behavioural uptake of single and combined interventions of water, sanitation hygiene and nutrition, in an efficacy trial (WASH benefits) in rural Bangladesh. *Trials*, vol. 19, issue 1, 358.

Petticrew, Mark, and Helen Roberts (eds.). (2006). *Systematic Reviews in the Social Sciences*. Malden/Oxford/Carlton: Blackwell Publishing Ltd.

Prowse, Martin, and Birte Snilstveit (2010). Impact evaluation and interventions to address climate change: a scoping study. *Journal of Development Effectiveness*, vol. 2, issue 2, pp. 228–262.

Puri, Jyotsna (2019). Transformational Change: The Challenge of a Brave New World. In *Sustainable Global Value Chains*, Michael Schmidt, and others (eds.), (*2*), pp. 305–326. Cham, Switzerland: Springer International Publishing.

Puri, Jyotsna, and others (2021). Assessing the likelihood for transformational change at the Green Climate Fund. IEU learning paper, No. 05 (March). Songdo, South Korea: Independent Evaluation Unit, Green Climate Fund.

Rankin, Kristen, and others (2016). Adolescent sexual and reproductive health: An evidence gap map. Evidence Gap Map Report 5. International Initiative for Impact Evaluation.

Ross Strategic and Community Science (2017). Climate Investment Funds: Synthesis of transformational change concepts. Available upon request (Contact ciftclp@worldbank.org).

Sawilowsky, Shiomo S. (2009). New Effect Size Rules of Thumb. *Journal of Modern Applied Statistical Methods*, vol. 8, issue 2, pp. 597–599.

Stanley, T.D., and Hristos Doucouliagos (2014). Meta-regression approximations to reduce publication selection bias. *Research Synthesis Methods*, vol. 5, issue 1, pp. 60–78.

Stuart, Christina, and Silvia Escudero. (2017). *Energy and Climate Change Adaptation in Developing Countries*. Echsborn: European Union Energy Initiative Partnership Dialogue Facility (EUEI PDF). Available at https://sun-connect-news.org/fileadmin/DATEIEN/Dateien/New/euei_pdf_2017_energy_and_climate_change_adaptation_in_developing_countries__1_.pdf.

Termeer, Catrien J.A.M., Art Dewulf, and G. Robbert Biesbroek (2017). Transformational change: governance interventions for climate change adaptation from a continuous change perspective. *Journal of Environmental Planning and Management*, vol. 60, issue 4, pp. 558–576.

Thomalla, Frank, and others (2018). Transforming Development and Disaster Risk. *Sustainability*, vol. 10, issue 5, pp. 1–12.

Thornton, Thomas F., and Claudia Comberti (2017). Synergies and trade-offs between adaptation, mitigation and development. *Climatic Change*, vol. 140, issue 1, pp. 5–18.

Triyana, Margaret, and Anuraj H. Shankar (2017). The effects of a household conditional cash transfer programme on coverage and quality of antenatal care: a secondary analysis of Indonesia's pilot programme. *BMJ Open*, vol. 7, issue 10.

United Nations Environment Programme (2021). *Adaptation Gap Report 2020.* Nairobi: UNEP.

van den Berg, Rob D., and Lee Cando-Noordhuizen (2017). Action on Climate Change: What Does It Mean and Where Does It Lead To? In *Evaluating Climate Change Action for Sustainable Development*, Juha I. Uitto, Jyotsna Puri, and Rob D. van den Berg (eds.) (pp. 13–34). Washington, DC: Springer International Publishing.

Waddington, Hugh, and others (2019). Citizen engagement in public services in low- and middle-income countries: A mixed-methods systematic review of participation, inclusion, transparency and accountability (PITA) initiatives. *Campbell Systematic Reviews*, vol. 15, issue 1–2.

Watts, Nick, and others (2018). The Lancet Countdown on health and climate change: from 25 years of inaction to a global transformation for public health. *The Lancet*, vol. 391, No. 10120, pp. 581–630.

Westphal, Michael I., and Joe Thwaites (2016). *Transformational climate finance: An exploration of low-carbon energy*. Available at https://www.wri.org/publication/transformational-climate-finance.

Wienges, Sebastian, and others (2017). Guidance Note on How to Design Transformational Mitigation Projects a Checklist for Project Developers and Managers. Available at https://transparency-partnership.net/sites/default/files/51351_-_tc_publication_v2-web_final.pdf.

Williams, Anna, Joseph Dickman, and Regan Smurthwaite (2021). Advancing Evaluation and Learning on Transformational Change: Lessons From the Climate Investment Funds' Transformational Change Learning Partnership. *American Journal of Evaluation*, vol. 42, issue 1, pp. 90–109.

Williams, Anna, Joseph Dickman, and Regan Smurthwaite (2021). Advancing evaluation and learning on transformational change: Lessons from the Climate Investment Funds' Transformational Change Learning Partnership. *American Journal of Evaluation*, vol. 42, issue 1, pp. 90-109. Available at https://journals.sagepub.com/doi/10.1177/1098214020970283.

Wolfram, Catherine, Orie Shelef, and Paul Gertler (2012). How Will Energy Demand Develop in the Developing World? *Journal of Economic Perspectives*, vol. 26, issue 1, pp. 119–138.

World Bank (2021a). Access to Electricity (% of Population) | Data. Sustainable Energy for All (SE4ALL) Database. Available at https://data.worldbank.org/indicator/EG.ELC.ACCS.ZS.

World Bank (2021b). Access to Electricity (% of Population)—South Asia | Data. Sustainable Energy for All (SE4ALL) Database. Available at https://data.worldbank.org/indicator/EG.ELC.ACCS.ZS?locations=8S.

World Energy Council (2015). *2015 World Energy Issues Monitor: Energy price volatility - the new normal*. Available at https://www.worldenergy.org/assets/images/imported/2015/01/2015-World-Energy-Issues-Monitor.pdf.

World Resources Institute (2018). *Climate Access Indicators Tool (CAIT): Paris Contributions Map*. Available at http://cait.wri.org/indc/.

Zhang, Yan, and Jiekuan Zhang (2019). Estimating the impacts of emissions trading scheme on low-carbon development. *Journal of Cleaner Production*, vol. 238, issue 1.

Zhao, Daiqing, Wenjun Wang, and Zhigang Luo (2019). *A Brief Overview of China's ETS Pilots: Deconstruction and Assessment of Guangdong's Greenhouse Gas Emission Trading Mechanism*. Singapore: Springer.

Zhou, Yu, and others (2020). Addressing climate change through a market mechanism: a comparative study of the pilot emission trading schemes in China. *Environmental Geochemistry and Health*, vol. 42, issue 3, pp. 745–767.

INFORMATION ABOUT THIS EVIDENCE SYNTHESIS

A. ROLES AND RESPONSIBILITIES

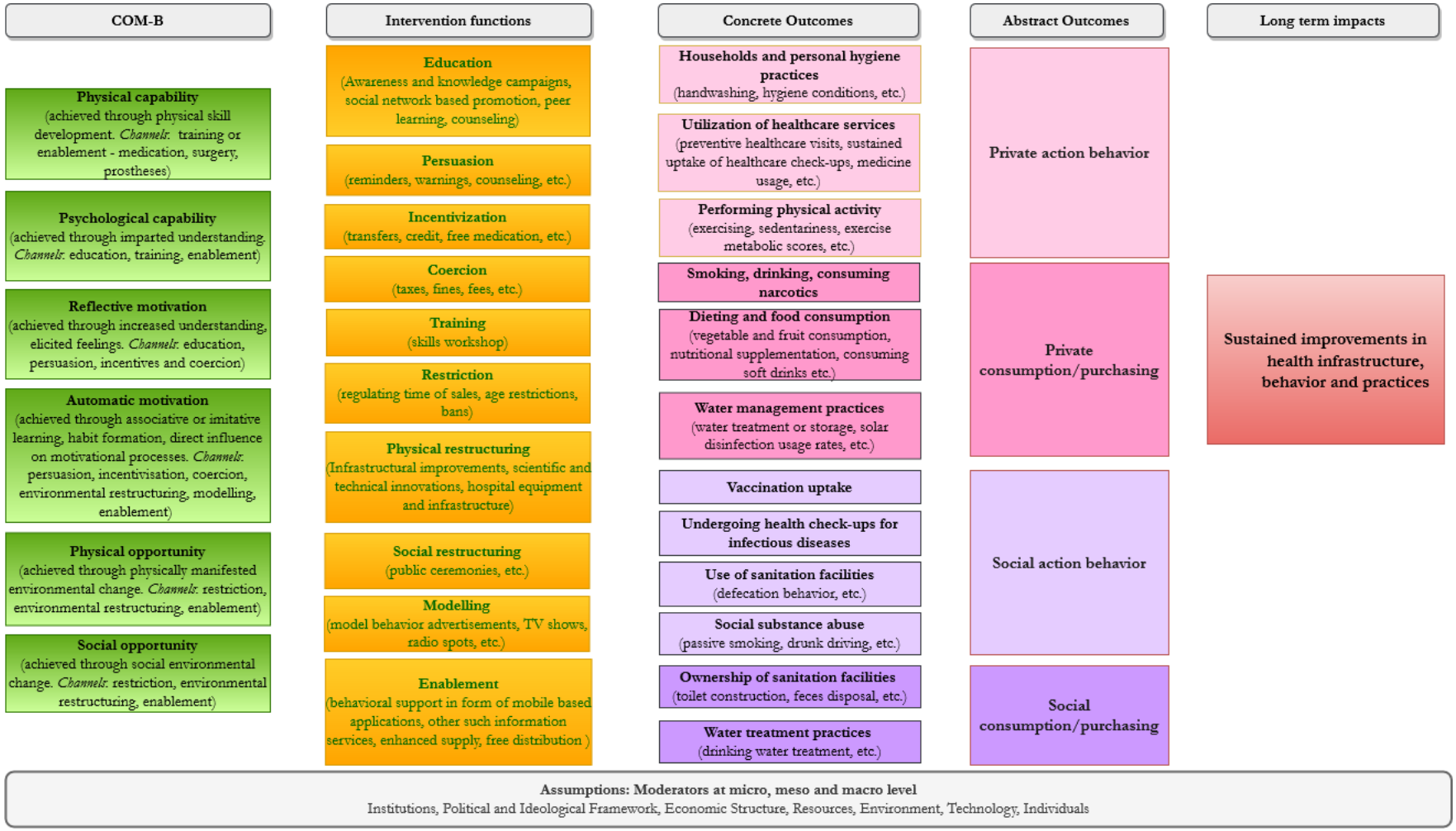
PROJECT LEAD, INCEPTION, CONCEPTION, FUNDING, AND QUALITY CONTROL	
Prof. Dr. Jyotsna Puri	(former) Independent Evaluation Unit, Green Climate Fund (GCF)
Dr. Martin Prowse	Independent Evaluation Unit, Green Climate Fund (GCF)
Neha Sharma	Climate Investment Funds (CIF)
PROJECT CONCEPTION AND IMPLEMENTATION	
Center for Evaluation and Development (C4ED)	
Prof. Dr. Markus Frölich	Quality assurance and technical support
Dr. Arne R. Weiss	Lead investigator on the project implementation and on the energy sector
Dr. Atika Pasha	Lead investigator on the public health sector
Zharas Aitmambet	Project management
Mathilda Featherston-Lardeux	Research assistance
Consultants of the Africa Centre for Systematic Reviews & Knowledge Translation, Makerere University (Africa Centre MakCHS)	
Dr. Ekwaro A. Obuku	Africa Centre MakCHS, Lead Consultant
Moses Ocan	Africa Centre MakCHS, Supporting Consultant
Robert Apunyo	Africa Centre MakCHS, Project Manager
Thomas Katairo	Africa Centre MakCHS, Research Associate
Ojiambo Kevin Ouma	Africa Centre MakCHS, Research Associate
EXTERNAL SUPPORT	
Dr. Frank Renkewitz	University of Erfurt, technical support on systematic review, meta-analysis, and detection of publication bias
John Eyers	Information specialist, support in designing search strategies for multiple databases
Martin Rudasingwa	University of Heidelberg, technical support on systematic review, specifically the health sector
THIRD-PARTY REVIEW	
Howard White	Campbell Collaboration
Ashrita Saran	Campbell Collaboration

