Ecole Doctorale des Sciences Economiques, Juridiques, Politiques et de gestion
Centre d'Etudes et de Recherche sur le Développement International (CERDI)

Université Clermont Auvergne, CNRS, IRD, CERDI, F-63000 Clermont-Ferrand, France

THREE ESSAYS ON SOCIAL SAFETY NETS IN DEVELOPING COUNTRIES

Thèse présentée et soutenue publiquement le 21 Octobre 2019
Pour l'obtention du titre de Docteur en Sciences Economiques

par
Jules GAZEAUD

sous la direction de
Mme. Catherine ARAUJO et M. Vianney DEQUIEDT

Membres du Jury

Catherine ARAUJO	Chargée de Recherche, CNRS	Directrice
Simone BERTOLI	Professeur, UCA	Président
Vianney DEQUIEDT	Professeur, UCA	Directeur
Douglas GOLLIN	Professeur, Oxford University	Rapporteur
Flore GUBERT	Directrice de Recherche, IRD	Suffragante
Elisabeth SADOULET	Professeur, UC Berkeley	Rapportrice

Remerciements

En débutant cette thèse à l'automne 2015, je ne savais pas vraiment dans quoi je me lançais. Près de quatre ans plus tard, c'est avec beaucoup de bonheur que je l'achève. On compare souvent la thèse à un exercice solitaire, long et éprouvant. Si je souscris volontiers aux deux derniers adjectifs, je suis heureux de pouvoir dire aujourd'hui que cette aventure a été au moins tout autant passionnante et épanouissante. Plus important encore, la préparation de cette thèse a été tout sauf un exercice solitaire. Je suis conscient de la chance qui a été la mienne d'avoir été aussi bien entouré pendant cette période. Famille, collègues et amis votre soutien a été décisif pour mener à bien ce projet. Je souhaite ici tous vous remercier. Cette thèse vous est dédiée.

Sans hésiter, mes premiers mots vont à mes directeurs de thèse. Merci Catherine et Vianney pour votre confiance, en particulier lorsque l'occasion de participer à l'évaluation d'impact aux Comores s'est présentée, puis lorsque je devais sans cesse prolonger mes séjours sur le terrain à cause des retards. Je suis extrêmement reconnaissant des libertés que vous m'avez accordé lorsque tout allait bien, et de votre soutien sans faille dans les moments plus difficiles. Vos encouragements et votre bienveillance m'ont permis de tirer le meilleur de cette expérience. Pour tout cela, je vous suis extrêmement reconnaissant. Merci également à Simone Bertoli, Douglas Gollin, Flore Gubert et Elisabeth Sadoulet d'avoir accepté de faire partie de mon jury.

Merci à mes co-auteurs Eric Mvukiyehe, Olivier Sterck et Victor Stéphane pour avoir contribué directement à mes travaux de recherche. Mention spéciale à Eric pour m'avoir appris à être une personne moins stressée et que *a good dissertation is a finished dissertation*, et à Olivier pour le séjour de recherche à Oxford et les joies des matchs de foot sous la neige. Plus sérieusement, merci à tous les deux de m'avoir fait autant confiance et d'être toujours restés disponibles et de bon conseil. Les séjours aux Comores et à Oxford m'ont énormément appris sur le développement en général et la recherche en particulier.

L'espace me manque pour remercier individuellement toutes les personnes qui ont compté sur le plan professionnel. Je voudrais néanmoins tout particulièrement m'adresser à Mme Doulfat et Bahtine El Maarouf. Merci pour votre accueil et votre générosité, notamment lors de mes premiers pas aux Comores.

J'ai tant appris à travers notre collaboration et nos discussions. Merci aussi à Simone Bertoli pour avoir été autant disponible pour discuter de ma recherche. La rigueur et la qualité de vos commentaires m'ont beaucoup inspiré, aussi je suis ravi que vous présidiez mon jury. Merci enfin à Pascal Jaupart dont l'aide, les conseils et les encouragements pour le job market ont sans doute fait la différence.

Le Chapitre 2 est la partie émergée de l'évaluation d'impact aux Comores et d'une collaboration fructueuse avec la Banque Mondiale, le FADC (Fonds d'Appui au Développement Communautaire) et l'INSEED (Institut National des Etudes Economiques et Démographiques). Je souhaite ici remercier toutes les personnes qui ont permis de mener à bien ce projet. Rachel Ravelosoa et Andrea Vermehren de la Banque Mondiale. Anrifouddine Ahmed, Zabah Ali, Salim Amri, Abdoul-Latuf Abdallah, Tidjara Djoumoi, Mohamed Soighir du FADC. Abbas Azali, Ahmed Djoumoi, Ounais Hamidou, Ali Issa, Armia Pidjani, Djaffar Soudjah de l'INSEED.

Une de mes plus belles expériences lors de cette thèse a sans aucun doute été la collecte des données du Chapitre 2. Pendant l'année et demie passée aux Comores, j'ai d'abord pu concevoir le plan d'échantillonnage et les questionnaires, pour ensuite recruter, former et accompagner enquêteurs et chefs d'équipes dans les ménages sélectionnés. Ce travail de longue haleine m'a permis de stimuler mon regard sur le processus de production des données, et de mieux réfléchir aux hypothèses sur lesquelles reposent les travaux de micro-économie appliquée. Je voudrais ici remercier les enquêteurs qui ont travaillé dur, souvent dans des contextes extrêmement difficiles, pour mener à bien ce projet. Merci Abdallah, Achirafi, Ahmed, Ajmal, Ali, Amir, Anfane, Anicha, Anzimati, Assimine, Badria, Baraka, Barouf, Bendjadid, Ben Djaloud, Binti, Chaimat, Chery, Djaima, Djaoide, Djoueria, Echat, El Chakour, Fahadine, Faissoil, Fatima, Hadidja, Hindata, Houssam, Ibrahim, Kachfa, Kamaldine, Karima, Kiwamiddine, Koulthoum, Maissara, Mahmoud, Malik, Mariama, Massoudi, Miftahou, Mikitadi, Minihadji, Missirna, Moinaecha, Monpapa, Mounir, Moustabchir, Mohamed, Moinaecha, Nayer, Nazmati, Nourdati, Roukia, Said, Salmata, Samir, Soilda, Soumaya, Soilihi. Merci tout particulièrement à Emile et Quentin pour le renfort très précieux lors de l'enquête finale, et à Bendjadid, Houssam, Kiwamiddine, Mohamed et Mounir pour avoir été de formidables assistants de recherche lors de l'enquête qualitative. Merci enfin aux ménages enquêtés pour leur gentillesse et pour avoir pris le temps de répondre à nos trop nombreuses questions. Cette recherche n'aurait pas été possible sans votre collaboration.

Que ce soit aux Comores, à Clermont-Ferrand ou à Oxford, j'ai toujours pu bénéficier d'un environnement stimulant et de conditions de travail épanouissantes. Merci à l'ensemble des membres du CERDI, du FADC, de l'INSEED et de RSC/ODID pour leur accueil. Merci tout particulièrement à Olivier Santoni pour son aide avec les données géo-référencées. Par ailleurs, merci à l'Université Clermont Auvergne, à l'école doctorale SEJPG, au CERDI et au CNRS pour le soutien financier.

En parlant de soutien, comment ne pas conclure ces remerciements en ayant une pensée pour toutes les personnes qui, au travers de pauses café ou déjeuner, de soirées posées ou exaltées, de soccers le mardi ou le jeudi, de sorties à vélo ou à ski, d'expéditions dauphins ou baleines, m'ont permis de m'aérer l'esprit et de déconnecter de la recherche. Merci en particulier à Claire, Clément, Hugues et Pauline pour tous les bons moments passés ensembles. Vous êtes des amis irremplaçables. Merci enfin à ma famille pour m'avoir toujours soutenu et sans qui rien de tout cela n'aurait été possible. Vous êtes le socle parfait pour tracer mon propre chemin.

Résumé

Cette thèse fournit trois essais sur le *design* et l'évaluation des filets sociaux de sécurité. Le Chapitre 1 contribue à la littérature sur les performances des méthodes de ciblage en général et sur le *Proxy Means Testing* en particulier. En utilisant une enquête expérimentale en Tanzanie, ce chapitre cherche à mesurer si les performances de ciblage du *Proxy Means Testing* sont biaisées lorsque les données de consommation sont sujettes à des erreurs de mesure non-aléatoires. Les résultats indiquent que les performances du *Proxy Means Testing* sont assez vulnérables aux erreurs de mesure non-aléatoires quand l'objectif est de cibler les ménages pauvres dans l'absolu, mais qu'elles restent en grande partie non affectées quand l'objectif est de cibler une part fixe de la population. Le Chapitre 2 étudie l'impact sur la migration d'un programme argent-contre-travail aux Comores. Ce programme a alloué de manière aléatoire à des ménages pauvres des transfers monétaires en échange de leur participation dans des travaux publics. En utilisant des données de première main, ce chapitre montre que le programme a augmenté la migration vers Mayotte – l'île Française voisine et plus riche. Entre 2016 et 2018, les ménages traités ont reçu jusqu'à 320USD et, par conséquent, étaient trois points de pourcentage plus susceptibles d'avoir un membre du ménage qui migre à Mayotte (une hausse statistiquement significative de 38 pourcent comparé au groupe de contrôle). Ce résultat semble être expliqué par la réduction des contraintes de liquidité et de risque à la migration. Le Chapitre 3 explore les effets productifs des programmes argent-contre-travail dans le contexte du *Productive Safety Net Project* en Ethiopie. Avec plus de 8 millions de bénéficiaires, le *Productive Safety Net Project* est parmis les plus grands programmes de filets sociaux d'Afrique. Il est aussi souvent considéré comme le programme d'adaptation au changement climatique le plus large d'Afrique avec ses activités concentrées sur l'amélioration des terres et des mesures de conservation des sols et des eaux. Des estimations en différence-en-différence couvrant toute l'Ethiopie sur la période 2000-2013 ne montrent aucune évidence pour supporter que les travaux publics ont eu des impacts mesurables sur la productivité agricole et sur la résilience aux chocs climatiques.