B. DECLARATIONS OF INTEREST

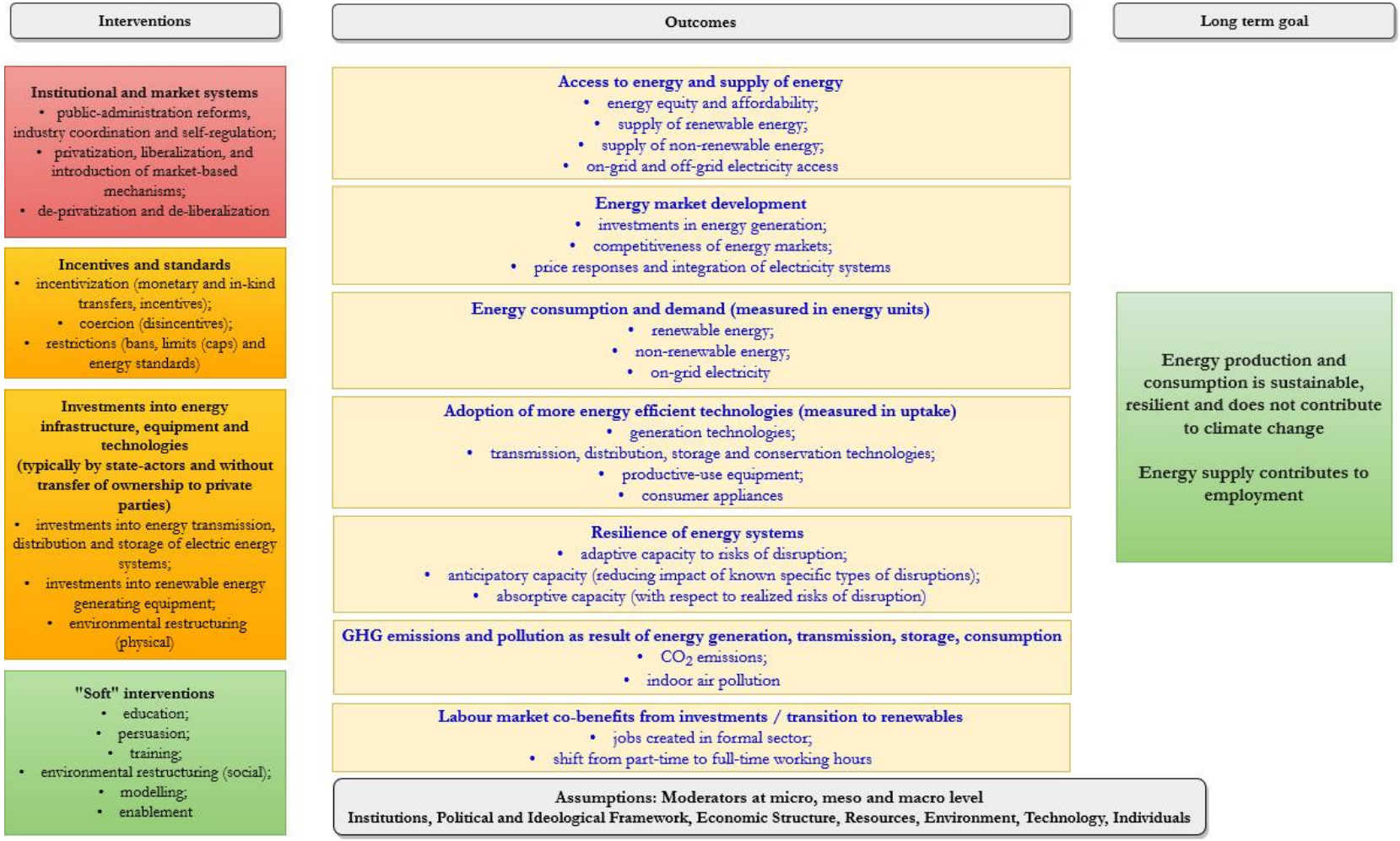
We declare no conflict of interest.

APPENDICES

APPENDIX 1. THEORY OF CHANGE FOR THE PUBLIC HEALTH SECTOR



APPENDIX 2. THEORY OF CHANGE FOR THE ENERGY SECTOR



APPENDIX 3. INCLUSION AND EXCLUSION CRITERIA FOR THE PUBLIC HEALTH SECTOR

INCLUDED	CLARIFICATORY EXAMPLES OF WHAT WILL BE CAPTURED	EXCLUDED
1) Population		
Adults (> 18 years) in LMICs, as a proxy for developing countries. Studies combining a population in LMICs and HICs, if the analysis distinguishes the two samples		Interventions targeting children (<18 years) and studies in HICs, a combination of both LMICs and HICs if effects are not reported separately
2) Interventions		
Interventions targeting the five broad areas of: *a)* *Substance abuse* *b)* *Utilization of health services* *c)* *Hygiene practices* *d)* *Nutrition and dietary habits* *e)* *Physical activity* following the behavioural framework under Michie and others (2011), as defined below: • Education (awareness and knowledge campaigns to promote frequent handwashing) • Persuasion (reminders or warnings per phone to not drink and drive or harms of smoking) • Incentivization (monetary and in-kind rewards such as free or subsidized medication, consultation, vaccines, etc.) • Coercion (price increases for alcohol or tobacco, or punishment) • Training (community skills-building workshops for hygiene practices) • Restriction (prohibiting the sales of alcohol to those under 18, permitting sales only at particular times, smoking only in specific areas in a bar, etc.) • Environmental restructuring – Physical (improvements or creation of infrastructure to facilitate access to	Within each broad intervention area: a) *Substance abuse*: i) Narcotics and stimulant drug use ii) Alcohol consumption iii) Smoking *b)* *Utilization of health services:* i) Uptake of health care-related services (participation in health care counselling, that is, for ANC or PNC, vaccination, etc.) ii) Uptake of health care checkups (HIV testing, ANC, PNC) iii) Compliance with health care service offered (use of oral rehydration solution), malarial profilax, insecticide-treated bednets, institutional delivery, vaccination, immunization, etc.) *c)* *Hygiene practices:* i) Drinking water treatment and storage ii) Use of sanitation facilities iii) Handwashing and personal hygiene *d)* *Nutrition and dietary habits:* i) Nutritional supplementation	• Within each of the five broad intervention areas: *a)* *Substance abuse:* i) Violence due to substance abuse *b)* *Utilization of health services:* i) Hospital/heath care quality improvement ii) Insurance uptake *c)* *Hygiene practices:* i) Food storage ii) Waste disposal and treatment iii) Menstrual hygiene *d)* *Nutrition and dietary habits* i) Agricultural and related food fortification (biofortification) *e)* *Physical activity:* i) Exercise among athletes • Natural interventions (those that are not within the control of humans), such as sudden climate-related shocks / natural disasters/ migration • Laboratory/clinical trial/field lab interventions targeting a hospital, clinic or laboratory (for example, quality of care, PBF (performance based financing), health professional training/ education) • Following policy changes from governmental authorities: a) (De)Regulation/ decentralization/ privatization/ simplification of procedures

INCLUDED	CLARIFICATORY EXAMPLES OF WHAT WILL BE CAPTURED	EXCLUDED
health services or use of health facilities, or reduce unhealthy practices) – Social (setting defaults, prompts) • Modelling (model behaviour advertisements, TV shows, or posters) • Enablement[27] (behavioural support for smoking cessation, such as mobile-based applications and services, or those that encourage health checkups)	ii) Dieting iii) Nutritional equity/food equity iv) Diet diversification v) Food fortification *e) Physical activity:* i) Exercise or fitness ii) Sedentariness	(deregulation and liberalization of health sector) b) Governmental change/laws or regulation (legislature, bills, or policies; coordination of government at different levels—national, subnational, etc.)
3) Outcomes		
Outcomes related to the five sectors of substance abuse, the utilization of health services, hygiene practices, nutrition and dietary practices, and physical activity will be categorized under the sections below:		
Action behaviour (actions taken by individual to improve their health status)		
Social (HIV testing, open defecation, drunk driving, passive smoking, etc.)		
Private (ANC, PNC, institutional delivery, handwashing, etc.)		
Consumption/purchasing (purchasing drugs or alcohol, nutritious food)		
Social (toilet construction, etc.)		
Private (expenditure on drugs, alcohol bought, spending on tobacco/cigarettes, etc.)		
4) Study design		
• Impact evaluation methods (experimental and quasi-experimental), Heckman selection, fixed effects with established control group • Systematic reviews		• Correlation analysis and random effects, input-output models, general-equilibrium models, and other methods not based on control groups • Time series analysis of indicators at the aggregated (macro) level • Non-systematic reviews
5) Publication date		
After 01.01.2000		Before 01.01.2000
6) Number of beneficiaries (scale)		
Results need to be representative of a large-scale intervention through two ways (Muralidharan and Niehaus, 2017).		• Interventions targeting less than 1,000 beneficiaries; single household/village/community/ firm interventions

[27] Capability beyond training and education; opportunity beyond environmental restructuring

INCLUDED	CLARIFICATORY EXAMPLES OF WHAT WILL BE CAPTURED	EXCLUDED
• Scale of intervention: at least 1,000 individual beneficiaries (automatically fulfilled if more than 1,000 treated individuals in sample); if the number of beneficiaries not given, or for radio and other media-based interventions, intervention needs to target an entire administrative area larger than villages (for example, districts, regions, etc.) • Scale of population represented: sample of treated individuals should be randomly drawn from a sampling frame of at least 1,000 treated individuals or from an administrative area larger than a village		• Interventions targeting small group of individuals (for example, a training programme of energy sales agents in a small city) • Studies that are not representative of at least a population of 1,000 treated individuals or an administrative area larger than a village
7) Language		
English		Other languages
8) Timing of data collection		
Outcomes were measured at least one year after the first full implementation of the causally identified intervention component of interest. When baseline values are used for identifying the treatment effect, then the time between the baseline and the endline needs to be at least one year		Outcomes are only measured for a period of less than one year

APPENDIX 4. INCLUSION AND EXCLUSION CRITERIA FOR THE ENERGY SECTOR

INCLUDED	CLARIFICATORY EXAMPLES OF WHAT WILL BE CAPTURED	EXCLUDED
1) Population		
Adults and adolescents, as well as social systems, in LMICs, as a proxy for developing countries; studies combining a population in LMICs and HICs, if the analysis distinguishes the two samples		Interventions targeting children (under the age of 12) and studies in HICs, combination of both LMICs and HICs, if effects are not reported separately
2) Interventions in the energy sector[28]		
Institutional and market systems *(Interventions that change the institutional structure of energy systems or markets)*		• Interventions outside the energy sector, those that do not target climate change mitigation or adaptation through energy production or usage, are typically excluded: for example, afforestation programmes (such as REDD+), labour market reforms, labour-market trainings, agronomic trials, engineering studies / trials (for example, the energy efficiency of engines or different fuels), geological, geographical studies; investments into infrastructure other than the energy infrastructure. • Economic growth is not an intervention. Studies only relevant, if economic growth is a mediator of the effects of an intervention in the energy sector on a relevant energy outcome (for example, GHG
• Public administration reforms, industry coordination, and industry self-regulation	Technical assistance, restructuring of government units, changes in management practices in the public sector; business associations, industry bodies	
• Privatization, liberalization, and introduction of market-based mechanisms	Energy and emissions trading platforms, frameworks for private sector involvement [public-private partnerships (PPPs)]	
• De-privatization and de-liberalization		
Incentives and standards		
Incentivization (monetary and in-kind transfers, incentives)	Distribution of more energy-efficient consumer appliances, subsidies, block tariffs, tax rebates, feed-in tariffs	
Coercion (disincentives)	Taxes and fees, permits, green quotas, fines	
Restrictions [bans, limits (caps) and energy standards]	Command and control, damage control, prohibition	
"Soft" interventions *(Interventions that do not change the incentives of actors)*		

[28] In the spirit of Arnott and others (2014), interventions are coded as "behavioural" or "structural". In light of cross-sectoral learnings, behavioural interventions are those that directly target the behavioural change of individuals or households and measure a behavioural outcome (see footnote in outcomes). These will be classified according to the Behaviour Change Wheel (BCW, Michi and others 2011). Structural interventions changes are those that do not, or only indirectly, lead to individual or household-level behavioural change. Behavioural interventions fall mainly in the category of "soft" interventions. The subcategories within this category are, therefore, classified according to seven of the "invention functions" of the BCW. Structural interventions fall mainly under the category of "institutional and market systems". The remaining two categories, "incentives and standards" and "investments into energy infrastructure", are mixed, with both structural interventions and behavioural interventions expected. The subcategories within "incentives and standards", as well as the subcategory, "physical environmental restructuring" within "investments into energy infrastructure", reflect the BCW intervention functions.

INCLUDED	CLARIFICATORY EXAMPLES OF WHAT WILL BE CAPTURED	EXCLUDED
• Education	Awareness and knowledge campaigns	emissions) to be causally identified • Investments and distributions of small energy-generating appliances (smaller than 200W), like solar lanterns, solar household-systems • Investments into non-renewable and nuclear
• Persuasion	Reminders or warnings	
• Training	community skills-building workshops on efficient- energy use	
• Social environmental restructuring	Social norms, peer pressure, feedback	
• (Role) modelling	Model behaviour advertisements, TV shows or posters	
• Enablement	Setting defaults, making options salient, creating options	
Investments into energy infrastructure, equipment, and technologies (typically by state actors and without transfer of ownership to private parties) and other physical environmental restructuring		
• Investments into energy transmission, distribution, and storage of electric energy systems	Batteries for storage, pumped-storage hydroelectricity	
• Investments into renewable energy-generating equipment	Dams for hydropower	
• Other physical environmental restructuring	Changes to the physical environment other than investments into infrastructure and the distribution of appliances, such as changing traffic signals	
3) Outcomes[29]		
Access to energy and supply of energy		• Implementation of a new business model in the energy sector; engineering or technological innovations other than related to energy efficiency, economic growth, biomass production, agricultural yields • Carbon sequestration through afforestation, (de)forestation
Energy equity and affordability	Energy inclusiveness, energy affordability (widening of access), energy cost reductions	
Supply of renewable energy (measured in units of energy) (traditional and next-generation)	Generation and supply of traditional renewable energy, next-generation renewable energy, including biofuels	
Supply of non-renewable energy (measured in units of energy)	Generation and supply of energy from coal, oil, gas, LPG, liquefied natural gas (LNG), kerosene, petrol, diesel, nuclear	

[29] In light of cross-sectoral learnings, behavioural outcomes, at the level of individuals and households, are coded according to the framework in health, that is, along the four categories of "action behaviour—private", "action behaviour—social", "consumption/purchasing behaviour—private", and "consumption/purchasing behaviour—social". Behavioural outcomes are mainly expected in the categories of "energy consumption and demand", "the adoption of more energy-efficient technologies", along with the sub-category, "consumer appliances".