Summary

This thesis provides three empirical essays on the design and evaluation of social safety nets. Chapter 1 adds to the literature on the performances of targeting methods in general and Proxy Means Testing in particular. Using a unique survey experiment conducted in Tanzania, it investigates whether and to what degree Proxy Means Testing targeting performances are biased when household consumption data are subject to non-random errors. The results indicate that Proxy Means Testing performances are quite vulnerable to non-random errors when the objective is to target absolutely poor households, but remain largely unaffected when the objective is to target a fixed share of the population. Chapter 2 studies the impact on migration of a cash-for-work program in Comoros that randomly offered poor households cash transfers in exchange for their participation in public works projects. Using first-hand data, this chapter shows that the program increased migration to Mayotte – the neighboring and richer French Island. Between 2016 and 2018, treated households received up to US$320 in cash and, as a result, were three percentage points more likely to have a household member migrating to Mayotte (a statistically significant 38 percent increase relative to the control group). This result appears to be driven by the alleviation of liquidity and risk constraints to migration. Chapter 3 explores the productive effects of cash-for-work programs in the context of the Productive Safety Net Project in Ethiopia. With more than 8 million beneficiaries, the Productive Safety Net Project is among the largest safety net programs in Africa. It is also often considered as Africa's largest climate change adaptation program due to its focus on activities such as land improvements and soil and water conservation measures. This chapter relies on satellite and geo-referenced data to evaluate the effects of these activities and overcome the lack of household data. Difference-in-differences estimates covering whole Ethiopia over the 2000-2013 period show no evidence to support that public works had measurable impacts on agricultural productivity and resilience to climate shocks.

Contents

Introduction **1**

1 Proxy Means Testing Vulnerability to Measurement Errors? **9**
1.1 Introduction . . . 9
1.2 Measurement Errors and PMT Performances . . . 12
1.2.1 Consumption measurement errors . . . 12
1.2.2 The impact of non-random error in the dependent variable on parameter estimates . . . 15
1.3 Survey Experiment . . . 16
1.3.1 Sample . . . 16
1.3.2 Experimental design . . . 17
1.3.3 Data . . . 18
1.4 Empirical Strategy . . . 19
1.4.1 Identification strategy . . . 19
1.4.2 Estimation procedure and construction of the outcomes of interest . . . 20
1.5 Results . . . 23
1.5.1 PMT estimates . . . 23
1.5.2 PMT predictive performances . . . 23
1.5.3 PMT targeting performances . . . 28
1.6 Conclusion . . . 34

2 Cash Transfers and Migration: Experimental Evidence from Comoros **38**
2.1 Introduction . . . 38
2.2 Conceptual framework . . . 41
2.2.1 The liquidity channel . . . 43
2.2.2 Opportunity cost channel . . . 44
2.2.3 Credit constraint channel . . . 45
2.2.4 Risk-aversion channel . . . 47
2.3 Background of the cash-for-work program . . . 51
2.3.1 Context . . . 51
2.3.2 The Comoros Social Safety Net Program (SSNP) . . . 54
2.4 Experimental design and data . . . 55

2.4.1 Empirical strategy . . . 55
2.4.2 Data . . . 57
2.5 Results . . . 61
2.5.1 Program take-up . . . 61
2.5.2 Impact on migration . . . 62
2.5.3 Threats to our interpretation . . . 64
2.6 Channels . . . 65
2.6.1 Liquidity channel . . . 66
2.6.2 Opportunity cost channel . . . 67
2.6.3 Credit constraint channel . . . 69
2.6.4 Risk-aversion channel . . . 71
2.7 Conclusion . . . 73

3 The (lack of) Value of Public Works: Evidence from Ethiopia 78
3.1 Introduction . . . 78
3.2 Background . . . 82
3.3 Data . . . 83
3.3.1 Crop Production . . . 84
3.3.2 Treatment variables . . . 86
3.4 Empirical analysis . . . 87
3.5 Results . . . 93
3.5.1 Treatment effects on crop productivity . . . 93
3.5.2 Robustness checks . . . 96
3.6 Conclusion . . . 98

Conclusion 103

A Appendix to Chapter 1 105

B Appendix to Chapter 2 127
B1 Mathematical Appendix . . . 131
B2 Sub-analysis outlined in the PAP . . . 136

C Appendix to Chapter 3 141

List of Figures

1.1 Comparing distributions of PMT scores by survey design 26
1.2 Correlation between PMT score's percentile predicted by the benchmark PMT formula (formula 8) and the seven other formulas . . 31

2.1 Outcomes of the benchmark model as a function of the wage differential $w_d - w_o$ and of initial savings s_0 43
2.2 Effect of a cash transfer through the liquidity channel 44
2.3 Effect of a conditionality 46
2.4 Decision to finance migration using credit 48
2.5 Decision to migrate if migration is risky 50
2.6 Migration route to Mayotte 52
2.7 Timeline diagram 58
2.8 Liquidity channel 67
2.9 Treatment effect over time 70
2.10 Risk-aversion channel 73

3.1 Woredas covered by the PSNP 87
3.2 NDVI trends in treatment and control woredas 89
3.3 Propensity scores distribution by treatment groups 92
3.4 Treatment effects over time 95

A1 Comparing distributions of consumption by survey design . . . 106

C1 Treatment effects by elevation deciles 142

List of Tables

1.1 PMT Regressions 25
1.2 Predictive Performances 27
1.3 Targeting Performances, $1.25 Poverty Line 30
1.4 Targeting Performances, 30% Poverty Threshold 32
1.5 Interaction of PMT formula and select household characteristics . 33

2.1 Poverty rates in treated villages 54
2.2 Household characteristics at baseline 60
2.3 Treatment effects on labor market outcomes 61
2.4 Treatment effects on migration to Mayotte 63
2.5 Other migration patterns 68
2.6 Summary statistics on project workers and migrants 69
2.7 Timing of cash transfers and migration 71

3.1 Correlation between NDVI and survey-based agricultural output 85
3.2 Pre-treatment trends 90
3.3 Pre-treatment characteristics by sub-samples 93
3.4 Impacts on agricultural output 94
3.5 Triple difference 96

A1 Summary statistics 107
A2 Description of variables 108
A3 Balance Table 109
A4 Survey experiment consumption modules 110
A5 PMT Regressions: food consumption only 111
A6 PMT Regressions: consumption per adult equivalent 112
A7 PMT Regressions: extended list of covariates 113
A8 PMT Regressions: no stepwise procedure 114
A9 Predictive Performances: food consumption only 115
A10 Predictive Performances: consumption per adult equivalent . . . 116
A11 Predictive Performances: extended list of covariates in PMT . . . 117
A12 Predictive Performances: no stepwise procedure in PMT 118

A13 Targeting Performances ($1.25 Poverty Line): food consumption only . . . 119
A14 Targeting Performances ($1.25 Poverty Line): consumption per adult equivalent . . . 120
A15 Targeting Performances ($1.25 Poverty Line): extended list of covariates in PMT . . . 121
A16 Targeting Performances ($1.25 Poverty Line): no stepwise procedure in PMT . . . 122
A17 Targeting Performances (30% Poverty Threshold): food consumption only . . . 123
A18 Targeting Performances (30% Poverty Threshold): consumption per adult equivalent . . . 124
A19 Targeting Performances (30% Poverty Threshold): extended list of covariates in PMT . . . 125
A20 Targeting Performances (30% Poverty Threshold): no stepwise procedure in PMT . . . 126

B1 IV estimates . . . 127
B2 Differential attrition test . . . 127
B3 Attrition reasons . . . 128
B4 Indirect treatment effects . . . 128
B5 Liquidity channel . . . 129
B6 Risk aversion channel . . . 130
B7 Treatment effects on remittances . . . 137
B8 Treatment effects by migration reasons . . . 138
B9 Summary statistics on remittances sent by migration reason . . . 138
B10 Heteregeneous Effects . . . 139
B11 Endogenous stratification . . . 140

C1 Satellite and survey-based cultivated area . . . 142
C2 Determinants of the treatment . . . 143
C3 Treatment and control Woredas by regions . . . 144
C4 Main results with standardized outcome . . . 144
C5 Main results with extended controls . . . 144
C6 Main results with no restriction to the common support region . . . 145
C7 Impacts on land conservation . . . 145
C8 Impacts on migration . . . 146

Introduction

In August 2009, Kiwamiddine Chibaco was on vacation in his birthplace of Comoni, a poor village lost in the south-east of Anjouan, Comoros. At that time, he was both the pride and hope of his family, being the first to graduate from high school. The future seemed bright: Kiwamiddine was about to start a college curriculum in Madagascar, in a few years he would become a doctor, later he would return to Comoros to take care of his relatives. The problems started when Ali Chibaco, a successful farmer residing in the neighboring French island of Mayotte, died from a tragic accident in the ocean. He was Kiwamiddine's eldest brother and protector. Anything that Kiwamiddine would need for his studies, the brother would take care of it. In the years following the accident, Kiwamiddine tried very hard to keep up with his plan. His mother also redoubled her efforts, going back and forth to sell coconuts, tomatoes, and lettuce. But with nine other children to feed and no one left to support them, there never seemed to be enough money. Kiwamiddine eventually managed to go to Madagascar, though with limited financial support from his relatives. In Madagascar, he suffered from recurrent deprivation and enormous stress. He endured odd jobs. He repeated a year. He was forced to marry a Malagasy wife. Ultimately, he had to return to Comoros before completing his studies. When I met Kiwamiddine, in summer 2018, he conceded that he had lost his dream of becoming a doctor. But he strongly believed that without the accident of his brother he would be one today.

The story of Kiwamiddine Chibaco is unfortunately emblematic of the existence of many people living in the developing world, where a bad break can have long-term consequences. In their book *Poor Economics*, Abhijit Banerjee and Esther Duflo compare poor individuals to "*barefoot hedge-fund managers*" (Banerjee and Duflo, 2011). Of course, the poor handle infinitesimally smaller amounts of money than hedge-fund managers. But because they are exposed to large shocks and are generally liable for the majority of their losses, Duflo and Banerjee argue that they are actually dealing with a level of risks that hedge-fund managers never face. To reduce their exposure to risk and its consequences, they have developed various risk-management and risk-coping strategies. These include, for example, the participation to risk-sharing networks (Townsend, 1994; Udry, 1994; Dercon et al., 2006), domestic (Lucas and Stark, 1985; Rosenzweig and Stark, 1989; Gröger and Zylberberg, 2016) and international migration (Gubert, 2002; Yang and Choi, 2007), or the diversification of income sources (Barrett et al., 2001; De Janvry and Sadoulet, 2001). However, vulnerability remains high and these strategies can be harmful as they often lead individuals to forgo profitable investment opportunities or to use costly instruments to escape forced solidarity (Baland et al., 2011; Jakiela and Ozier, 2015; Boltz et al., 2019). This could curb overall development, as in the case of low-return investments in agriculture that hinder productivity gains (Rosenzweig and Binswanger, 1993) and therefore industrialization (Gollin et al., 2002).

Against this background, social safety nets have gradually emerged as one of the most, if not the most popular public policy to fight poverty and vulnerability in developing countries (Grosh et al., 2008; Del Ninno and Mills, 2015). According to the latest World Bank estimates, around 2.5 billion people are covered by safety net programs, and an average of US$ 106 per citizen is spent annually on these programs in low and middle-income countries (World Bank, 2018). Social safety nets (also sometimes called social assistance programs) are usually defined as all the noncontributory transfer programs designed and implemented by governments, international organizations, and nongovernmental organizations. They include instruments such as school feeding programs, pensions, child grants, and public works. Their distinctive feature compared to other components of the social protection system like insurance is their noncontributory nature, that is beneficiaries do not have to contribute financially to receive the transfers. Although there is a growing interest in universal transfers such as basic income (Banerjee et al., 2019), social safety nets are generally targeted towards the poor and vulnerable population.

A large evidence base has shown that social safety nets can achieve significant welfare gains (Case and Deaton, 1998; Fiszbein and Schady, 2009; Subbarao et al., 2012; Banerjee et al., 2015; Haushofer and Shapiro, 2016; Bandiera et al., 2017). But many unresolved questions remain, especially regarding the many details conditioning the success of safety net programs (Duflo, 2017). This thesis contributes to this literature by providing three original essays on the design and evaluation of social safety nets. In particular, it pays attention to aspects of programs that are still largely unexplored, and hopefully can contribute to the design of more informed and effective public policies.

Chapter 1 *"Proxy Means Testing Vulnerability to Measurement Errors?"* adds to the literature on the performances of targeting methods in general and Proxy Means Testing (PMT) in particular. Among researchers and development practitioners there has been much debate about the efficacy of PMT – a popular device to target social safety nets to poor households. I provide empirical evidence on one largely ignored aspect of PMT targeting, namely its vulnerability to non-random measurement errors in survey-based consumption data. While PMT usually assumes random measurement errors in consumption, this assumption has been challenged by recent literature. According to the typical textbook on the impact of measurement errors, this would lead to biased estimates. However, the magnitude of the bias and its implications on PMT accuracy are not clear.

I leverage a unique survey experiment conducted in Tanzania to examine whether and to what degree targeting performances are biased when the household consumption data used in PMT models are subject to non-random errors. The survey experiment randomly assigned eight different designs of consumption module to more than 4,000 households in Tanzania. One resource intensive design is believed to approximate a gold standard for consumption estimates. My empirical strategy compares the performances of PMT relying on these gold standard consumption data with those of PMT using more error-prone consumption data. The results show that non-random errors affect the targeting of absolutely poor households more than it affects the targeting of relatively poor households. In other words, a bias in consumption data leads to a bias in estimates of absolute poverty (and in the ability to identify absolutely poor households) but leaves the ranking of households largely unchanged.

Overall, this chapter contributes to the ongoing debate on the methods to target poor households. The results indicate that PMT performances are quite vulnerable to non-random errors when the objective is to target absolutely poor households, but remain largely unaffected when the objective is to target a fixed share of the population.

Chapter 2 "*Cash Transfers and Migration: Experimental Evidence from Comoros*" (joint with Eric Mvukiyehe and Olivier Sterck) studies the effects of cash transfers on migration. Given the widespread promotion of cash transfers to foster development, understanding how they affect migration is crucial, not only for academics but also for policy-makers who have preferences over migration outcomes. Using experimental data, we study the impact on international migration of a cash-for-work intervention targeted at very poor households in Comoros.

To guide the analysis, we model the decision to migrate and identify four channels through which a cash transfer intervention could affect migration. First, cash transfers relax the budget constraint and can therefore facilitate the migration of households facing a liquidity constraint (liquidity channel). Second, cash transfers that are conditional on remaining in the origin country increase the opportunity cost of migrating and can therefore reduce migration (opportunity-cost channel). Third, cash transfers can facilitate access to credit and thereby increase migration of credit constrained households as soon as they are selected to benefit from cash transfers (credit-constraint channel). Finally, as migration is a risky investment, cash transfers can encourage the migration of individuals whose preferences are characterized by decreasing absolute risk aversion (DARA) while restraining those of individuals whose preferences are characterized by increasing absolute risk aversion (IARA) (risk-aversion channel).

In the empirical analysis, we show that cash-for-work opportunities increased migration to Mayotte – the neighboring and richer French Island. Between 2016 and 2018, treated households received up to US$320 in cash and as a result were three percentage points more likely to have a household member migrating to Mayotte (a statistically significant 38 percent increase relative to the control group). We find suggestive evidence that the liquidity and the risk-aversion channels drive the results. In contrast, the opportunity-cost and the credit-constraint channels seem irrelevant in this study.

These findings confirm that many households do not migrate because of binding liquidity constraints. It also adds to a nascent literature showing that risk is an important deterrent of migration decisions. Our findings suggest that social safety nets can ease risk bearing and thereby risky migrations.

Chapter 3 "*The (lack of) Value of Public Works: Evidence from Ethiopia*" (joint with Victor Stéphane) explores the productive effects of cash-for-work programs. While cash-for-work programs have become increasingly popular in low-income countries, a key question is whether they are superior to other types of safety nets. Cash-for-work programs are more expensive to run due to higher administrative costs. While these higher costs are generally justified by the productive effects of public works, empirical grounds on such arguments remain scant. This lack of evidence is problematic because it prevents comprehensive cost-effectiveness exercises and comparisons with alternative safety nets such as unconditional transfers.

This chapter attempts to start filling this gap in the context of the Productive Safety Net Project (PSNP) in Ethiopia. With more than 8 million beneficiaries, the PSNP is among the largest social protection programs in Africa. It is also often considered as Africa's largest climate change adaptation program due to its focus on activities such as land improvements and soil and water conservation measures. To evaluate the effects of these activities and overcome the lack of household data, we use satellite and geo-referenced data. In particular, we combine the Normalized Difference Vegetation Index (NDVI) with highly disaggregated information on land use, crop types, and crop calendars to build a proxy for crop productivity. We examine PSNP's productive value using difference-in-differences estimates covering whole Ethiopia over the 2000-2013 period.

We find no evidence to support that public works had measurable impacts on agricultural productivity and resilience to climate shocks. We provide several robustness checks to assess the validity of our results. In all cases, the effect of the PSNP on agricultural productivity remains small, and non significant. Our results suggest that public works will not always generate measurable effects, and thus call for a more attentive examination of the double dividend that development practitioners typically attribute to public works.

Bibliography

Baland, J.-M., Guirkinger, C., and Mali, C. (2011). Pretending to be poor: Borrowing to escape forced solidarity in Cameroon. *Economic Development and Cultural Change*, 60(1):1–16.

Bandiera, O., Burgess, R., Das, N., Gulesci, S., Rasul, I., and Sulaiman, M. (2017). Labor markets and poverty in village economies. *The Quarterly Journal of Economics*, 132(2):811–870.

Banerjee, A., Duflo, E., Goldberg, N., Karlan, D., Osei, R., Parienté, W., Shapiro, J., Thuysbaert, B., and Udry, C. (2015). A multifaceted program causes lasting progress for the very poor: Evidence from six countries. *Science*, 348(6236):1260799.

Banerjee, A. V. and Duflo, E. (2011). *Poor economics: A radical rethinking of the way to fight global poverty*. Public Affairs.

Banerjee, A., Niehaus, P., and Suri, T. (2019). Universal basic income in the developing world. *Annual Review of Economics*, 11.

Barrett, C. B., Reardon, T., and Webb, P. (2001). Nonfarm income diversification and household livelihood strategies in rural Africa: Concepts, dynamics, and policy implications. *Food Policy*, 26(4):315–331.

Boltz, M., Marazyan, K., and Villar, P. (2019). Income hiding and informal redistribution: A lab-in-the-field experiment in Senegal. *Journal of Development Economics*, 137(1):78–92.