INCLUDED	CLARIFICATORY EXAMPLES OF WHAT WILL BE CAPTURED	EXCLUDED
On-grid and off-grid electricity access	Generation and supply of electricity, electricity coverage; adoption of grid access	• Labour market outcomes that are in measured in response to an intervention targeting climate change mitigation or energy access • Disaster risk reduction measures (that only indirectly affect energy system resilience) • Time use other than for formal employment • Willingness to pay (instead of direct adoption)
Energy market development		
• Investments (measured in monetary terms, typically private) in energy generation renewable energy (traditional and next-generation), fossil fuels, nuclear energy, electricity		
• Competitiveness of energy markets (market power of energy suppliers, composition)	Number of suppliers of energy products and services, concentration indices, service quality standards, power outages, variation in voltage, use of technological innovations	
• Price responses and integration of electricity systems	Liquidity, pricing regulation and instruments, price adjustments, spatial connectivity of electricity systems, linkages within the power supply chain	
Energy consumption and demand for (measured in energy units)		
• Renewable energy (traditional and next-generation)	Consumption of and demand for traditional renewable energy, next- generation renewable energy, including biofuels	
• Non-renewable energy	Consumption of and demand for energy from coal, oil, gas, LPG, LNG, kerosene, petrol, diesel, nuclear; diesel generators (off-grid)	
• On-grid electricity	Consumption of and demand for on-grid electricity (produced by a mix of energy sources; otherwise, it is categorized under renewable or non-renewable)	
Adoption of more energy-efficient technologies (measured in uptake, not in monetary units or through demand for energy)		
• Generation technologies	Adoption of energy-efficient generation technologies	
• Transmission, distribution, storage, and conservation technologies	Adoption of energy-efficient transmission, distribution, storage, and conservation technologies	

INCLUDED	CLARIFICATORY EXAMPLES OF WHAT WILL BE CAPTURED	EXCLUDED
• Productive-use equipment	Adoption of energy-efficient technologies and equipment in manufacturing, infrastructure, services	
• Consumer appliances	Adoption of more energy-efficient consumer appliances (lighting, transportation, cooking)	
Resilience of energy systems (adaptation)		
• Adaptive capacity towards risks of disruption (for example, security of energy supply through diversification of energy sources, lower energy imports - for example, forecasting)	Energy security, reliance on energy imports, excess generating capacity, oil, gas, and LNG storage reservoirs	
• Anticipatory capacity (reducing impact of known specific types of disruptions through preparedness and planning - for example, proactive action to reduce vulnerability)	Energy use planning, peak energy use, smoothing of energy consumption, decentralization of energy systems, integrating energy resilience into systems planning (heat, power, transportation systems)	
• Absorptive capacity (with respect to realized risks of disruption - exercised during and after a disturbance)	Ability of households to cope with the energy production side blackouts, power quality, reliability of energy systems	
GHG emissions and pollution as a result of energy generation, transmission, storage, consumption		
• GHG emissions (for example, carbon capture at power plants)		
• Indoor air pollution (for example, from cookstoves)		
Labour market co-benefits (jobs creation, demand for workforce) from investments into energy infrastructure, equipment, and technologies	Employment, unemployment, number of new jobs, local-level multiplier effects on the labour market	
• Employment in the formal sector		
• Shift from part-time to full-time working hours		
• Working hours in salaried / formal employment		
4) Study design		
• Impact evaluation methods (experimental and quasi-experimental), Heckman selection, fixed effects with established control group • Systematic reviews		• Correlation analysis and random effects without a clear control group (as is often the case with the time series analysis, such as those relying on Granger causality), input-

INCLUDED	CLARIFICATORY EXAMPLES OF WHAT WILL BE CAPTURED	EXCLUDED
		output models, general-equilibrium models, simulations, and other methods not based on control groups • Qualitative studies • Non-systematic reviews • Book chapters, perspectives, protocols
5) Publication date		
After 01.01.1990		Before 01.01.1990
6) Number of beneficiaries (scale)		
Results need to be representative of a large-scale intervention through two ways (Muralidharan & Niehaus, 2017). • Scale of intervention: at least 1,000 individual beneficiaries (automatically fulfilled, if more than 1,000 treated individuals in sample); if the number of beneficiaries not given, or for radio and other media-based interventions, the intervention needs to target an entire administrative area needs to be larger than villages (for example, districts, regions, etc.) • Scale of population represented: sample of treated individuals should be randomly drawn from a sampling frame of at least 1,000 treated individuals or from an administrative area larger than a village		• Interventions targeting less than 1,000 beneficiaries; single household/ village/ community/ firm interventions • Interventions targeting a small group of individuals (for example, a training programme of energy sales agents in a small city)
7) Timing of data collection		
Outcomes were measured at least one year after the first full implementation of the causally identified intervention component of interest. When baseline values were used for identifying the treatment effect, then time between the baseline and endline needs to be at least one year.		Outcomes are only measured for a period of less than one year

APPENDIX 5. SEARCH STRATEGY HEALTH SECTOR—WEB OF SCIENCE DATABASE

SEARCH	RESULTS #	SEARCH EXPRESSION
# 1	1,428,140	(TS=("long-term" OR "long term" OR longitudinal OR "over time" OR "multiple wave*" OR ((two OR three OR four OR five OR six OR seven OR eight OR nine OR ten) NEAR/0 wave*) OR transformat* OR prolonged OR wane* OR waning OR sustain* OR unsustain* OR "not sustainable" OR ("year*-long" NEAR/0 (stud* OR data*)) OR ("month*-long" NEAR/0 (stud* OR data*)) OR ((panel OR longitudinal) NEAR/2 (study OR analysis OR data)) OR ((paradigm* OR "ground-breaking" OR "ground breaking") NEAR/2 (shift* OR chang*)) OR large-scale OR largescale OR "large scale" OR year* OR month* OR time* OR "long-period" OR "long time period" OR "over a long period" OR "long run" OR "long-run" OR "follow*up")) AND LANGUAGE: (English) Indexes=SSCI Timespan=2000-2020
# 2	474,013	(TS=(("quasi experiment*" OR quasi-experiment* OR quasiexperiment* OR "random* control* trial*" OR "random* trial*" OR RCT OR randomi* OR (matching NEAR/2 (study OR procedure OR "using" OR use* OR observable*)) OR "propensity score" OR psm OR "regression discontinuity" OR "regression kink" OR "fuzzy regression" OR "sharp regression" OR "discontinuous design" OR rdd OR "difference in difference*" OR "difference-in-difference*" OR "diff in diff" OR "diff-in-diff" OR (random* NEAR/1 (allocat* OR assign* OR select*)) OR "research synthesis" OR "fixed effect*" OR "synthetic control" OR "rapid evidence assessment*" OR "systematic literature review*" OR "systematic* review*" OR metaanaly* OR "meta analy*" OR meta-analy* OR "control* evaluation" OR "control* treatment" OR "instrumental variable*" OR (as NEAR/2 instrument) OR (heckit NEAR/2 (model* OR estimat* OR procedure OR method)) OR (heckman* NEAR/5 (sample OR selection OR model OR correction)) OR ((treatment OR intervention OR comparison OR control OR subsidy) NEAR/0 group) OR ((counterfactual OR "counter factual" OR "counter-factual" OR random*) NEAR/2 (stud* OR analysis OR experiment*)) OR ((counterfactual OR "counter factual" OR "counter-factual" OR random*) NEAR/2 (outcome*)) OR causal* OR "control group*" OR "comparison group*" OR ((control OR treatment) NEAR/0 (communit* OR village*)) OR (experiment* NEAR/1 (stud* OR analysis OR design*)) OR IV OR ITT OR ((treatment OR intervention) NEAR/2 effect*) OR "intention-to-treat" OR "intention to treat" OR ("econometric analysis") OR (impact* NEAR/1 (evaluation OR stud*)) OR "controlled before?and?after" OR "controlled before?after" OR "quasi?experimental time series" OR "interrupted time series") NOT (granger OR "kuznets curve" OR "unidirectional causality" OR "uni-directional causality" OR "bidirectional causality" OR "cointegrated equation" OR cointegration OR "panel causality" OR "co-integration" OR "causality test*" OR "wavelet coherence" OR "spatial econometric" OR nexus OR "response surface" OR "choice experiment*" OR ((root OR "cross-section* dependence" OR "cross-sectional augmented") NEAR/0 test)))) AND LANGUAGE: (English) Indexes=SSCI Timespan=2000-2020
# 3	561,959	(TS=((Africa OR Caribbean OR "West Indies" OR "Middle East" OR "Central America" OR "Pacific Islands" OR Micronesia OR Polynesia OR Melanesia) OR (Asia NOT (Japan OR Korea OR "Hong Kong" OR Hong-Kong)) OR ("South America" OR "Latin America") OR (Afghanistan OR Albania OR Algeria OR "American Samoa" OR Angola OR Argentina OR Armenia OR Armenian OR Azerbaijan OR Bangladesh OR Byelarus OR Byelorussian OR Belarus OR Belorussian OR Belorussia OR Belize OR Benin OR Bhutan OR Bolivia OR Bosnia OR Herzegovina OR Hercegovina OR Botswana OR Brazil OR Bulgaria

SEARCH	RESULTS #	SEARCH EXPRESSION
		OR "Burkina Faso" OR "Burkina Fasso" OR "Upper Volta" OR Burundi OR Urundi OR "Cabo Verde" OR Cambodia OR "Khmer Republic" OR Kampuchea OR Cameroon OR Cameroons OR Cameron OR Camerons OR "Cape Verde" OR "Central African Republic" OR Chad OR China OR Colombia OR Comoros OR "Comoro Islands" OR Comores OR Mayotte OR Congo OR Zaire OR "Costa Rica" OR "Cote d'Ivoire" OR "Côte d'Ivoire" OR "Ivory Coast" OR Cuba OR Djibouti OR "French Somaliland" OR Dominica OR "Dominican Republic" OR "East Timor" OR "East Timur" OR "Timor Leste" OR Ecuador OR Egypt OR "United Arab Republic" OR "El Salvador" OR "Equatorial Guinea" OR Eritrea OR "Eswatini" OR Ethiopia OR Fiji OR Gabon OR "Gabonese Republic" OR Gambia OR Gaza OR Georgia OR "Georgia Republic" OR "Georgian Republic" OR Ghana OR Grenada OR Guatemala OR Guinea OR Guiana OR Guyana OR "Guinea-Bissau" OR Haiti OR Honduras OR India OR Indonesia OR Iran OR Iraq OR Jamaica OR Jordan OR Kazakhstan OR Kazakh OR Kenya OR Kiribati OR Kosovo OR Kyrgyzstan OR Kirghizia OR "Kyrgyz Republic" OR Kirghiz OR Kirgizstan OR "Lao PDR" OR Laos OR Lebanon OR Lesotho OR Basutoland OR Liberia OR Libya OR Macedonia OR Madagascar OR "Malagasy Republic" OR Malaysia OR Malaya OR Malay OR Sabah OR Sarawak OR Malawi OR Maldives OR Mali OR "Marshall Islands" OR Mauritania OR Mauritius OR "Agalega Islands" OR Mexico OR Micronesia OR Moldova OR Moldovia OR Moldovian OR Mongolia OR Montenegro OR Morocco OR Ifni OR Mozambique OR Myanmar OR Myanma OR Burma OR Namibia OR Nauru OR Nepal OR "Netherlands Antilles" OR Nicaragua OR Niger OR Nigeria OR Muscat OR Pakistan OR Palestine OR Paraguay OR Peru OR Philippines OR Philipines OR Phillipines OR Phillippines OR "Papua New Guinea" OR Romania OR Rumania OR Roumania OR Russia OR Russian OR Rwanda OR Ruanda OR "Saint Lucia" OR "St Lucia" OR "St. Lucia" OR "Saint Vincent" OR "St Vincent" OR "St. Vincent" OR Grenadines OR Samoa OR "Samoan Islands" OR "Navigator Island*" OR "Sao Tome" OR "São Tomé and Principe" OR Senegal OR Serbia OR "Sierra Leone" OR "Sri Lanka" OR "Solomon Islands" OR Somalia OR Sudan OR Suriname OR Surinam OR Swaziland OR "South Africa" OR Syria OR Syrian OR Tajikistan OR Tadzhikistan OR Tadjikistan OR Tadzhik OR Tanzania OR Thailand OR Togo OR "Togolese Republic" OR Tonga OR Tunisia OR Turkey OR Turkmenistan OR Turkmen OR Tuvalu OR Uganda OR Ukraine OR Uzbekistan OR Uzbek OR Vanuatu OR "New Hebrides" OR Venezuela OR Vietnam OR "Viet Nam" OR "West Bank" OR Gaza OR Yemen OR Zambia OR Zimbabwe) OR ((developing OR "less* developed" OR "less-developed" OR "under developed" OR underdeveloped OR "middle income" OR "middle-income" OR "low* income" OR "low*-income" OR underserved OR "under served" OR deprived OR poor*) NEAR/0 (countr* OR nation OR nations OR population* OR world OR state*)) OR ((developing OR "less* developed" OR "under developed" OR underdeveloped OR "middle income" OR "low*-income" OR "low* income" OR underserved OR "under served" OR deprived OR poor*) NEAR/0 (economy OR economies)) OR (low* NEAR/0 (gdp OR gnp OR "gross domestic" OR "gross national")) OR (low NEAR/3 middle NEAR/3 countr*) OR (lmic OR lmics OR "third world" OR "lami countr*" OR "global south") OR "former soviet" OR "post-soviet" OR "commonwealth of independent states" OR "non-OECD" OR ((transition* OR cis) NEAR/0 (countr* OR state* OR economy OR economies)))) AND LANGUAGE: (English) Indexes=SSCI Timespan=2000-2020
# 4	1,530,495	(TS=(((prevent* OR health OR primary OR community OR "peer group" OR group) NEAR/2 (care OR service OR program* OR session OR educat* OR re-educat* OR reducat* OR intervention OR train* OR retrain* OR re-train* OR check* OR knowledge OR support)) OR ((physical* OR game* OR leisure* OR fitness OR wellness OR health OR care) NEAR/2 (event* OR setting* OR program* OR venue* OR site* OR center OR centre OR check OR check-up OR checkup)) OR ((prevent* OR intervention* OR campaign* OR initiative*)