Case, A. and Deaton, A. (1998). Large cash transfers to the elderly in South Africa. *The Economic Journal*, 108(450):1330–1361.

De Janvry, A. and Sadoulet, E. (2001). Income strategies among rural households in Mexico: The role of off-farm activities. *World Development*, 29(3):467–480.

Del Ninno, C. and Mills, B. (2015). *Safety Nets in Africa: Effective Mechanisms to Reach the Poor and Most Vulnerable*. World Bank and Agence Française de Développement Publication.

Dercon, S., De Weerdt, J., Bold, T., and Pankhurst, A. (2006). Group-based funeral insurance in Ethiopia and Tanzania. *World Development*, 34(4):685–703.

Esther, E. (2017). The Economist as Plumber. Working Paper 23213, National Bureau of Economic Research.

Fiszbein, A. and Schady, N. R. (2009). *Conditional cash transfers: Reducing present and future poverty*. The World Bank.

Gollin, D., Parente, S., and Rogerson, R. (2002). The role of agriculture in development. *American Economic Review*, 92(2):160–164.

Gröger, A. and Zylberberg, Y. (2016). Internal labor migration as a shock coping strategy: Evidence from a typhoon. *American Economic Journal: Applied Economics*, 8(2):123–53.

Grosh, M. E., Del Ninno, C., Tesliuc, E., and Ouerghi, A. (2008). *For protection and promotion: The design and implementation of effective safety nets*. The World Bank.

Gubert, F. (2002). Do migrants insure those who stay behind? Evidence from the Kayes area (western Mali). *Oxford Development Studies*, 30(3):267–287.

Haushofer, J. and Shapiro, J. (2016). The short-term impact of unconditional cash transfers to the poor: Experimental evidence from Kenya. *The Quarterly Journal of Economics*, 131(4):1973–2042.

Jakiela, P. and Ozier, O. (2015). Does Africa need a rotten kin theorem? experimental evidence from village economies. *The Review of Economic Studies*, 83(1):231–268.

Lucas, R. E. and Stark, O. (1985). Motivations to remit: Evidence from Botswana. *Journal of Political Economy*, 93(5):901–918.

Rosenzweig, M. R. and Binswanger, H. P. (1993). Wealth, weather risk, and the composition and profitability of agricultural investments. *The Economic Journal*, 103(416):58–78.

Rosenzweig, M. R. and Stark, O. (1989). Consumption smoothing, migration, and marriage: Evidence from rural India. *Journal of Political Economy*, 97(4):905–926.

Subbarao, K., Del Ninno, C., Andrews, C., and Rodríguez-Alas, C. (2012). *Public works as a safety net: Design, evidence, and implementation*. The World Bank.

Townsend, R. M. (1994). Risk and insurance in village India. *Econometrica*, 62:539–539.

Udry, C. (1994). Risk and insurance in a rural credit market: An empirical investigation in northern Nigeria. *The Review of Economic Studies*, 61(3):495–526.

Yang, D. and Choi, H. (2007). Are remittances insurance? Evidence from rainfall shocks in the Philippines. *The World Bank Economic Review*, 21(2):219–248.

World Bank (2018). The State of Social Safety Nets 2018. World Bank, Washington, DC.

Chapter 1

Proxy Means Testing Vulnerability to Measurement Errors?

This chapter is currently "revise and resubmit" in the *Journal of Development Studies*.

1.1 Introduction

Social safety nets programs (SSNP) such as cash and in-kind transfers have become an important tool for achieving poverty alleviation in developing countries. Based on the World Bank Aspire database, the number of developing countries with SSNP doubled from 72 to 149 in the last two decades.[1] However, with an average spending of 1.6 percent of GDP, coverage is far from universal. Governments and development practitioners often use targeting tools in an effort to concentrate the benefits of SSNP on the poorest, but poor households targeting is an inherently inexact and challenging practice, especially in low-income countries which face a lack of verifiable records on earnings. This lack of records often makes means-testing impractical.

Against this backdrop, Proxy Means Testing (PMT thereafter) has become an increasingly popular targeting method. PMT has been implemented in large countries such as Indonesia, Pakistan, Mexico, and the Philippines, as well as in a number of smaller countries, ranging from Ecuador to Jamaica, and more recently to at least 20 African countries (Fiszbein and Schady, 2009; Cirillo and Tebaldi, 2016). In PMT, a survey-based measure of well-being (usually con-

[1] ASPIRE database (Consulted on: `www.worldbank.org/aspire`). See also Beegle et al. (2018) for a focus on Africa.

sumption) is regressed on household covariates to estimate a proxy for well-being, and this proxy is in turn used for targeting out of the sample. Typically, the implementation of PMT has two distinct phases. First, an in-depth survey is administered to a sample of households to collect data on consumption as well as some easily observable and verifiable correlates of consumption (such as demographic characteristics and home attributes). These data are used to estimate a regression of log consumption per capita on correlates of consumption. Second, a short survey is administered to all potential beneficiary households to collect information on the same correlates of consumption, compute PMT scores based on coefficients estimates, and determine the list of beneficiaries based on resulting PMT scores.

PMT is subject to a lively debate among policy makers, civilian stakeholders and academics. The most debated issue is probably the claim that PMT is one of the best mechanisms, if not the best mechanism available for identifying households living in poverty. Del Ninno and Mills (2015, p.20) argue that it "*can accurately and cost-effectively target the chronic poor*". A recent World Bank report recommends the use of PMT to target beneficiaries of social benefits in Namibia because it "*could provide better coverage at existing spending levels, providing a greater poverty and inequality impact*" (Sulla et al., 2017, p.63). In contrast, critics often point to PMT's high built-in errors, implementation issues and lack of transparency. For instance, Kidd and Wylde (2011, p.2) argue that "*PMT is inherently inaccurate, especially at low levels of coverage, and it relatively arbitrarily selects beneficiaries*", while "*other methods (...) may be better at including intended beneficiaries*". Other targeting methods include demographic targeting (targeting of specific categories such as elderly, widowed and children), community-based targeting or CBT (groups of community leaders and members determine eligibility), geographic targeting (location determines eligibility) and self-targeting (benefits and transaction costs are set so that only needy households enrol).[2]

This debate has been feeded by a surge of recent studies assessing the performances of PMT. In these studies, performances are typically displayed in terms of "errors of inclusion" (providing benefits to households which should not be eligible) and "errors of exclusion" (not providing benefits to households that should be eligible). Brown et al. (2018) provide a systematic assessment of PMT performances for nine countries in Sub-Saharan Africa (SSA). The authors find that PMT yields relatively low inclusion errors but high exclusion errors. In the context of Ghana's fertiliser subsidy programme, Houssou et al. (2018) show that PMT would be more efficient and more cost-effective than a univer-

[2] For a detailed overview on PMT and other targeting methods used in developing countries see Grosh (1994); Grosh et al. (2008); Del Ninno and Mills (2015); Devereux et al. (2017); Hanna and Olken (2018).

sal allocation. In Sri Lanka, Sebastian et al. (2018) indicate that switching from self-reported income to PMT could considerably improve the targeting performance of Samurdhi, Sri Lanka's flagship SSNP, and would significantly improve the poverty impact of the program. Comparisons of PMT with Community Based Targeting (CBT) suggest some gains in terms of accuracy but some losses in terms of community satisfaction with the beneficiary list (Alatas et al., 2012; Basurto et al., 2017; Karlan and Thuysbaert, 2016; Premand and Schnitzer, 2018; Stoeffler et al., 2016). For instance, Alatas et al. (2012) in Indonesia report that PMT allowed a 10 percent reduction in the error rate relative to CBT, while CBT resulted in 60 percent fewer complaints than PMT.[3]

An implicit assumption made by these studies is that consumption data underlying PMT regressions are error-free or measured with random errors. However, this assumption has been challenged by recent literature. In particular, Gibson et al. (2015) show that measurement errors in consumption have a mean-reverting negative correlation with true values. According to the typical textbook on the impact of measurement errors, this would lead to biased PMT estimates.[4] However, the magnitude of the bias and its implications on targeting accuracy are not clear.

The goal of this paper is to assess the effects of a violation of this assumption on PMT performances. As with many impact evaluations, the key challenge here is to construct the most credible counterfactual of what would happen with error-free or random measurement errors in consumption. I rely on a unique survey experiment that randomly assigned eight different designs of consumption module to more than 4,000 households in Tanzania. This experiment has been used to explore the relative performances of different survey designs in terms of mean consumption, inequality, poverty, the prevalence of hunger and measurement errors (Beegle et al., 2012; De Weerdt et al., 2016; Beegle et al., 2017; Gibson et al., 2015), but never with an explicit focus on the implications for targeting accuracy. One design of the consumption module involved the distribution of individual diaries to each adult member of households to track all commodity in-flows (harvests, purchases, gifts, destocking) and outflows (sales, gifts, restocking, food fed to animals). In addition, each adult member was provided with tight supervision by interviewers specifically trained to cross-check and query reported information. This resource intensive

[3]Some studies assess PMT targeting outcomes beyond accuracy and satisfaction. Cameron and Shah (2013) show that PMT had significant negative social consequences such as an increase in the prevalence of crime within communities and a decline of the participation in community groups. In the context of a subsidy program in Malawi, Basurto et al. (2017) report that local leaders allocate input subsidies to farmers with larger returns compared to PMT.

[4]See Bound et al. (2001) for a discussion on the impact of measurement errors on regression estimates. In section 1.2.2, I present in more details how Bound et al. (2001) speak to the present study.

design is believed to approximate a "gold standard" for consumption estimates in that it minimizes the prevalence of various sources of measurement errors. My empirical strategy compares the performances of PMT relying on the gold standard consumption data with those of PMT using the more error-prone consumption data.

This paper contributes to the ongoing debate on the methods to target poor households. It provides empirical evidence on one largely ignored aspect of PMT targeting, namely its vulnerability to non-random measurement errors in survey-based consumption data. I estimate that coefficients from PMT regressions are biased in the presence of non-random errors, which results in a reduction in both the predictive and targeting performances of PMT. The predictive performances of PTM decrease by 5 to 27 percent depending on how consumption data is collected. Moreover, using the typical \$1.25 poverty line, the incidence of targeting errors increase by a magnitude ranging from 10 to 34 percent. This latter result is largely driven by an increase in inclusion errors, which suggests that PMT typically overestimates poverty rates. More reassuringly, I find rather small and non-significant effects on targeting performances when poverty is defined in relative terms (such as with the typical 30 percent threshold used in many development projects). This means that non-random errors in consumption have, if anything, a limited impact on the ranking of households.

It is always difficult to extrapolate the results derived from one context and one may be concerned that the findings presented in this paper may not hold in other contexts. However, the focus on measurement errors due to survey design (as opposed to other type of errors such as fraud or fabrication) provides some reassurance that the results are not too specific. Indeed, it is quite reasonable to assume that survey design has a core mechanism that affects respondents answers regardless of the context.

1.2 Measurement Errors and PMT Performances

1.2.1 Consumption measurement errors

Consider the following typical survey questions about some consumed item X:

> *"How much X did your household consume in the past 14 days? How much came from purchases? How much did you spend? How much came from own-production? How much came from gifts and other sources?"*

Often, individuals trying to answer these questions will struggle to give accurate figures, leading to imprecise data.

Why should one expect consumption estimates to deviate from actual consumption?[5] First, it is well documented in the literature that retrospective reports on expenditures can cause both recall and telescoping errors. The longer the period of recall the greater the likelihood events are forgotten or not precisely remembered. A second source of error is the inability of respondents to accurately report individual consumption by other household members, which may be particularly salient in the context of SSA where households are larger and the unitary model has been challenged empirically. This source of error is likely to be more compelling for certain types of consumption such as alcohol, tobacco, meals eaten outside the home, telecommunication or personal toiletries. Lastly, for longer recall periods or items involving frequent transactions, respondents may resort to inference rather than memory to estimate consumption, resulting in what can be termed rule of thumb errors. This source of error has no obvious direction of bias but it is probably more important in hypothetical scenarios requiring high cognitive readiness.

These various sources of errors may be more or less prevalent depending on the design of data collection instruments. In recent years, a number of empirical studies confirmed that measurement of consumption is sensitive to survey design. I focus here on evidence on four key dimensions in which survey design vary: the method of data capture (diary versus recall questionnaires), the length of the recall period, the number of items on which data are collected and the level of respondent (individual versus household). This focus is motivated by the specific experiment exploited in this paper and described in the next section, which randomly assigned households to eight survey designs differing along the four dimensions above-mentionned.[6]

While diaries are generally believed to overcome some sources of error such as recall errors or rule of thumb errors, some concerns related to their implementation in the field have been raised. Specifically, in the case of illiterate, unmotivated or non-cooperative respondents, a diary survey with a lack of supervision may be equivalent to a recall survey if the information is gathered

[5] I only consider deviations caused by the insufficient ability of respondents to acquire, process and recall information. However, it should be noted that deviations can also arise from other sources, such as social desirability bias (e.g. under-reporting of "bad" consumption such as spending on alcohol or cigarettes), strategic responses (e.g. understatements of consumption because of the belief that responses may be used to determine eligibility for some future social program; negative answers bias in order to avoid follow-up questions) and untrained, inadvertent or strategic enumerators (e.g. enumerators guiding respondents to give answers that minimize interview length).

[6] For more detailed discussions on the sensitivity of consumption expenditures to survey design, see for instance Deaton (1997), Deaton and Grosh (2000), Gibson and Kim (2007) and Beegle et al. (2012).

by the enumerator at the end of the period. In Canada, where households reported their food expenditures during the past month and then filled in a diary during the following two weeks, Ahmed et al. (2006) identify substantial measurement errors in recall food consumption with properties inconsistent with random measurement error. However, it also found some discrepancies in the diary survey and concludes that the "superiority of the diary may not be as obvious as the literature suggests". Implementation of diary in developping countries may be even more challenging. Beegle et al. (2012) mention two diary household surveys conducted in Tanzania and Malawi where stylized facts are consistent with poor supervision, respondent fatigue and incomplete or unreliable data. The authors conclude that "the implications of variation in literacy, motivation, and other factors, although not well-documented, suggest it can be quite difficult to conduct high-quality diary survey".

There is a wide understanding that an inverse relationship exists between the length of time over which respondents are asked to recall events and the accuracy of the reported estimates. Events are less likely to be precisely remembered with time due to recall errors and telescoping. While these errors work in opposite directions, experimental studies of self-reported consumption show that under-reporting is more widespread than over-reporting. In an experiment in Ghana, Scott and Amenuvegbe (1991) varied recall periods and find that the reported spending on a basket of the 13 most frequently purchased items decreased by 2.9 percent for every additional day of recall. Similarly, Beegle et al. (2012) in Tanzania report that a 7-day recall design measured a 11 percent higher mean food consumption than a 14-day recall design.

Shorter versus longer lists of items included in questionnaires has also been shown to influence consumption estimates. Observational work by Lanjouw and Ravallion (1996) in Ecuador estimated a decline in poverty of seven percentage points between 1994 and 1995 while the country did not experience any policy to reduce poverty nor significant growth, suggesting that the observed decline in poverty was more related to the change of design in the questionnaire (more than 25 percent additional items was added between the two survey rounds). Jolliffe (2001) confirmed this positive relationship between the number of items and the level of recorded consumption in El Salvador. The author found that more detailed questions on consumption result in an estimate of mean household consumption 31 percent higher than estimates derived from a condensed version of the questionnaire.

Finally, the identity of the respondent to survey questions may influence consumption records due to the difficulty for a sole respondent to perfectly capture the consumption by other household members for items such as alcohol, tobacco, meals eaten outside the home, telecommunication or personal toiletries. As reported by Beegle et al. (2012), personal diaries have been used in Russia for a random sample of households during the 2003 Household Budget Survey, and this yielded 6–11 percent higher expenditure levels, even if the survey was plagued with non-respondent problems.

These examples of diverging consumption estimates when different survey designs are used in the same setting are indicative of measurement error. However, because of a lack of data on actual consumption, there is only scant evidence on the nature of measurement error in estimates of household consumption. One of the main contribution of the survey experiment conducted by Gibson et al. (2015) is that they collect benchmark consumption data allowing them to make such investigations.[7] They find that errors in measured consumption are non-random and negatively correlated with true values – a pattern that Bound and Krueger (1991) also found for earnings data and labelled *mean-reverting measurement error*. In what follows, I present how this pattern may affect PMT performances.

1.2.2 The impact of non-random error in the dependent variable on parameter estimates

A significant amount of attention has been devoted to measurement error and its effects on model estimates. Because this paper is primarily interested in measurement error in consumption, which is used as a left-hand-side variable in PMT regressions, I confine attention to the impact of errors in the dependent variable.[8] Assume the true model is:

$$y = \alpha + \beta \boldsymbol{X} + \varepsilon \tag{1.1}$$

where y is the dependent variable, $\boldsymbol{X}$ a vector of independent variables, β the associated coefficients and ε a pure random error. Instead of y, the observed value of the outcome variable is y^*, which is related to the true value y by:

$$y^* = \theta + \lambda y + v \tag{1.2}$$

The estimator of the response coefficient with the error-ridden dependent variable is:

$$\beta_{y^*\boldsymbol{X}} = \frac{cov(y^*, \boldsymbol{X})}{var(\boldsymbol{X})} = \frac{cov(\lambda\alpha + \lambda\beta\boldsymbol{X} + \lambda\varepsilon - v, \boldsymbol{X})}{var(\boldsymbol{X})} \tag{1.3}$$

One has to assume random error in order to get consistent estimates of β from equation 1.3. Random error is a special case which adds variability to the data

[7] As noted above, this paper rests on the same data as Gibson et al. (2015). More details on the design of the survey are presented in the next section.

[8] The framework presented in this section is adapted from Bound et al. (2001), Hausman (2001) and Gibson et al. (2015).

but does not affect average performance for the sample. The following assumptions are made under random error: $\theta = 0$, $\lambda = 1$ and $E(v) = cov(y, v) = cov(\boldsymbol{X}, v) = cov(\varepsilon, v) = 0$. In contrast, mean-reverting measurement error in y^* assumes $0 < \lambda < 1$ which makes estimates of β inconsistent – from equation 1.3 it is now equal to $\lambda\beta$.

Thus, with $0 < \lambda < 1$, estimates of equation 1.1 will be attenuated. In other words, mean-reverting measurement error in consumption data is expected to bias downward the coefficients of consumption correlates derived from PMT estimates. As noted in the introduction, some assessments of PMT targeting are already available in the literature. However, I am not aware of any previous work looking at the severity of this bias and to what extent it affects PMT performances.[9]

1.3 Survey Experiment

I exploit the same survey experiment as Beegle et al. (2012); De Weerdt et al. (2016); Beegle et al. (2017); Gibson et al. (2015). It is a unique experiment developed by the Living Standards Measurement Study (LSMS) Team in the World Bank in collaboration with the University of Dar es Salaam and the Economic Development Initiatives (EDI thereafter), a leading research company established in 2002 in Tanzania. This section summarizes the experiment and its implementation. More details can be found in Beegle et al. (2012).