SEARCH	RESULTS #	SEARCH EXPRESSION
		NEAR/2 (diabetes OR obesity OR cardiac)) OR ((acquatic OR resistance OR physical) NEAR/2 (training OR exercis* OR exert*)) OR ((lifestyle OR "life style" OR life-style) NEAR/2 (intervention* OR change OR improv* OR better* OR campaign*)) OR ((activity OR movement OR fitness) NEAR/2 (track* OR sens* OR monitor*)) OR pedometr* OR step count* OR mhealth OR (self NEAR/2 (help* OR manag* OR monitor* OR track*)) OR (walk* OR run* OR jog* OR swim* OR danc* OR garden* OR cycl* OR bicycl* OR bike* OR recreation*) OR ((food* OR diet* OR nutritio* OR nutrient*) NEAR/2 (choice OR mediterranean OR health* OR balance* OR fat OR fats OR salt* OR sugar* OR unhealthy OR therapy OR polic* OR diversi* OR balanc* OR prepar*)) OR ((beverage* OR drink* OR liquid*) NEAR/2 (sweet OR sweetened OR carbonated OR cola OR sugar OR caloric OR energy)) OR ((physical* OR game* OR leisure* OR fitness) NEAR/2 (event* OR setting* OR program* OR venue* OR site* OR center OR centre)) OR ((media OR community OR school OR family OR parent*) NEAR/2 (intervention* OR program* OR campaign* OR initiative*)) OR ((lifestyle OR "life style" OR life-style) NEAR/2 (intervention OR change OR improv* OR better*)) OR ((food NEAR/2 (ration* OR supplement* OR fortif*)) NEAR/2 (program* OR intervent* OR campaign* OR initiative*)) OR ((vitamin* OR mineral* OR iodin* OR iron OR zinc OR micronutrient* OR nutrient*) NEAR/2 (suppl* OR capsule* OR inject* OR deficiency*)) OR (food NEAR/2 label*) OR ((weight OR "weight control" OR "weight reduction") NEAR/2 (program* OR intervent* OR campaign* OR initiative*)) OR (water NEAR/2 (drink* OR provide OR provis* OR filter* OR sanitiz* OR sanitis* OR purifi* OR treat* OR guard OR manage* OR disinfect* OR steriliz* OR sterilis* OR boil* OR sedimentation OR biofilter* OR "anti-bacterial agent*" OR antimicrobial*)) OR "sodium hypochlorite" OR SODIS OR "water management" OR sanitizer OR sanitiser OR "sanitary engineering" OR ("household water treatment" NEAR/2 "safe storage") OR hwts OR "water safety plan*" OR "water supply" OR ((chemical OR heat) NEAR/2 treatment) OR ((ultraviolet OR UV) NEAR/2 (radiation OR treatment)) OR (chlorine NEAR/2 (dispenser OR filter)) OR ((sanitation OR handwash* OR "WASH" OR hwipc OR "toilet construction" OR "sewage construction" OR hygiene OR wastewater) NEAR/2 (facilit* OR station OR intervention* OR program* OR campaign* OR initiative*)) OR "pour toilet*" OR "flush toilet*" OR "pit latrine*" OR "composting toilet*" OR "on?site system*" OR "off?site system*" OR sewerage OR "septic tank*" OR "Community-Led Total Sanitation" OR CLTS OR "WASH" OR ("water sanitation" NEAR/2 hygiene) OR ((water OR sanitation) NEAR/2 hygiene) OR ("Participatory Hygiene" NEAR/2 "Sanitation Transformation") OR SARAR OR "Urban Led Total Sanitation" OR "community approach*" OR "supply side improvements" OR "hygiene promot*" OR "water closet*" OR ((hygiene OR mother OR mothers OR health) NEAR/2 club*) OR (("nicotine replacement" OR "smoking cessation" OR replacement OR anti-smoking OR antismoking OR no-smoking OR "no smoking" OR "non smoking" OR non-smoking OR nonsmoking OR drug OR anti-drug OR "substance abuse" OR alcohol) NEAR/2 (therap* OR intervention* OR program* OR initative* OR campaign* OR counsel* OR treat* OR polic*)) OR ((smoking OR tobacco OR cigar* OR bidi* OR beedi* OR hooka* OR waterpipe* OR kretek* OR shisha* OR chutta* OR dhumti* OR hookli* OR chillum* OR alcohol OR liquor* OR drug* OR marijuana OR cocaine OR crack OR heroin OR meth OR methamphetamine OR amphetamine OR narcotic OR opium OR MDMA OR LSD) NEAR/2 (restriction* OR ban* OR prevent* OR polic* OR prohibit*)) OR (nicotine NEAR/2 (patch* OR spray* OR inhaler* OR lozenge* OR gum*)) OR bupropion OR ((label* OR pack* OR packet* OR package*) NEAR/2 warn*) OR ((behavioural OR cessation) NEAR/2 (support OR aid OR therap* OR instruct*)) OR "cigarette* price" OR ((indoor OR workplace OR work-place OR office OR hospital OR employee*) NEAR/2 ((restriction* OR ban* OR polic* OR prohibit*) NEAR/2 smok*)) OR (("door to door" OR home OR clinic OR "preventive healthcare" OR monthly) NEAR/2 (visit OR checkup* OR check-up* OR test OR tests OR testing)) OR ((health OR healthcare OR "health care" OR medical) NEAR/2 (availab* OR provision*

SEARCH	RESULTS #	SEARCH EXPRESSION
		OR provid* OR promot* OR prevent* OR barrier* OR constrain* OR imped* OR facilitat* OR hinder* OR block* OR obstacle OR restrict* OR optimiz* OR optimis* OR adher* OR access* OR motivat* OR accept* OR availabl*)) OR ("insecticide treated" NEAR/2 (net* OR bednet* OR "bed net*")) OR "insurance provi*" OR "facility based delivery" OR "institutional* delivery" OR "village council meeting" OR (elimination NEAR/2 ("user fee" OR fee OR charge)) OR checkup* OR ((vaccin* OR immuniz* OR innocul* OR "antenatal care" OR ANC OR "postnatal care" OR PNC OR health OR "health care" OR healthcare OR medical) NEAR/2 (therap* OR intervention* OR program* OR initative* OR campaign* OR counsel* OR treat* OR polic* OR camp)) OR (train* NEAR/2 ("GPs" OR "service provider" OR "health worker" OR midwife)) OR ((education* OR information*) NEAR/2 (campaign* OR session* OR poster* OR leaflet* OR counseling)) OR (technical NEAR/2 information) OR "home based counselling" OR stickers OR broadcasting OR leaflets OR meetings OR "individual advice" OR "social mobilization" OR advocacy OR advocat* OR "behaviour change communication" OR "family based home health education" OR ((messag* OR SMS OR "short message service" OR email* OR e-mail* OR "electronic mail*" OR television OR tv OR televised OR radio OR newspaper OR movie OR in-store OR "in store" OR magazine* OR internet OR web OR print) NEAR/2 (campaign* OR commercial OR commercials* OR display OR displays OR retail OR store OR "point of purchase" OR "point-of-purchase" OR "point of sale" OR "point-of-sale")) OR "cash transfer" OR ((monetary OR financial) NEAR/2 incentive*) OR "tax reform" OR subsid* OR voucher* OR microcredit OR micro-credit* OR loan* OR financ* OR price* OR prices OR tax OR taxes OR taxation OR training OR "technical information" OR "capacity building" OR restrict* OR forbid* OR warning* OR ban OR bans OR banning OR prohibiti* OR ordinance OR ((mobile OR "smart phone" OR smartphone OR phone OR cellphone OR "cell phone" OR tablet* OR electronic) NEAR/2 (app OR apps OR application* OR messag*)) OR community motivation OR "Participatory Rural Appraisal" OR "Community Action Planning" OR remind* OR prompt* OR cue* OR cuing OR advertis* OR brand* OR marketing OR mass distribution campaign OR enable* OR infrastructur* OR "mass media campaign" OR "social media campaign" OR stickers OR broadcasting OR leaflet* OR promot* "social marketing" OR "role modelling" OR "technical assistance" OR "behavioural support session" OR "behavioural support" OR "self help material" OR aid)) AND LANGUAGE: (English) Indexes=SSCI Timespan=2000-2020
# 5	1,484,963	(TS=((physical* NEAR/2 (activ* OR inactiv* OR fit* OR train* OR exertion OR "effort exert*" OR "functioning score" OR condition)) OR "activ*" OR "walk" OR "walking" OR "run" OR "running" OR "jog" OR "jogging" OR "exercise" OR "fitness" OR (gym* NEAR/2 member*) OR "cycling" OR "sedentar*" OR "metaboli*" or "inactiv*" OR (lifestyle NEAR/2 (change OR health*)) OR "sport*" OR (("sedentar*" OR "sitting" OR "seat*" OR "lying" OR "reclin*" OR "recumben*" OR screen OR computer OR TV OR television) NEAR/2 time) OR "screentime" OR "self track*" OR "fitness track*" OR (keep* OR cardio* OR aerobic OR fitness) NEAR/2 (fit* OR activ* OR train*) OR "motor activity" OR "exercise" OR "nutrit*" OR "supplement*" OR "feed*" OR "nutrient" OR "intake" OR "diet*" OR "meal frequency" OR "vegetable consumption" OR "leafy" OR "fruit-based" OR "food fortification" OR "vitamin intake" OR "glucose tolerance" OR "soft drink consumption" OR "iycn knowledge" OR "appetite regulation" OR "eating" OR "energy consumption" OR "portion size" OR "ingestion" OR "*food" OR "(food OR fruit* OR vegetable* OR salt* OR fat* OR sugar*) NEAR/2 (intake* OR consum* OR eat* OR ate) OR "overweight" OR "over weight" OR "obesity" OR "overeat*" OR "over eat" OR "(waist* NEAR/2 (circumference* OR measur*)) OR "(weight OR bmi OR body mass index) NEAR/2 (gain* OR loss* OR lose* OR lost OR change*)) OR "body weight" OR "(diet* NEAR/2 (chang* OR education OR behavio*r OR pattern*)) OR (water NEAR/2 (increas* OR intake* OR consum*)) OR

SEARCH	RESULTS #	SEARCH EXPRESSION
		("drinking water" NEAR/2 (increas* OR intake* OR consum* OR decreas* OR reduction)) OR "undernutrition" OR "undernourish*" OR "under-nutrition*" OR "under-nourish" OR ((sugar* OR fizzy OR carbonated OR cola) NEAR/2 (beverage* OR drink*)) OR (liquid* NEAR/2 carbohydrate*) OR cordial* OR "hygien*" OR "sanita*" OR "wash*" OR "open defecation" OR "water treatment" OR "water disinfect*" OR "water quality" OR "handwash*" "hand-washing" OR ((toilet OR latrine) NEAR/2 usage) OR "SODIS" OR "handwashing" OR "hand disinfection" OR "hand cleansing" OR "handscrubbing" OR "chlorine test" OR (("faeces" OR "feces" OR fecal OR "faecal" OR "defecat*" OR "excrement*" OR "human waste" OR "night soil" OR "excreta") NEAR/2 ("dispos*" OR "manag*")) OR "drinking water" OR "drinkwater" OR "soap" OR "water quality" OR "water supply" OR "water contamination" OR "water storage" OR "water source" OR "point of use" OR "environmental health" OR "water access" OR "potable water" OR (infection NEAR/2 (control OR prevention OR management)) OR "cross infection" OR "waste disposal" OR "adaptive behaviour" OR "behaviour change" OR "adaption" OR "adoption" OR "adopting" OR "rejecting" OR "rejection" OR "observance" OR "conformity" OR "compliance" OR "diffusion" OR "substance abuse" OR "substance*" OR "drug abuse" OR "drug*" OR "smok*" OR "alcohol*" OR "drunk" OR "cigar*" OR "narcotic*" OR "stimulant*" OR "bidis" OR "drink*" OR ((alcohol OR drink* OR cigarette OR tobacco) NEAR/2 (consumption OR use OR abuse OR misuse* OR intoxicat* OR harmful OR excess* OR binge* OR heavy OR temperance OR abstinence OR exposure OR quit OR quitting OR reduction OR stop OR stopping)) OR "smoke free" OR "smoke-free" OR "abstinence" OR (smoking NEAR/2 (cessation OR quit OR quitting OR reduction OR stop OR stopping)) OR "tobacco smoke" OR "second hand smoke" OR "nicotine replacement" OR "demand cigarette" OR "waterpipe smoking" OR "addiction" OR "addictive behaviour" OR "tobacco" OR "nicotine" OR "temperan*" OR "drinking behaviour" OR "drink*" OR "beer" OR "wine" OR "ethanol" OR addict* OR recidivism OR "ANC" OR "ante-natal care" OR "antenatal care" OR "postnatal care" OR "post-natal care" OR "PNC" OR "institutional deliver*" OR "deliver*" OR "hospital delivery" OR "childbirth*" OR "child birth*" OR "bed net*" OR "bednet*" OR "itn use" OR "oral rehydration solution" OR (hiv NEAR/2 (test* OR learn*)) OR "check-up*" OR "checkup*" OR "check up*" OR "vaccin*" OR "immuni*" OR "deworm*" OR ((medicine* OR "medical device*" OR condom* OR "therapy session*") NEAR/2 (use OR usage OR purchas*)) OR "visitation" OR "counsel*" OR (("health care" OR healthcare OR "health service*") NEAR/2 (utilization OR utilisation OR entry)) OR inocul* OR innocul* OR "knowledge practice" OR (patient NEAR/2 (complian* OR adheren* OR dropout* OR attrition OR capacitance)) OR "treatment refusal" OR ((drug OR medication) NEAR/2 adherence) OR ltfu OR "loss* to follow-up" OR "self-directed exercise*" OR "self directed exercise*" OR "uptake" OR "take up" OR "satisfaction" OR "compliance" OR "comply" OR "complie*" OR "refus*" OR "modif*" OR "adaptation" OR "adaptive behaviour" OR "behaviour change" OR "adaption" OR "adoption" OR "adopting" OR "rejecting" OR "rejection" OR "observance" OR "conformity" OR "diffusion" OR (knowledge NEAR/2 attitudes NEAR/2 practice))) AND LANGUAGE: (English) Indexes=SSCI Timespan=2000-2020
#6	12,448	#5 AND #4 AND #3 AND #2 AND #1 Indexes=SSCI Timespan=2000-2020

APPENDIX 6. SEARCH STRATEGY ENERGY SECTOR—WEB OF SCIENCE DATABASE

SEARCH	RESULTS #	SEARCH EXPRESSION
# 1	11,886,194	TS=("long-term" OR "long term" OR longitudinal OR "over time" OR "multiple wave*" OR ((two OR three OR four OR five OR six OR seven OR eight OR nine OR ten) NEAR/0 wave*) OR transformat* OR prolonged OR wane* OR waning OR sustain* OR unsustain* OR "not sustainable" OR ("year*-long" NEAR/0 (stud* OR data*)) OR ("month*-long" NEAR/0 (stud* OR data*)) OR ((panel OR longitudinal) NEAR/2 (study OR analysis OR data)) OR ((paradigm* OR "ground-breaking" OR "ground breaking") NEAR/2 (shift* OR chang*)) OR large-scale OR largescale OR "large scale" OR year* OR month* OR time* OR "long-period" OR "long time period" OR "over a long period" OR "long run" OR "long-run" OR "follow*up") AND LANGUAGE: (English) Indexes=SSCI Timespan=1990-2020
# 2	3,008,045	TS=(("quasi experiment*" OR quasi-experiment* OR quasiexperiment* OR "random* control* trial*" OR "random* trial*" OR RCT OR randomi* OR (matching NEAR/2 (study OR procedure OR "using" OR use* OR observable*)) OR "propensity score" OR psm OR "regression discontinuity" OR "regression kink" OR "fuzzy regression" OR "sharp regression" OR "discontinuous design" OR rdd OR "difference in difference*" OR "difference-in-difference*" OR "diff in diff" OR "diff-in-diff" OR (random* NEAR/1 (allocat* OR assign* OR select*)) OR "research synthesis" OR "fixed effect*" OR "synthetic control" OR "rapid evidence assessment*" OR "systematic literature review*" OR "systematic* review*" OR metaanaly* OR "meta analy*" OR meta-analy* OR "control* evaluation" OR "control* treatment" OR "instrumental variable*" OR (as NEAR/2 instrument) OR (heckit NEAR/2 (model* OR estimat* OR procedure OR method)) OR (heckman* NEAR/5 (sample OR selection OR model OR correction)) OR ((treatment OR intervention OR comparison OR control OR subsidy) NEAR/0 group) OR ((counterfactual OR "counter factual" OR "counter-factual" OR random*) NEAR/2 (stud* OR analysis OR experiment*)) OR ((counterfactual OR "counter factual" OR "counter-factual" OR random*) NEAR/2 (outcome*)) OR causal* OR "control group*" OR "comparison group*" OR ((control OR treatment) NEAR/0 (communit* OR village*)) OR (experiment* NEAR/1 (stud* OR analysis OR design*)) OR IV OR ITT OR ((treatment OR intervention) NEAR/2 effect*) OR "intention-to-treat" OR "intention to treat" OR ("econometric analysis") OR (impact* NEAR/1 (evaluation OR stud*)) OR "controlled before?and?after" OR "controlled before?after" OR "quasi?experimental time series" OR "interrupted time series") NOT (granger OR "kuznets curve" OR "unidirectional causality" OR "uni-directional causality" OR "bidirectional causality" OR "cointegrated equation" OR cointegration OR "panel causality" OR "co-integration" OR "causality test*" OR "wavelet coherence" OR "spatial econometric" OR nexus OR "response surface" OR "choice experiment*" OR ((root OR "cross-section* dependence" OR "cross-sectional augmented") NEAR/0 test))) AND LANGUAGE: (English) Indexes=SSCI Timespan=1990-2020
# 3	2,767,320	TS=((Africa OR Caribbean OR "West Indies" OR "Middle East" OR "Central America" OR "Pacific Islands" OR Micronesia OR Polynesia OR Melanesia) OR (Asia NOT (Japan OR Korea OR "Hong Kong" OR Hong-Kong)) OR ("South America" OR "Latin America") OR (Afghanistan OR Albania OR Algeria OR "American Samoa" OR Angola OR Argentina OR Armenia OR Armenian OR Azerbaijan OR Bangladesh OR Byelarus OR Byelorussian OR Belarus OR Belorussian OR Belorussia OR Belize OR Benin OR Bhutan OR Bolivia OR Bosnia OR Herzegovina OR Hercegovina OR Botswana OR Brazil OR Bulgaria

SEARCH	RESULTS #	SEARCH EXPRESSION
		OR "Burkina Faso" OR "Burkina Fasso" OR "Upper Volta" OR Burundi OR Urundi OR "Cabo Verde" OR Cambodia OR "Khmer Republic" OR Kampuchea OR Cameroon OR Cameroons OR Cameron OR Camerons OR "Cape Verde" OR "Central African Republic" OR Chad OR China OR Colombia OR Comoros OR "Comoro Islands" OR Comores OR Mayotte OR Congo OR Zaire OR "Costa Rica" OR "Cote d'Ivoire" OR "Côte d'Ivoire" OR "Ivory Coast" OR Cuba OR Djibouti OR "French Somaliland" OR Dominica OR "Dominican Republic" OR "East Timor" OR "East Timur" OR "Timor Leste" OR Ecuador OR Egypt OR "United Arab Republic" OR "El Salvador" OR "Equatorial Guinea" OR Eritrea OR "Eswatini" OR Ethiopia OR Fiji OR Gabon OR "Gabonese Republic" OR Gambia OR Gaza OR Georgia OR "Georgia Republic" OR "Georgian Republic" OR Ghana OR Grenada OR Guatemala OR Guinea OR Guiana OR Guyana OR "Guinea-Bissau" OR Haiti OR Honduras OR India OR Indonesia OR Iran OR Iraq OR Jamaica OR Jordan OR Kazakhstan OR Kazakh OR Kenya OR Kiribati OR Kosovo OR Kyrgyzstan OR Kirghizia OR "Kyrgyz Republic" OR Kirghiz OR Kirgizstan OR "Lao PDR" OR Laos OR Lebanon OR Lesotho OR Basutoland OR Liberia OR Libya OR Macedonia OR Madagascar OR "Malagasy Republic" OR Malaysia OR Malaya OR Malay OR Sabah OR Sarawak OR Malawi OR Maldives OR Mali OR "Marshall Islands" OR Mauritania OR Mauritius OR "Agalega Islands" OR Mexico OR Micronesia OR Moldova OR Moldovia OR Moldovian OR Mongolia OR Montenegro OR Morocco OR Ifni OR Mozambique OR Myanmar OR Myanma OR Burma OR Namibia OR Nauru OR Nepal OR "Netherlands Antilles" OR Nicaragua OR Niger OR Nigeria OR Muscat OR Pakistan OR Palestine OR Paraguay OR Peru OR Philippines OR Philipines OR Phillipines OR Phillippines OR "Papua New Guinea" OR Romania OR Rumania OR Roumania OR Russia OR Russian OR Rwanda OR Ruanda OR "Saint Lucia" OR "St Lucia" OR "St. Lucia" OR "Saint Vincent" OR "St Vincent" OR "St. Vincent" OR Grenadines OR Samoa OR "Samoan Islands" OR "Navigator Island*" OR "Sao Tome" OR "São Tomé and Principe" OR Senegal OR Serbia OR "Sierra Leone" OR "Sri Lanka" OR "Solomon Islands" OR Somalia OR Sudan OR Suriname OR Surinam OR Swaziland OR "South Africa" OR Syria OR Syrian OR Tajikistan OR Tadzhikistan OR Tadjikistan OR Tadzhik OR Tanzania OR Thailand OR Togo OR "Togolese Republic" OR Tonga OR Tunisia OR Turkey OR Turkmenistan OR Turkmen OR Tuvalu OR Uganda OR Ukraine OR Uzbekistan OR Uzbek OR Vanuatu OR "New Hebrides" OR Venezuela OR Vietnam OR "Viet Nam" OR "West Bank" OR Gaza OR Yemen OR Zambia OR Zimbabwe) OR ((developing OR "less* developed" OR "less-developed" OR "under developed" OR underdeveloped OR "middle income" OR "middle-income" OR "low* income" OR "low*-income" OR underserved OR "under served" OR deprived OR poor*) NEAR/0 (countr* OR nation OR nations OR population* OR world OR state*)) OR ((developing OR "less* developed" OR "under developed" OR underdeveloped OR "middle income" OR "low*-income" OR "low* income" OR underserved OR "under served" OR deprived OR poor*) NEAR/0 (economy OR economies)) OR (low* NEAR/0 (gdp OR gnp OR "gross domestic" OR "gross national")) OR (low NEAR/3 middle NEAR/3 countr*) OR (lmic OR lmics OR "third world" OR "lami countr*" OR "global south") OR "former soviet" OR "post-soviet" OR "commonwealth of independent states" OR "non-OECD" OR ((transition* OR cis) NEAR/0 (countr* OR state* OR economy OR economies))) AND LANGUAGE: (English) Indexes=SSCI Timespan=1990-2020
# 4	13,340,573	TS=("electricity reform*" OR "electricity sector reform*" OR "institutional framework*" OR (reform* NEAR/2 (electricity OR power OR investment OR administrative OR regulatory OR institutional)) OR "independent regulator*" OR "cap-and-trade" OR "emission* allowance*" OR "emission* credit*" OR PPP OR "public-private partnership*" OR "public-private-partnership*" OR (emission* NEAR/1 (trad* OR tax* OR certificat*)) OR privatiz* OR privatis* OR "deregulation" OR "market-oriented reform*" OR "market reform*" OR "electricity dispatch reform" OR (competition NEAR/1 (electricity OR power)) OR (restructuring NEAR/3 (electricity OR power)) OR (restructur* NEAR/2