1.3.1 Sample

The sample for the experiment consists of 4,032 households spread across seven Tanzanian districts: one district in the regions of Dodoma, Pwani, Dar es Salaam, Manyara, and Shinyanga and two districts in the Kagera Region. While the districts in the regions of Dodoma and Dar es Salaam are urban areas, other

[9] One exception is Brown et al. (2018), which exploits panel data in Ethiopia, Malawi, Nigeria, Tanzania and Uganda to reduce any bias due to measurement errors. The authors use time-mean consumption instead of current consumption and find that PMT performances slightly improve. However, Griliches and Hausman (1986) argue that a crucial parameter in such cases is the correlation over time in the true values of the dependent variable (y in equation 1.1) and in the measurement errors (v in equation 1.2). Specifically, if true values of y are highly correlated over time while the measurement errors v are more or less uncorrelated, moving from cross-sectional estimates to panel estimates would actually intensify the bias due to measurement errors in y.

districts are rural.[10] Within these seven districts, a probability-proportional-to-size sample of 24 villages was selected using data from the 2002 Census. In each selected village, Enumeration Areas (EA) were listed in cooperation with local informants, and one of these EA was randomly chosen for the experiment. These EA are best thought of as sub-villages or neighbourhoods. Finally, in each selected EA, all households were listed, and 24 households were randomly sampled for the survey experiment. According to Beegle et al. (2012), "*the sample was constructed to be representative at the district level, but not at the national level*", however "*the basic characteristics of the sampled households generally match the nationally representative estimates from the 2006/2007 Household Budget Survey*".

1.3.2 Experimental design

In each sub-village, three households were randomly assigned to each of the eight consumption modules summarized in Table A4. Households were assigned to a single module to prevent potential cross-module spillovers. The designs of these eight modules vary along five key dimensions: the method of data capture (diary versus recall questionnaires), the length of the recall period, the number of items in the recall list, the level of respondent (individual versus household) and the degree of supervision received. These eight survey designs were strategically selected to reflect the most common methods used in low-income countries and the scope of variation one is likely to find in practice (Beegle et al., 2012).

Modules 1–5 rely on a recall design and modules 6–8 on diaries. Modules 1 and 2 use a long list of 58 commodities with a recall period of 14 and 7 days respectively. Module 3 uses a subset list consisting of the 17 most important commodities and representing 77 percent of the food consumption expenditure in Tanzania (based on the national Household Budget Survey 2000-2001).[11] Module 4 includes a list of 11 comprehensive categories, which corresponds to an aggregated version of the list of 58 commodities. Module 5 inquires about "usual" consumption over the list of 58 commodities. In particular, households reported the number of months in which the commodity is typically consumed, the quantity usually consumed, and the average value of what is consumed in those months. Modules 6 and 7 are household diaries (i.e. a single diary was

[10] According to Beegle et al. (2012), "districts were purposively selected to capture variations between urban and rural areas as well as across socio-economic dimensions to inform survey design related to labor statistics and consumption expenditure for low-income settings". Table A1 shows basic descriptive statistics.

[11] To make data comparable, reported expenditures for module 3 were scaled up by a factor equal to 1/0.77, as is commonly done in practice.

used to record all household consumption) with different intensity of supervision. Households assigned to module 6 were visited by a trained survey staff every other day, while those assigned to module 7 were only visited weekly. Module 8 is a personal diary in which each adult member was provided with his or her own diary while children were placed on the diaries of the adults who knew most about their daily activities. Each adult was visited every other day.

Non-food items were divided into two categories based on frequency of purchase. Frequently purchased items such as charcoal, soap, cigarettes and communications were collected using a 14-day recall period for modules 1–5 and the 14-day diary for modules 6–8. Non-frequent expenditures such as durables, education and health were collected using the same design across modules (i.e. a one or 12-month recall period depending on the item in question).

1.3.3 Data

The data were collected between September 2007 and August 2008 by EDI. Each interviewer implemented all eight modules in equal proportion in order to avoid confounding module effects with interviewer effects. In each EA, households assigned to the recall modules were surveyed in the span of the 14 days the survey team was in the EA to collect the data based on the diaries. Interviewers were provided with an extensive training starting in June 2007 and including intensive sessions on how to check and correct individual diaries for the issue of double-counting. The survey was administered on paper but maximum control was made possible by the relatively small number of a dozen interviewers and the long 12-month period of data collection. Specifically, back-checks as well as direct observations were carried out on regular basis by supervisors. The same double blind data entry protocol was used for all modules in order to avoid any systematic error to arise and bias the results. Refusal and attrition were negligible: there were only 13 replacements due to refusals and only three households that started a diary were dropped because they did not complete their final interview. Another five households were dropped because of missing data, yielding a final sample size of 4,025 households. A summary of key statistics for the sample is reported in Table A1.

1.4 Empirical Strategy

This paper seeks to quantify the impact of non-random measurement errors in consumption on PMT estimates and to assess how PMT performances are impacted. As with many impact evaluation, the key challenge here consists in constructing the most credible counterfactual of what would happen without measurement errors in consumption. Ideally, we would like to have error-free and error-prone consumption data for each household. Most studies on measurement error rely on validation data such as administrative records for income (Bound and Krueger, 1991). However, the lack of data on actual consumption makes validation studies impractical for consumption. The survey experiment described in section 1.3 offers a rare opportunity to study measurement errors in consumption.

1.4.1 Identification strategy

A key assumption of the identification strategy is that the personal diary (module 8) provides "gold standard" (or "benchmark") data on consumption. In the personal diary, there is a smaller scope for recall errors, telescoping and missed individual consumption. In addition, three measures have been undertaken to avoid double counting – the main stated weakness of personal diaries. First, the personal diary has been designed as a record of food brought into the household instead of food consumed, which is likely to reduce the scope for double-counting purchased or self-produced items. Second, as discussed, interviewers were trained to cross-check individual diaries for similar items and apply the appropriate corrections when the same item was accidentally recorded by two individuals. Third, each adult member was visited every other day in order to provide him or her with adequate supervision. Reassuringly, some statistics, such as the daily consumption, show no diary fatigue.

The identification strategy exploits this benchmark consumption and the random assignment of the different survey designs across households. Table A3 shows the results of randomization balance checks across a set of core household characteristics. Overall, randomized assignment of households to the eight different designs seems successful. Six of the differences are statistically significant at conventional levels. These differences, while significant, are not too worrying because they are small in size. Consequently, any systematic difference in measured consumption across modules can be attributed to measurement error due to alternative survey designs. Comparisons of error prone survey designs with the benchmark give estimates of the effect of measurement error on PMT targeting.

1.4.2 Estimation procedure and construction of the outcomes of interest

While I recognize that poverty is multidimensional in nature, I rely on per capita consumption as the main welfare indicator for the analysis because it is generally considered as a good predictor of neediness (Deaton, 1997) and because it is used in most PMT targeting exercises. Per capita consumption is aggregated on an annual basis using data collected on food consumption and frequent non-food consumption.[12] Total food consumption from module 3 is scaled up by a factor equal to 1/0.77 (i.e. 29.87 percent) to make data comparable across modules.

First, I create a set of variables that are long-term determinants or correlates of poverty, encompassing household's demographic characteristics (household size, number of children, etc.), home attributes (floor type, wall type, etc.) and household head's features (education, occupation, etc.). These variables have been selected to be representative of the variables typically included in PMT targeting.[13] Then, using an OLS estimator with a backward stepwise selection of the variables, I estimate the relationship between this set of variables and log consumption per capita (the so-called PMT formula) by module type.[14] The following regression is estimated eight times (one for each sample of households assigned to module type k):

$$y_{ik} = \alpha_k + \beta_k X_i + \varepsilon_{ik} \tag{1.4}$$

where y_{ik} is the log consumption per capita of household i (with $i = 1, ..., N_k$; N_k the sample size of households assigned to module k; $k = 1, ..., 8$), X_i the set of correlates of consumption. Estimates from equation 1.4 are then used to predict PMT scores of household i for each PMT formula k:

$$\hat{y}_{ik} = \hat{\alpha}_k + \hat{\beta}_k X_i \tag{1.5}$$

[12] Results are robust using food consumption only (see Tables A5, A9, A13 and A17) or consumption per adult equivalent (see Tables A6, A10, A14 and A18). In both cases, the consumption of non-frequently purchased items such as durables, education and health was excluded because it was collected using the same design across modules (i.e. there is no benchmark) and because it is usually not included in PMT. That said, it would be quite reasonable to assume that measurement errors for these items are more prevalent because of the longer recall period (one month or 12 months depending on the items considered). Unfortunately, I am not able to check this assumption because there are no benchmark data on actual non-frequent consumption.

[13] In an extended version, I also include variables on assets and livestock that are good correlates of consumption but are more difficult to verify and may be vulnerable to strategic responses. The results are similar (see Tables A7, A11, A15 and A19).

[14] Tables A8, A12, A16 and A20 show that results are largely similar without the stepwise option.

Note that each PMT formula is used to compute PMT scores "in and out of sample" (e.g. formula 1 is used to predict PMT scores of households assigned to module 1 but also for the sample of households assigned to the other modules). As a result, I obtain eight PMT scores per household (one from each of the eight PMT formulas) which form the basis to assess the impact of measurement errors on PMT performances. Then, I restrict the analysis to the sample of households for which benchmark consumption data are available, i.e. those assigned to the personal diary (module 8), and compare how each formula perform to predict their consumption and which households are poor. Under the identifying assumption that the personal diary approximates a benchmark for true consumption, formula 8 can be interpreted as the closest to the counterfactual scenario, i.e. the PMT formula one would obtain if consumption was measured without errors.