SEARCH	RESULTS #	SEARCH EXPRESSION
		(firm* OR effect* OR industry)) OR "unbundling" OR ((corporatisation OR corporatization) NEAR/2 (state-owned)) OR ("open access" NEAR/2 (transmission OR distribution) NEAR/1 (network* OR system* OR grid* OR line*)) OR ("private participation" NEAR/2 (transmission OR distribution)) OR "independent power produc*" OR IPPs OR "business association*" OR "industry body" OR "industry self-regulation" OR (("cooking stove*" OR cookstove) NEAR/1 improve*) OR (("prepaid metering" OR "graded tariff*" OR "pay-as-you-go" OR "pay as you go" OR "flexible tariff*") NEAR/1 (electricity OR power)) OR ((carbon OR fuel) NEAR/1 (tax* OR pricing)) OR (((renewable OR reusable) NEAR/1 energy) NEAR/1 subsid*) OR (("fossil-fuel subsid*" OR "fossil fuel subsid*") NEAR/1 removal*) OR (("tax credit*" OR "tax incentive*" OR credit*) NEAR/1 "renewable energy") OR "feed-in tariffs" OR "infrastructure treasury guarantees" OR "power purchase agreement*" OR PPA OR PPAs OR ((energy OR power OR electricity OR "water heating" OR heating OR solar OR wind OR hydro* OR micro-hydro* OR microhydro* OR pico-hydro* OR renewable*) NEAR/4 ("loan guarantee*" OR "credit guarantee*" OR "soft loan*" OR "concessional loan*" OR "loan program*" OR "lending facilit*" OR "project financ*" OR financ* OR "private equity" OR "investment fund*" OR insurance OR micro-insurance OR microinsurance OR co-insurance OR coinsurance OR "risk-sharing" OR "risk sharing")) OR ("market commitment" NEAR/1 advanced) OR (("micro-loan*" OR "micro loan*" OR "micro financ*" OR "microfinanc*" OR "credit guarantee fund*" OR "credit line") NEAR/3 (energy OR electricity OR power OR solar OR wind OR hydro* OR micro-hydro* OR microhydro* OR pico-hydro* OR renewable*)) OR "clean technology fund" OR (("results based" OR "results-based") NEAR/1 (financ* OR funding)) OR CDM OR "clean development mechanism" OR ((energy OR power) NEAR/1 "local financ* intermed*") OR REFINe OR (("R&D" OR "Research and Development") NEAR/1 fund*) OR "green bonds" OR (("adapt* programme" OR "development polic*" OR "sector investment") NEAR/1 loan) OR "waterfall payment mechanism*" OR ((bended OR carbon) NEAR/1 finance) OR "carbon support" OR "block pricing" OR "block tariff*" OR (price NEAR/2 (shock OR schedule OR "determination mechanism*")) OR (incentive* NEAR/1 (monetary OR financial)) OR "utility bill" OR (voucher* NEAR/2 (distribution OR status OR allocation OR non-transferable OR discount* OR recipient*)) OR "subsidy program*" OR "cross subsid*" OR "cross-subsid*" OR "subsidy payback period" OR "electricity tariff*" OR "feeder load" OR "backup tariff*" OR "tariff order" OR "tariff rationalisation" OR "tariff rationalization" OR (rebalanc* NEAR/2 price*) OR "energy efficien*" OR (emission* NEAR/1 (control* OR standard* OR target*)) OR ("end-of pipe emission*" NEAR/1 control*) OR ((renewable* OR green OR clean) NEAR/1 "portfolio standard*") OR "green quota*" OR "green certificate*" OR "renewable energy quota*" OR "renewable obligation*" OR "renewable energy source*" OR ((policy OR policies OR govern* OR legisl* OR law* OR legal) NEAR/4 (energy OR electricity OR power OR renewable OR "fossil fuel" OR nuclear OR gas OR oil OR LNG OR LPG OR biofuel OR biogas OR biomass OR firewood OR carbon OR biodiesel OR bioethanol OR gasoline OR diesel OR kerosene OR heating OR wind OR hydro* OR geothermal OR thermal)) OR "energy standard*" OR "behavioural intervention*" OR "behavioural intervention*" OR training OR (technology NEAR/2 demonstration) OR (campaign* NEAR/1 awareness) OR ((promot* OR awareness OR inform*) NEAR/2 (reusable OR green OR renewable OR solar OR wind OR hydro* OR "micro-hydro*" OR microhydro* OR "pico-hydro*" OR thermal OR geothermal OR saving OR efficien* OR sustainab*) NEAR/1 (energ* OR electricity OR power)) OR ((promot* OR awareness OR inform*) NEAR/2 (biofuel OR biogas OR "cooking stove" OR cookstove)) OR "information incentive*" OR (defaults NEAR/1 ("green electricity" OR "carbon offsets")) OR (("social norm*" OR implementation) NEAR/1 intervention*) OR "social network*" OR (peer NEAR/2 (learn* OR effect* OR influence*)) OR (interaction* NEAR/1 (social OR preference)) OR "social learning" OR (neighbor* NEAR/2 ("connection behaviour*" OR "connection behaviour*")) OR (simplif* NEAR/1 (choice OR environment)) OR disclosure OR warning* OR reminder* OR precommitment OR

SEARCH	RESULTS #	SEARCH EXPRESSION
		(inform* NEAR/1 choice*) OR sign* OR suggest* OR (information NEAR/2 nudge*) OR "respon* to nudge*" OR ((government* OR public) NEAR/1 ("tech* investment" OR research OR development OR "research and development" OR "R&D")) OR ("superconduct* elec*" NEAR/1 (elem* OR equip*)) OR ((mini OR micro OR smart) NEAR/1 (grid NEAR/1 (expan* OR improv* OR increas*))) OR ((LNG OR "liquefied natural gas") NEAR/1 infrastructur*) OR "gas pipeline*" OR (centrali* NEAR/2 "energy access") OR "grid flexibilit*" OR (decentrali* NEAR/2 (energy OR power OR electricity OR heating)) OR "dry cell battery" OR ((household* OR rural) NEAR/2 (electrification)) OR (grid* NEAR/1 (smart OR small OR mini OR micro OR electricity)) OR ((transmission OR distribution OR electricity) NEAR/1 (line* OR network* OR station* OR substation OR infrastructur*)) OR ((electricity OR grid) NEAR/1 expansion) OR "infrastructur* investment*" OR (electric* NEAR/1 infrastructure) OR "grid-electricity generation" OR "energy conservation" OR "fuel conversion" OR "utilization of reusable energy" OR electrif* OR (investment NEAR/1 (energy OR "renewable energy" OR "reusable energy" OR electric* OR "clean energy")) OR "solar power" OR (wind NEAR/1 (turbine* OR power OR wheel OR mill* OR energy)) OR (solar NEAR/1 (plant OR "power plant" OR panel* OR battery OR "water heaters" OR power OR grid* OR electricity)) OR "photovoltaic tech*" OR "offshore wind turbine*" OR ((hydro* OR "micro-hydro*" OR microhydro* OR "pico-hydro*") NEAR/1 (batter* OR fuel OR power OR electric* OR energy)) OR (water NEAR/1 (power OR electric* OR energy)) OR (tidal NEAR/1 (power OR energy OR electric* OR "stem gener*" OR barrage)) OR (wave NEAR/1 (energy OR power)) OR "wave energy converter" OR (damless NEAR/1 (power OR energy)) OR (marin* NEAR/1 (power OR electric* OR energy)) OR (("liquid-dominat*" OR reservoir* OR electric* OR power OR energy OR "heat pump*" OR plant* OR system* OR generat*) NEAR/1 (thermal OR geothermal)) OR biofuel* OR biodiesel OR bioethanol* OR ((algae* OR algal*) NEAR/1 (fuel OR oil)) OR "fuel from waste" OR "fuel cell" OR "carbon-neutral fuel" OR ((first-gener* OR "first-gener*" OR "second-gener*" OR "second-gener*") NEAR/1 biofuel*) OR "fuel switch" OR "low carbon tech*" OR ((small OR micro) NEAR/1 dams) OR ((thermal OR photovoltaic*) NEAR/1 solar NEAR/1 (power OR energy)) OR CSP OR "concentrated solar power" OR "home energy system*" OR "solar home system*" OR "carbon capture and storage" OR CCS OR "solar technolog*" OR (biogas NEAR/1 (stove* OR domestic OR digester* OR program*)) OR (solar NEAR/1 (microgrid* OR grid* OR electric* OR energy)) OR "hydroelectric dam*" OR "hydropower dam*" OR "hydropower" OR "hydroelectric") AND LANGUAGE: (English) Indexes=SSCI Timespan=1990-2020
# 5	2,272,129	TS=(((energy OR electricity OR power OR ((solar OR thermal OR geothermal OR wind OR wave OR hydro* OR "micro-hydro*" OR microhydro* OR "pico-hydro*" OR nuclear) NEAR/1 (power OR energy OR electric*))) NEAR/2 (production OR reliabilit* OR supply OR provision OR penetration OR generation OR expansion OR consumption OR use OR access)) OR ((biofuel OR biogas OR biomass OR biodiesel OR firewood OR bioethanol OR gasoline OR diesel OR kerosene OR "fossil fuel*" OR fuel OR coal OR "natural gas" OR "shale gas" OR LNG OR LPG) NEAR/2 (consumption OR use OR access)) OR (levelized NEAR/2 "cost of energy") OR ("avoided cost" NEAR/2 (energy OR "new generation")) OR (access NEAR/2 ((reusable OR modern OR clean OR renewable) NEAR/1 energy)) OR (price NEAR/2 electricity) OR "non electrified household*" OR "non-electrified household*" OR "residential energy consumption" OR ((energy OR fuel) NEAR/2 (expenditure* OR expense* OR cost*)) OR (access NEAR/2 (electricity OR power OR energy OR biogas OR solar OR thermal OR geothermal OR wind)) OR ((transit* OR switch* OR shift* OR "phasing in") NEAR/2 ("clean energy" OR "renewable energy")) OR "from coal to gas" OR (access NEAR/2 ((clean OR renewable OR reusable OR green) NEAR/2 energy)) OR (access* NEAR/2 energy) OR ((renewable OR reusable) NEAR/2 energy NEAR/2 (supply OR provision)) OR "solar home system*" OR "solar

SEARCH	RESULTS #	SEARCH EXPRESSION
		technolog*" OR ((((solar OR thermal OR geothermal OR wind OR wave OR hydro* OR "micro-hydro*" OR microhydro* OR "pico-hydro*") NEAR/1 (power OR energy OR electric*)) OR power OR heating OR electricity) NEAR/2 ("generation capacity")) OR (baseload OR "base load") OR (capacity NEAR/1 installed) OR "household electricity" OR "residential electricity" OR "generation plant*" OR "electrified grid points" OR ((effect* OR impact OR household* OR rural OR "grid-based" OR "grid based") NEAR/3 electrification) OR "electrification rate" OR "household* with* electricity" OR "non-electrified household*" OR "electrified household*" OR (investment* NEAR/2 (energy OR power OR electricity OR solar OR wind OR hydro* OR "micro-hydro*" OR microhydro* OR "pico-hydro*" OR renewable* OR thermal OR geothermal OR biofuel OR biogas OR biodiesel OR bioethanol OR nuclear OR "fossil fuel*" OR gas OR LNG OR LPG OR liquefied)) OR "industrial economics" OR ((outage OR blackout) NEAR/2 (electricity OR power OR planned OR unplanned)) OR "day* without electricity" OR ("electricity service" NEAR/2 (reliability OR quality)) OR "independent power producer*" OR (separation NEAR/2 (generation OR transmission OR distribution)) OR (access NEAR/2 ((transmission OR distribution) NEAR/2 (network* OR line* OR grid))) OR "market concentration" OR (voltage NEAR/1 variation*) OR ("technological innovation*" NEAR/2 (energy OR power OR electricity)) OR "forward capacity auction" OR "marginal price" OR "price elasticity" OR "cost covering price" OR ((competitive OR discriminatory OR regulation* OR instrument*) NEAR/2 pricing) OR ("cross subsid* " OR "cross-subsid*") OR (pric* NEAR/2 (distort* OR adjustment* OR mechanism*)) OR "grid extension" OR "connection decision" OR "decision to connect" OR "choose to connect" OR "choice to connect" OR (connect* NEAR/2 ("electrical grid*" OR grid*)) OR "connection rate" OR "connection behaviour" OR "connection behaviour" OR "grid-connected" OR (connection NEAR/2 (price* OR fee* OR cost*)) OR ((discount OR status OR electricity) NEAR/2 connection) OR "metered individually" OR "electrical connectivity" OR ((electricity OR power) NEAR/3 unconnected) OR "consumption response" OR ((renewable OR reusable) NEAR/2 (energy OR electricity OR "energy service*") NEAR/2 (demand OR "peak demand" OR usage OR use OR consumption OR switch*)) OR ((energy OR biofuel OR biogas OR biomass OR biodiesel OR bioethanol OR gasoline OR firewood OR "traditional fuel*") NEAR/2 (demand OR consumption OR usage OR use OR switch*)) OR "off-grid solar electricity" OR "digester owner" OR ((waste NEAR/2 conver*) NEAR/2 (energy OR power OR electricity)) OR "waste-to-energy" OR "waste to energy" OR ((anaerobic OR biogas OR biofuel) NEAR/2 (digest* OR plant)) OR ((demand OR consumption OR usage OR use OR replac* OR switch*) NEAR/2 (diesel OR kerosene OR petrol* OR coal OR charcoal OR gas OR LNG OR LPG OR oil OR nuclear OR "fossil fuel*" OR fuel)) OR "fossil fuel fired electricity generation" OR "captive power" OR "kerosene displacement" OR "kerosene lamp" OR "kerosene lantern" OR ((electricity OR power) NEAR/2 (demand OR consumption OR usage OR use OR switch* OR connect*)) OR (("energy efficien*" OR "energy-efficien*" OR "energy-saving" OR "energy sav*") NEAR/2 technolog*) OR ((fuel OR "conventional fuel" OR "fuel usage" OR "fuel use") NEAR/3 (efficienc* OR inefficienc* OR productivity OR saving*)) OR “technical efficienc*" OR "technical inefficienc*" OR "sectoral efficienc*" OR "sectoral inefficienc*" OR ((efficienc* OR inefficienc* OR productivity OR loss* OR saving*) NEAR/3 ("generation plant*" OR "electricity generation" OR "generation segment")) OR (((power OR electricity) NEAR/2 (transmission OR distribution)) NEAR/3 (efficienc* OR inefficienc* OR productivity)) OR ((transmission OR distribution) NEAR/2 loss*) OR (("energy efficien*" OR "energy inefficienc*" OR "energy saving") NEAR/2 (building* OR cit* OR industr*)) OR (("industrial consum*" OR "industrial demand") NEAR/2 (electricity OR power)) OR (fuel NEAR/10 (efficienc* OR inefficienc* OR econom*)) OR "miles per gallon" OR "miles/gallon" OR "mpg" OR "liters per 100 kilometers" OR "liters per 100 km" OR "litres per 100 kilometers" OR "litres per 100 km"OR "l/100km" OR "negative pric*" OR "A-rated appliance*" OR "incandescent bulb*" OR "incandescent lightbulb*" OR "energy efficient bulb*" OR "incandescent lightbulb*" OR

SEARCH	RESULTS #	SEARCH EXPRESSION
		((generation OR nameplate OR installed OR "peak load") NEAR/2 capacity) OR (storage NEAR/2 regasification) OR (energy NEAR/2 (security OR resilien* OR relian* OR dependen*)) OR (import* NEAR/2 (energy OR oil OR gas OR LNG)) OR "peak load" OR "peak deficit" OR "peak demand" OR "peak energy use" OR ((sustainab* OR planning OR smooth*) NEAR/2 ((energy OR power OR electricity) NEAR/2 (use OR usage OR consumption OR demand OR system*))) OR "backup power" OR "backup electricity" OR "grid supply-demand imbalances" OR ((shortage OR deficit) NEAR/2 (energy OR power OR electricity)) OR "low emission transport" OR ((car OR vehicle) NEAR/2 (green OR "carbon neutral" OR electric OR sharing)) OR ("low emission" NEAR/2 ("energy supply" OR provision)) OR (reduc* NEAR/2 (GHG OR "greenhouse gas" OR "greenhouse-gas" OR "greenhousegas" OR "carbon intens*" OR fuel OR "carbon dioxide" OR "household carbon dioxide" OR CO2 OR Nox OR SOx OR CH_4)) OR decarbonization OR decarbonisation OR (avoid* NEAR/2 (GHG OR "greenhouse gas*" OR "greenhouse-gas*" OR "greenhousegas*" OR emission*)) OR ("carbon capture" NEAR/2 (sequestration OR storage)) OR "low carbon" OR "low-carbon" OR (decarbon* NEAR/2 (electricity OR energy)) OR "low-carbon" OR (("carbon" OR "carbon-dioxide" OR CO2) NEAR/1 emission*) OR (("carbon" OR "GHG" OR "emission*") NEAR/2 avoid*) OR "household air pollution" OR "respiratory disease symptom*" OR "indoor air pollution" OR "indoor air quality" OR "overnight pm concentration" OR "employment rate*" OR "unemployment rate*" OR "employment effect*" OR "unemployment effect*" OR "employment impact*" OR "unemployment impact*" OR "formal employment" OR "unemployment" OR "waged-employment" OR "waged employment" OR (("excess supply" OR surplus OR supply OR outcome OR increase OR reduction) NEAR/2 (labor OR labour)) OR "working-age" OR (employ* OR unemploy* NEAR/2 ("men" OR "women" OR "female" OR "male" OR "worker*" OR "formal*" OR structure)) OR "market work" OR (propensity NEAR/5 work) OR ((day* OR hour* OR time OR transition) NEAR/3 (work* OR employment OR unemployment)) OR "work* behaviour*" OR "work* behaviour*") AND LANGUAGE: (English) Indexes=SSCI Timespan=1990-2020
#6	2,595,557	TS=((renewable OR (("liquid-dominat*" OR reservoir* OR electric* OR power OR energy OR "heat pump*" OR plant* OR system* OR generat*) NEAR/1 (thermal OR geothermal)) OR " thermal insulation" OR solar OR energy OR electricity OR "water heating" OR "home heating" OR on-grid OR off-grid OR "wind turbine*" OR ("hydropower" OR "hydro-power" OR "hydroelectric*" OR "hydro-electric*" OR "micro-hydro*" OR microhydro* OR "pico-hydro*") OR biofuel OR biogas OR biomass OR biodiesel OR bioethanol OR gasoline OR diesel OR kerosene OR "cooking stove*" OR cookstove* OR firewood OR "fossil fuel*" OR fuel OR coal OR "natural gas" OR "shale gas" OR LNG OR LPG OR electrif* OR grid OR "micro-grid*" OR "micro grid*" OR microgrid* OR ((clean OR reusable OR green OR sustainable OR wind OR traditional OR modern OR plant OR generation OR transmission OR distribution OR dispatch OR network OR line OR storage OR sector OR industry OR demand OR supply OR access OR regulat* OR framework) NEAR/1 power) OR "power market*") NOT (crop* OR "energy intake" OR "gas exchange" OR "agroforestr*" OR ((biodiesel OR bioethanol OR biogas OR biomethane OR hydrogen OR biomass* OR ethanol*) NEAR/1 (production OR purification OR yield)) OR soil* OR seed* OR ((palm OR cooking OR methanol OR vegetable) NEAR/1 oil) OR "biomass burning" OR "fatty acid*" OR "molar ratio" OR "transition state*" OR (reaction NEAR/0 (time OR temperature)) OR "energy surface*" OR "activation energy" OR "catalyst concentration" OR "specie*" OR "kinetic*" OR "pyrolysis" OR "sediment*" OR "energy balance" OR "plasma" OR "spectromet*" OR "physiological*" OR "fertiliz*" OR "neutron*" OR "thermodynamic*")) AND LANGUAGE: (English) Indexes=SSCI Timespan=1990-2020

SEARCH	RESULTS #	SEARCH EXPRESSION
#7	2,046	#6 AND #5 AND #4 AND #3 AND #2 AND #1 Indexes=SSCI Timespan=1990-2020

APPENDIX 7. STUDIES INCLUDED IN THE META-ANALYSIS (PUBLIC HEALTH)

Utilization of health services

#	AUTHORS	TITLE	YEAR
1	Akresh R; De Walque D; Kazianga H	Alternative Cash Transfer Delivery Mechanisms: Impacts on Routine Preventative Health Clinic Visits in Burkina Faso	2012
2	Andrade M V; Chein F; Souza L R D; Puig-Junoy J	Income Transfer Policies and the Impacts on the Immunization of Children: The Bolsa Familia Program	2012
3	Banerjee A V; Duflo E; Glennerster R; Kothari D	Improving Immunisation Coverage in Rural India: Clustered Randomised Controlled Evaluation of Immunisation Campaigns with and without Incentives	2010
4	Boudreaux C; Chanthala P; Lindelow M	Assessing the Elimination of User Fees for Delivery Services in Laos	2014
5	Cahyadi N; Hanna R; Olken B A; Prima R A; Satriawan E; Syamsulhakim E	Cumulative Impacts of Conditional Cash Transfer Programs: Experimental Evidence from Indonesia	2018
6	Carvalho N; Thacker N; Gupta S S; Salomon J A	More Evidence on the Impact of India's Conditional Cash Transfer Program, Janani Suraksha Yojana: Quasi-experimental Evaluation of the Effects on Childhood Immunization and OtherR and Child Health Outcomes	2014
7	Choulagai B P; Onta S; Subedi N; Bhatta D N; Shrestha B; Petzold M; Krettek A	A Cluster-randomized Evaluation of an Intervention to Increase Skilled Birth Attendant Utilization in Mid- and Far-western Nepal	2017
8	Corbett E L; Makamure B; Cheung Y B; Dauya E; Matambo R; Bandason T; Munyati S S; Mason P R; Butterworth A E; Hayes R J	HIV Incidence during a Cluster-randomized Trial of Two Strategies Providing Voluntary Counselling and Testing at the Workplace, Zimbabwe	2007
9	Daniels J; Komárek A; Makusha T; Van Heerden A; Gray G; Chingono A; Richter L	Effects of a Community Intervention on HIV Prevention Behaviours among Men who Experienced Childhood Sexual or Physical Abuse in Four African Settings: Findings from NIMH Project Accept (HPTN 043)	2014
10	Edmond K M; Foshanji A I; Naziri M; Higgins-Steele A;;Burke J M; Strobel N; Farewar F	Conditional Cash Transfers to Improve Use of Health Facilities by Mothers and Newborns in Conflict Affected Countries, a Prospective Population Based Intervention Study from Afghanistan	2019
11	Ezeanolue E E; Obiefune M C; Ezeanolue C O; Ehiri J E; Osuji A; Ogidi A G; Ogedegbe G	Effect of a Congregation-based Intervention on Uptake of HIV Testing and Linkage to Care in Pregnant Women in Nigeria (Baby Shower): A Cluster-Randomised Trial	2015
12	Geldsetzer P; Mboggo E; Larson E; Lema I A; Magesa L; Machumi L; Bärnighausen T	Community Health Workers to Improve Uptake of Maternal Health care Services: A Cluster-Randomized Pragmatic trial in Dar es Salaam, Tanzania	2019
13	Hemminki E; Long Q; Zhang W H; Wu Z; Raven J; Tao F; Tang S	Impact of Financial and Educational Interventions on Maternity Care: Results of Cluster-Randomized Trials in Rural China, CHIMACA	2013

#	Authors	Title	Year
14	Kirkwood B R; Manu A; ten Asbroek A H; Soremekun S; Weobong B; Gyan T; Hill Z	Effect of the Newhints Home-visits Intervention on Neonatal Mortality Rate and Care Practices in Ghana: A Cluster-Randomised Controlled Trial	2013
15	Lamichhane P; Sharma A; Mahal A	Impact Evaluation of Free Delivery Care on Maternal Health Service Utilisation and Neonatal Health in Nepal	2017
16	Lee H; Park S J; Ndombi G O, Nam E W	Community-based Maternal and Child Health Project on 4+ Antenatal Care in the Democratic Republic of Congo: A Difference-in-differences Analysis	2019
17	Midhet F; Becker S	Impact of Community-based Interventions on Maternal and Neonatal Health Indicators: Results from a Community Randomized Trial in Rural Balochistan, Pakistan	2010
18	Njuguna J	Impact of a Conditional Cash Transfer Program Aimed at Promoting Maternal and Child Health Services in Kakamega County, Kenya	2019
19	Odeny T A; Hughes J P; Bukusi E A; Akama E; Geng E H; Holmes K K; McClell R S	Text Messaging for Maternal and Infant Retention in Prevention of Mother-to-child HIV Transmission Services: A Pragmatic Stepped-wedge Cluster-randomized Trial in Kenya	2019
20	Parvez S M; Azad R; Rahman M; Unicomb L; Ram P K; Naser A M; Luby S P	Achieving Optimal Technology and Behavioural Uptake of Single and Combined Interventions of Water, Sanitation Hygiene and Nutrition, in an Efficacy Trial (WASH Benefits) in Rural Bangladesh	2018
21	Powell-Jackson T; Hanson K	Financial Incentives for Maternal Health: Impact of a National Programme in Nepal	2012
22	Quayyum Z; Khan M N U; Quayyum T; Nasreen H E; Chowdhury M; Ensor T	Can Community Level Interventions Have an Impact on Equity and Utilization of Maternal Health Care —Evidence from Rural Bangladesh	2013
23	Stoller N E; Gebre T; Ayele B; Zerihun M; Assefa Y; Habte D; Emerson P M	Efficacy of Latrine Promotion on Emergence of Infection with Ocular Chlamydia Trachomatis after Mass Antibiotic treatment: A Cluster-randomized Trial	2011
24	Tiruneh G T; Zemichael N F; Betemariam W A; Karim A M	Effectiveness of Participatory Community Solutions Strategy on Improving Household and Provider Health Care Behaviours and Practices: A Mixed-method Evaluation	2020
25	Wang P C; Mwango A; Moberley S; Brockman B J; Connor A L; Kalesha-Masumbu P; McCarthy E	A Cluster-Randomised Trial on the Impact of Integrating Early Infant HIV Diagnosis with the Expanded Programme on Immunization on Immunization and HIV Testing Rates in Rural Health Facilities in Southern Zambia	2015

Hygiene practices

#	Authors	Title	Year
1	Arnold B; Arana B; Mäusezahl D; Hubbard A; Colford Jr J. M	Evaluation of a Pre-existing, 3-year Household Water Treatment and Handwashing Intervention in Rural Guatemala	2009
2	Biran A; Schmidt W P; Varadharajan K S; Rajaraman D; Kumar R; Greenland K; Gopalan B; Aunger R; Curtis V	Effect of a Behaviour change Intervention on Handwashing with Soap in India (SuperAmma): A Cluster-randomised Trial	2014