In a first part, I compare the predictive performances of the alternative PMT formulas. I estimate an equation of the following form using OLS:

$$\hat{y}_{ik} = \gamma_k M_{ik} + v_{ik} \tag{1.6}$$

where $\hat{y}_{ik}$ is the PMT score of household i derived from formula k (see equation 1.5) and M_{ik} a vector of dummy variables indicating if $\hat{y}_{ik}$ is derived from formula k. I also compute the mean squared prediction error $\hat{\mu}_{ik} = (y_i - \hat{y}_{ik})^2$, where y_i is the individual diary consumption, and regress it on the same variables:

$$\hat{\mu}_{ik} = \gamma_k M_{ik} + v_{ik} \tag{1.7}$$

In both estimates, standard errors are clustered at the village level to account for the correlation between the error terms of observations from the same village. The comparison of γ_k ($k = 1, ..., 7$) with γ_8 gives the impact of measurement errors on PMT predictive performances by survey design.

In a second part, I compare the performances of the alternative PMT formulas against different measures of targeting accuracy. As discussed in the introduction, there are two types of targeting errors: Inclusion Errors (IE), i.e. identifying a non-poor household as poor, and Exclusion Errors (EE), i.e. identifying a poor household as non-poor. The Inclusion Error Rate (IER), defined as the proportion of the non-poor households identified as poor, for module k, can be written as:

$$IER_k = \frac{\sum_{i=1}^{N_k} \mathbb{1}(\hat{y}_{ik} \leq z \mid y_i > z)}{\sum_{i=1}^{N_k} \mathbb{1}(y_i > z)} \tag{1.8}$$

where N_k is the sample size, z the poverty line, y_i the measured per capita consumption of household i, $\hat{y}_{ik}$ its PMT score using PMT formula k and $\mathbb{1}(.)$ an indicator function which takes the value one when the condition in parentheses is true and zero otherwise.[15]

Similarly, the Exclusion Error Rate (EER), defined as the proportion of the poor households not identified as poor, can be written as:

$$EER_k = \frac{\sum_{i=1}^{N_k} \mathbb{1}(\hat{y}_{ik} > z \,|\, y_i \leq z)}{\sum_{i=1}^{N_k} \mathbb{1}(y_i \leq z)} \tag{1.9}$$

The IER and the EER do not consider how far from the poverty line beneficiary and non-beneficiary households lie. For instance, the EER would be the same if a given household i, excluded by mistake, is just below or very far below the poverty line. Hence, mean squared errors, which allocate a higher weight for errors farther from the poverty line, is perhaps richer for measuring targeting errors. The Mean Squared IE (MSIE) and Mean Squared EE (MSEE) for module k are given by:

$$MSIE_k = \frac{\sum_{i=1}^{N_k} \mathbb{1}(\hat{y}_{ik} \leq z \,|\, y_i > z) * (z - y_i)^2}{\sum_{i=1}^{N_k} \mathbb{1}(y_i > z)} \tag{1.10}$$

$$MSEE_k = \frac{\sum_{i=1}^{N_k} \mathbb{1}(\hat{y}_{ik} > z \,|\, y_i \leq z) * (z - y_i)^2}{\sum_{i=1}^{N_k} \mathbb{1}(y_i \leq z)} \tag{1.11}$$

The IER, the EER, the Targeting Error Rate (TER), defined as the weighted sum of the IER and the EER (weights are the share of poor/non-poor households), the MSIE, the MSEE and the MSTE, defined as the weighted sum of the MSIE and the MSEE, form the basis to assess the targeting performances of the alternative PMT formulas. From the rates and means defined above, I construct variables which can fit in typical regression frameworks. Specifically, for the IER, the EER and the TER, I create dummies equal to one if household i with consumption derived from formula k is mistargeted, and zero otherwise. For instance, IE_{ik} is equal to one for all households i which are considered as poor by mistake using PMT formula k. Similarly, for the MSIE, the MSEE and the MSTE, I create variables equal to the squared targeting error if household i with consumption derived from PMT formula k is mistargeted, and zero otherwise. For instance, IE^2_{ik} is equal to the squared inclusion error (i.e. $(y_i - \hat{y}_{ik})^2$) for all households i which are considered as poor by mistake using PMT formula k. Each of these outcomes of interest is estimated with the same specification as equations 1.6 and 1.7 using a linear probability model. Importantly,

[15] I use the same definition as Alatas et al. (2012). IER could also be defined as Brown et al. (2018), i.e. the proportion of those identified as poor who are not poor. The latter definition is less practical in the present study. Indeed, the sample of households identified as poor is likely to vary across PMT formulas.

the poverty line z in equations 1.8–1.11 can be defined in absolute or in relative terms. With a poverty line defined in absolute terms, e.g. PPP$1.25, beneficiaries are those with a PMT score below PPP$1.25. With a poverty line defined in relative terms, e.g. the poorest 30 percent, beneficiaries are those with a PMT score equal or below the PMT score of the 30th percentile. I start by assessing PMT targeting performances with respect to the typical PPP$1.25 poverty line. Then I use a poverty line defined in relative terms using the 30 percent threshold used in many SSNP. Specifically, for each PMT formula, I rank households from lowest to highest PMT scores and consider as eligible those with PMT scores equal or below the PMT score of the 30th percentile.

1.5 Results

1.5.1 PMT estimates

Table 1.1 presents the results of PMT regressions by module type. Adjusted-R^2 range from 0.45 for the sample assigned to household diary with infrequent supervision (module 7), to 0.64 for the sample assigned to the usual month recall (module 5). Column 9 displays the PMT on the full sample of households and Adjusted-R^2 is 0.54, which is slightly lower than the 0.59 Adjusted-R^2 obtained by Brown et al. (2018) in Tanzania using LSMS-ISA data, and somewhat higher than the 0.40 obtained in Indonesia by Alatas et al. (2012). Interestingly, while coefficients and variables selected through the backward stepwise procedure vary somewhat across specifications, signs do not change (with a few exceptions). Overall, coefficients are larger for households assigned to the benchmark (module 8), which is consistent with the assumption that non-random measurement errors in consumption data bias downward PMT estimates (see Section 1.2.2 above).

1.5.2 PMT predictive performances

Simple comparisons of distributions of PMT scores using different formulas in Figure 1.1 show that the benchmark PMT formula yields relatively higher scores. The distribution of scores is shifted to the right compared to other formulas.[16] This is confirmed by the results of regressions presented in Table 1.2. Overall, formulas 1–7 yield significantly lower PMT scores and higher

[16] Figure A1 similarly compares raw distributions of consumption before PMT regressions and shows that households assigned to module 8 have higher scores.

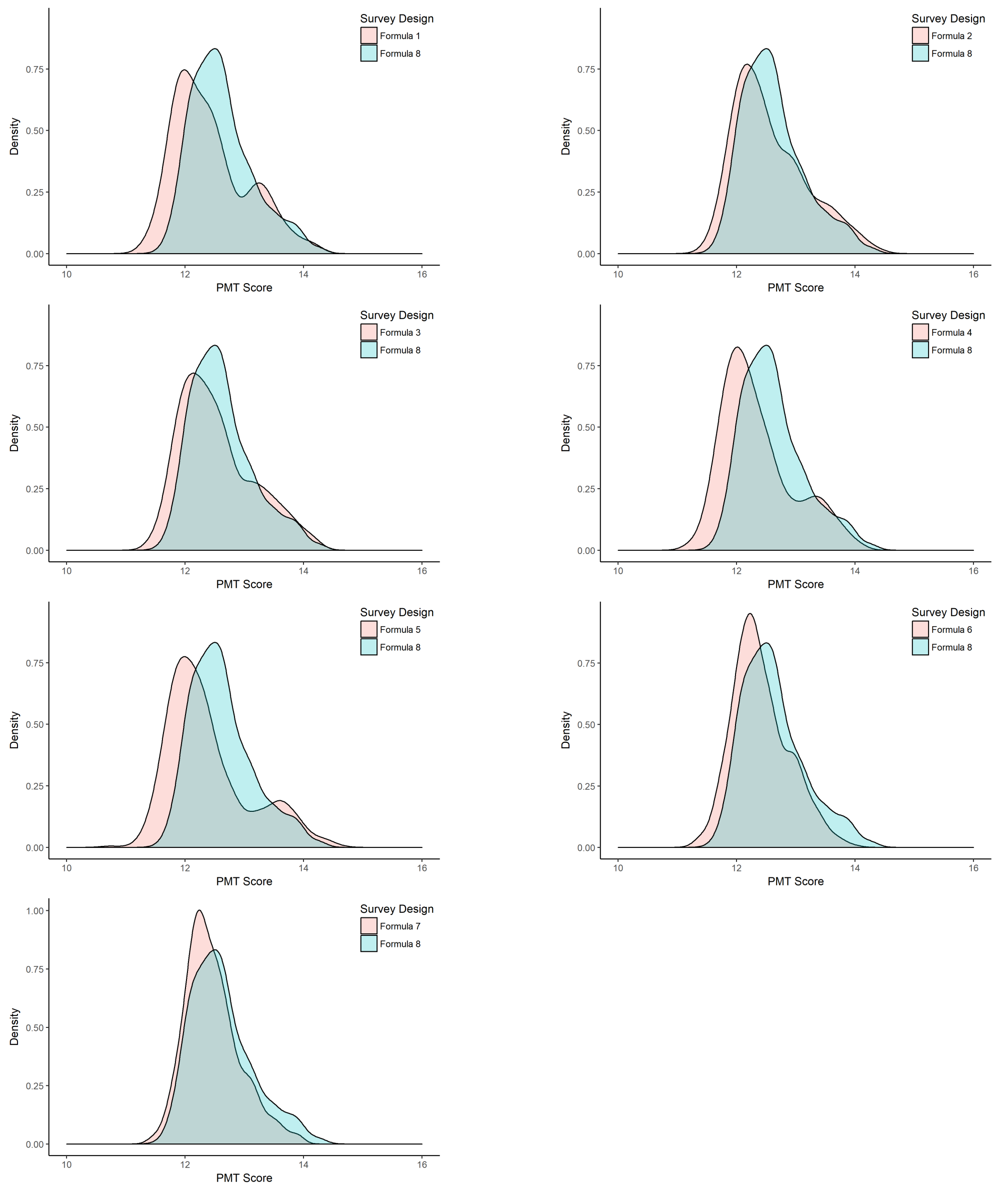
squared prediction errors than formula 8 (derived from the benchmark personal diary). The results in column 1 show that formulas 1–7 predict between 5 and 27 percent lower PMT scores compared with the benchmark formula 8. Similarly, formulas 1–7 produce mean prediction errors between 12 and 49 percent higher compared with the benchmark. PMT formulas derived from the long list 7-day recall (module 2) and the subset list (module 3) appear to yield slightly better predictions compared with formulas derived from the collapsed list (module 4) and usual month (module 5). Non-random measurement errors in consumption have thus a significant and rather large impact on the predictive performances of PMT. In the next section I will investigate how these relate to targeting performances.