#	Authors	Title	Year
3	Briceño B; Coville A; Gertler P; Martinez S	Are There Synergies from Combining Hygiene and Sanitation Promotion Campaigns: Evidence from a Large-scale Cluster-randomized Trial in Rural Tanzania	2017
4	Cameron L; Shah M; Olivia S	Impact Evaluation of a Large-scale Rural Sanitation Project in Indonesia	2013
5	Chankova S; Hatt L; Musange S	A Community-based Approach to Promote Household Water Treatment in Rwanda	2012
6	Crocker J; Abodoo E; Asamani D; Domapielle W; Gyapong B; Bartram J	Impact Evaluation of Training Natural Leaders during a Community-Led Total Sanitation Intervention: A Cluster-Randomized Field Trial in Ghana	2016
7	Crocker J; Geremew A; Atalie F; Yetie M; Bartram J	Teachers and Sanitation Promotion: An Assessment of Community-Led Total Sanitation in Ethiopia	2016
8	Freeman M C; Majorin F; Boisson S; Routray P; Torondel B; Clasen T	The Impact of a Rural Sanitation Programme on Safe Disposal of Child Faeces: A Cluster-Randomised Trial in Odisha, India	2016
9	Galiani S; Gertler P; Orsola-Vidal A	Promoting Handwashing Behaviour in Peru: The Effect of Large-scale Mass media and Community Level Interventions	2012
10	Guiteras R; Levinsohn J; Mobarak A M	Encouraging Sanitation Investment in the Developing World: A Cluster-randomized Trial	2015
11	Harter M; Contzen N; Inauen J	The Role of Social Identification for Achieving an Open-defecation Free Environment: A Cluster-randomized, Controlled Trial of Community-Led Total Sanitation in Ghana	2019
12	Huda T M N; Unicomb L; Johnston R B; Halder A K; Sharker M A. Y; Luby S P	Interim Evaluation of a Large Scale Sanitation, Hygiene and Water Improvement Programme on Childhood Diarrhea and Respiratory Disease in Rural Bangladesh	2012
13	Luby S P; Agboatwalla M; Bowen A; Kenah E; Sharker Y; Hoekstra R M	Difficulties in Maintaining Improved Handwashing Behaviour, Karachi, Pakistan	2009
14	Parvez S M; Azad R; Rahman M; Unicomb L; Ram P K; Naser A M; Luby S P	Achieving Optimal Technology and Behavioural Uptake of Single and Combined Interventions of Water, Sanitation Hygiene and Nutrition, in an Efficacy Trial (WASH Benefits) in Rural Bangladesh	2018
15	Patil S; Arnold B; Salvatore A; Briceno B; Colford J M Jr; Gertler P J	A Randomized, Controlled Study of a Rural Sanitation Behaviour Change Program in Madhya Pradesh, India	2013
16	Stoller N E; Gebre T; Ayele B; Zerihun M; Assefa Y; Habte D; Emerson P M	Efficacy of Latrine Promotion on Emergence of Infection with Ocular Chlamydia Trachomatis after Mass Antibiotic Treatment: a Cluster-randomized Trial	2011
17	Tian X; Yan L; Zhao G; Wang L; Cheng Y; Lu Y; Southerland J	Evaluation of a Multi-layered Health Promotion Approach in Rural China	2019
18	Wichaidit W; Biswas S; Begum F; Yeasmin F; Nizame F A; Najnin N; Ram P K	Effectiveness of a Large-scale Handwashing Promotion Intervention on Handwashing Behaviour in Dhaka, Bangladesh	2019

APPENDIX 8. STUDIES INCLUDED IN META-ANALYSIS (ENERGY)

Emissions

#	AUTHORS	TITLE	YEAR
1	Dong F; Dai Y; Zhang S; Zhang X; Long R	Can a Carbon Emission Trading Scheme Generate the Porter Effect? Evidence from Pilot Areas in China	2019
2	Dong Z Q; Wang H; Wang S X; Wang L H	The Validity of Carbon Emission Trading Policies: Evidence from a Quasi-natural Experiment in China	2020
3	Gao Y; Li M; Xue J; Liu Y	Evaluation of Effectiveness of China's Carbon Emissions Trading Scheme in Carbon Mitigation	2020
4	Hu Y; Cheng H	Displacement Efficiency of Alternative energy and Trans-provincial Imported Electricity in China	2017
5	Qi S; Cheng S; Cui J	Environmental and Economic Effects of China's Carbon Market Pilots: Empirical Evidence Based on a DiD Model	2021
6	Shen J; Tang P; Zeng H	Does China's Carbon Emission Trading Reduce Carbon Emissions? Evidence from Listed Firms	2020
7	Tang K; Zhou Y; Liang X; Zhou D	The Effectiveness and Heterogeneity of Carbon Emissions Trading Scheme in China	2020
8	Wang Q; Gao; Dai S	Effect of the Emissions Trading Scheme on CO_2 Abatement in China	2019
9	Wang L; Liu C; Yang X	Research on Carbon Emission Reduction Effect of China's Carbon Trading Pilot—Based on Different in Different Method	2020
10	Wen Y; Hu P; Li J; Liu Q; Shi L; Ewing J; Ma Z	Does China's Carbon Emissions Trading Scheme Really Work? A Case Study of the Hubei Pilot	2020
11	Yi L; Bai N; Yang L; Li Z; Wang F	Evaluation on the Effectiveness of China's Pilot Carbon Market Policy	2020
12	Zhang H; Duan M; Deng Z	Have China's Pilot Emissions Trading Schemes Promoted Carbon Emission Reductions? —The Evidence from industrial Subsectors at the Provincial Level	2019
13	Zhang W; Li J; Li G; Guo S	Emission Reduction Effect and Carbon Market Efficiency of Carbon Emissions Trading Policy in China	2020
14	Zhang Y; Li S; Luo T; Gao J	The Effect of Emission Trading Policy on Carbon Emission Reduction: Evidence from an Integrated Study of Pilot Regions in China	2020
15	Zhang Y; Zhang J	Estimating the Impacts of Emissions Trading Scheme on Low-carbon Development	2019

#	Authors	Title	Year
16	Zhang H; Zhang R; Li G; Li W; Choi Y	Has China's Emission Trading System Achieved the Development of a Low-Carbon Economy in High-Emission Industrial Subsectors?	2020

Electrification

#	Authors	Title	Year
1	Akpandjar G; Kitchens C	From Darkness to Light: The Effect of Electrification in Ghana, 2000–2010	2017
2	Chakravorty U; Emerick K; Ravago M L	Lighting up the Last Mile: The Benefits and Costs of Extending Electricity to the Rural Poor	2016
3	Dasso R; Fernandez F	The Effects of Electrification on Employment In Rural Peru	2015
4	Dinkelman T	The Effects of Rural Electrification on Employment: New Evidence from South Africa	2011
5	Grogan L; Sadanand A	Rural Electrification and Employment in Poor Countries: Evidence from Nicaragua	2013
6	Grogan L	Household Electrification, Fertility, and Employment: Evidence from Hydroelectric Dam Construction in Colombia	2016
7	Grogan L	Time Use Impacts of Rural Electrification: Longitudinal Evidence from Guatemala	2018
8	Lee K; Miguel E; Wolfram C	Experimental Evidence on the Economics of Rural Electrification	2020
9	Lenz L; Munyehirwe A; Peters J; Sievert M	Does Large-Scale Infrastructure Investment Alleviate Poverty? Impacts of Rwanda's Electricity Access Roll-Out Program	2017
10	Lipscomb M; Mobarak A M; Barham T	Development Effects of Electrification: Evidence from the Topographic Placement of Hydropower Plants in Brazil	2013
11	Litzow E L; Pattanayak S K; Thinley T	Returns to Rural Electrification: Evidence from Bhutan	2019
12	Samad H A; Zhang F	Benefits Of Electrification and the Role of Reliability: Evidence from India	2016
13	Samad H A; Zhang F	Heterogeneous Effects of Rural Electrification Evidence from Bangladesh	2017
14	Samad H A; Zhang F	Electrification and Women's Empowerment: Evidence from Rural India	2019
15	Tagliapietra S; Occhiali G; Nano E; Kalcik R	The Impact of Electrification on Labour Market Outcomes in Nigeria	2020

APPENDIX 9. TABLES DEFINING TERMS USED IN BEHAVIOURAL CHANGE WHEEL

COM-B	Components	Theoretical Domains
CAPABILITY	Psychological capability	Knowledge
		Cognitive & interpersonal skills
		Memory, attention, & decision processes
		Behaviour regulation
Do we have the knowledge, skills, and abilities required to engage in a particular behaviour?	Physical capability	Physical strength & skills
OPPORTUNITY	Physical opportunity	Opportunities provided by the **surrounding environment** (for example, time, location, resource)
What external factors make the execution of a particular behaviour possible?	Social opportunity	Opportunities as a result of **social factors** (for example, cultural norms, social cues)
MOTIVATION	Reflective	Reflective processes (for example, making plans, evaluating things that have already happened)
		Social / Professional role & identity
		Drivers & goals
		Belief about capabilities & consequences
		Optimism
Internal processes that influence our decision-making and behaviours	Automatic	Automatic processes (for example, desires, impulses, inhibitions)
		Reinforcement
		Emotion

Function	Definition	Cases
Education	Increasing understanding or awareness (Not only to inspire a particular behaviour, but also to provide knowledge about competing behaviours)	Central to many adaptation interventions, for example, improving the resilience of agricultural production techniques or of housing in flood-prone locations
Persuasion	Using communication to induce positive or negative feelings, or stimulate action	Reminders or warnings via phone or other ICTs; positive or negative feelings induced to stimulate action; can be used to increase the uptake and use of climate information portals and systems
Incentivization	Creating expectation of reward	Rewarding energy efficiency measures through a financial incentive structure

FUNCTION	DEFINITION	CASES
Coercion	Creating expectation of punishment or cost (as opposed to Incentivization)	Higher road taxes for older, inefficient vehicles
Training	Imparting skills to encourage the behaviour of activity being trained	Central to adaptation interventions, including the use of climate information systems, climate-smart agriculture and standard operating procedures for disasters
Restriction	Using rules to reduce the opportunity to engage in the target behaviour (or increase the target behaviour by reducing the opportunity to engage in competing behaviours)	Restricted use of chlorofluorocarbons - production inputs that not only contribute to the thinning of the ozone layer, but also to climate change
Environmental restructuring	Changing the physical or social context to encourage or discourage behavioural change	Physical: Use of the smart design of cities to change everyday behaviour, especially on transport and energy use Social: Engage civic pride to protect natural resources at risk from climate impacts
Modelling	Providing a model example for people to aspire to or imitate	Use of lead farmers when implementing climate-resilient agricultural activities
Enablement	Increasing means/reducing barriers to increase capability (beyond education or training) or opportunity (beyond environmental restructuring)	Providing access to on-grid electricity in rural areas
Communications / Marketing	Using print, electronic, telephonic, or broadcast media	Mass or social media campaigns
Guidelines	Creating documents that recommend or mandate practice, including all changes to service provision	Producing and disseminating new mandatory agricultural practices or protocols
Fiscal	Using the tax system to reduce or increase financial costs	Increasing duty or increasing taxes on fossil fuels
Regulation	Establishing rules or principles of behaviour or practice	Creating mandatory tax increases for the refusal of clean-energy source installation
Legislation	Making or changing laws	Establishing mandatory limits on household energy-usage quantities with hiked tax rates and cut-off threshold
Environmental /Social Planning	Designing and/or controlling the physical or social environment	Re-design street lanes with sectioned-off bike lanes in addition to carpool lanes; limit single-driver lanes
Service Provision	Delivering a service	Provide a collective agricultural produce transportation option from the rural area to market cities

APPENDIX 10. ONLINE SUPPLEMENTAL MATERIALS

List of online supplemental materials

1. Interactive EGM for the public health sector
2. Interactive EGM for the energy sector
3. Online appendix A: Data extraction form for public health sector EGM
4. Online appendix B: Data extraction form for energy sector EGM
5. Online appendix C: Data extraction form for meta-analysis in public health sector
6. Online appendix D: Data extraction form for meta-analysis in energy sector
7. Online appendix E: Calculations of Cohen's d
8. Online appendix F: Quality assessment procedure and criteria for impact evaluations
9. Online appendix G: Public health sector studies screened (21477 studies, in RIS format)
10. Online appendix H: Energy sector studies screened (7092 studies, in RIS format)
11. Online appendix I: Additional graphs with descriptive statistics for public health sector EGMs
12. Online appendix J: Risk of bias assessment report for public health sector
13. Online appendix K: Risk of bias assessment report for energy sector
14. Online appendix L: Public health sector meta-analysis: forest plots
15. Online appendix M: Public health sector meta-analysis: funnel plots
16. Online appendix N: Bias by intervention function in public health sector
17. Online appendix O: Consumption / purchasing behaviour with largely private benefits without Parvez (2018) by intervention function
18. Online appendix P: Public health sector meta-analysis: tables

Independent Evaluation Unit
Green Climate Fund
175 Art center-daero, Yeonsu-gu
Incheon 22004, Republic of Korea
Tel. (+82) 032-458-6450
ieu@gcfund.org
https://ieu.greenclimate.fund

GREEN CLIMATE FUND | Independent Evaluation Unit

CIF CLIMATE INVESTMENT FUNDS