Table 1.1: PMT Regressions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Module 1	Module 2	Module 3	Module 4	Module 5	Module 6	Module 7	Module 8	All
Hhsize	-0.249***	-0.289***	-0.176***	-0.231***	-0.212***	-0.158***	-0.184***	-0.249***	-0.197***
	(0.033)	(0.031)	(0.041)	(0.030)	(0.036)	(0.019)	(0.022)	(0.027)	(0.014)
Hhsize^2	0.011***	0.014***	0.010***	0.009***	0.011***	0.007***	0.007***	0.009***	0.009***
	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.001)	(0.001)	(0.002)	(0.001)
Elderly		0.110*						-0.156***	
		(0.056)						(0.048)	
Young Children	-0.076***	-0.072**	-0.150***	-0.063**	-0.132***	-0.089***	-0.059**		-0.100***
	(0.029)	(0.030)	(0.030)	(0.030)	(0.029)	(0.028)	(0.027)		(0.012)
Children			-0.049*						-0.023**
			(0.029)						(0.012)
Mud/Dirt Floor		-0.118*				-0.146***			-0.074**
		(0.060)				(0.056)			(0.031)
Thatch Roof				-0.159***		-0.134**	-0.187***		-0.066**
				(0.060)		(0.052)	(0.057)		(0.026)
Mud Walls	-0.308***		-0.326***	-0.253***	-0.266***	-0.249***		-0.164**	-0.193***
	(0.082)		(0.058)	(0.071)	(0.065)	(0.076)		(0.081)	(0.031)
N Rooms		0.042**		0.041**			0.072***	0.098***	0.039***
		(0.020)		(0.018)			(0.016)	(0.017)	(0.008)
Electricity		0.234**	-0.138**	0.186**		0.307***	0.343***	0.393***	0.072*
		(0.096)	(0.069)	(0.091)		(0.080)	(0.096)	(0.092)	(0.044)
Urban	0.126*	0.289***						0.283***	0.101**
	(0.076)	(0.072)						(0.085)	(0.040)
Water			0.122**						
			(0.061)						
Flushed Toilet					0.185*	0.206**	0.356***		0.161***
					(0.098)	(0.097)	(0.090)		(0.052)
Cooking	0.575***	0.505***	0.590***	0.665***	0.787***			0.216**	0.418***
	(0.112)	(0.086)	(0.070)	(0.090)	(0.073)			(0.105)	(0.050)
Married	-0.238**		-0.223***			-0.236**	-0.353***	0.189***	-0.149***
	(0.117)		(0.082)			(0.094)	(0.117)	(0.060)	(0.042)
Widowed	-0.171*	-0.129*							-0.105***
	(0.102)	(0.074)							(0.039)
Age		0.023**		0.019*					
		(0.009)		(0.010)					
Age^2		-0.000***		-0.000**	-0.000***				-0.000***
		(0.000)		(0.000)	(0.000)				(0.000)
Male	0.347***		0.315***	0.195***		0.370***	0.332***		0.179***
	(0.119)		(0.081)	(0.063)		(0.097)	(0.117)		(0.039)
Primary		0.120**				0.142**	0.197***		0.077***
		(0.058)				(0.056)	(0.054)		(0.025)
Secondary			0.287***		0.449***		0.291***		0.154***
			(0.109)		(0.097)		(0.092)		(0.041)
Primary Max	0.222**				0.260**				
	(0.096)				(0.105)				
Secondary Max	0.249***	0.260***	0.151**	0.165**				0.235***	0.106***
	(0.068)	(0.069)	(0.070)	(0.066)				(0.071)	(0.028)
Ajusted-R^2	0.58	0.60	0.62	0.59	0.64	0.46	0.45	0.47	0.54
Observations	503	504	504	504	504	502	501	503	4025

Notes: This table reports regressions of per capita consumption (in log) as reported in different survey designs. Sequential selection of variables has been done using backward stepwise regression. Definition of the variables are provided in Table A2. The sample in columns 1–8 is restricted to households assigned to a certain consumption module. Results for the full sample are reported in column 9. OLS estimator is used for all regressions. Standard errors in parentheses are clustered at the village level. *** p<0.01, ** p<0.05, * p<0.1.

Figure 1.1: Comparing distributions of PMT scores by survey design



Notes: Each sub-figure compares the distribution of PMT scores derived from Formula 8 (the benchmark) with distributions of PMT scores of households derived from Formula k (with $k = \{1, 7\}$).

Table 1.2: Predictive Performances

	(1)	(2)
	$\hat{y}_{ik}$	$\hat{\mu}_{ik}$
Formula 1	-0.195***	0.100***
	(0.014)	(0.017)
Formula 2	-0.050***	0.034***
	(0.011)	(0.012)
Formula 3	-0.089***	0.073***
	(0.015)	(0.015)
Formula 4	-0.269***	0.105***
	(0.013)	(0.019)
Formula 5	-0.239***	0.145***
	(0.020)	(0.024)
Formula 6	-0.210***	0.089***
	(0.014)	(0.018)
Formula 7	-0.159***	0.080***
	(0.017)	(0.018)
F-statistics	193.98***	7.20***
Observations	4024	4024
Number of Households	503	503
Mean in *Formula* 8	12.621	0.293

Notes: This table reports regressions of predictive performances of PMT by survey design. $\hat{y}_{ik}$ is the predicted value of the log consumption per capita (PMT score) of household i for formula k. $\hat{\mu}_{ik}$ is the squared prediction error for household i and formula k. Formula k (with $k = \{1, 8\}$) is a dummy variable taking the value of 1 if PMT Formula k is used to derive $\hat{y}_{ik}$. All coefficients are interpretable relative to formula 8, which is the omitted category and the benchmark to assess the impact of measurement error on the predictive performances by survey design. OLS estimator is used for both regressions. Robust standard errors clustered at the village level in parentheses. F-test is performed on the null hypothesis that the coefficients of all controls are jointly zero. *** p<0.01, ** p<0.05, * p<0.1.

1.5.3 PMT targeting performances

Regressions in Table 1.3 compare the targeting performances of each of the seven formulas against the benchmark formula derived from the sample of households assigned to the personal diary (module 8), using the PPP$1.25 poverty line. The results in column 1 show that measurement errors in formulas 1–7 increase the TER by a magnitude ranging from 2.4 and 8.3 percentage points. Given that the TER derived from formula 8 is 24.7 percent, these effects are equivalent to an increase in TER of 10 to 34 percent. In columns 2 and 3, I examine the error rates separately for the non-poor and the poor (defined as the households above/below the PPP$1.25 poverty line). The results show that the IER increase and the EER decrease for all formulas compared with formula 8, which is not surprising given that Table 1.2 found that formulas 1–7 predict lower PMT scores. This means that the number of poor households is overestimated when formulas 1–7 are used.

I further investigate whether this pattern (higher TER and EER and lower IER) holds when the poverty line is defined in relative terms. Figure 1.2 looks at whether a household position in the distribution of PMT scores is influenced to some extent by the formula being used to predict her score. Each point in the graphs represent the percentile of a household in the consumption distribution when PMT formula 8 is used against the percentile in the consumption distribution for the same household when PMT formula k (with $k = \{1,7\}$) is used. If measurement errors had no distributive impacts, each point should be on the diagonal. Households are relatively well distributed around the diagonal, even though large deviations exist for some households. Spearman correlations range from 0.83 for PMT scores derived from formula 7 (household diary with infrequent supervision) to 0.93 for PMT scores derived from formula 2 (7 day recall with the long list of items) and formula 4 (7 day recall with the collapsed list of items). Table 1.4 refines the insights from Figure 1.2 by investigating the results of regressions. Poverty is defined in relative terms, using a typical 30 percent threshold. Interestingly, the coefficients are now much smaller in magnitudes and not statistically significant (except formula 5, derived from the usual month recall module). Point estimates correspond to a 0.4 to 3.2 percentage point increase in TER using formulas 1–7. Similarly, both the IER and EER estimates find small and statistically insignificant coefficients (columns 2 and 3). These results provide evidence that if anything measurement errors in consumption does not affect to a great extent the distribution of poor households. In other words, measurement errors in consumption seem to have relatively weak implications on the distribution of PMT scores.

The results presented in columns 4–6 in tables 1.3 and 1.4 suggest the effects of measurement errors in consumption on the MSTE, the MSIE and the MSEE are similar to those found for the TER, the IER and the EER.

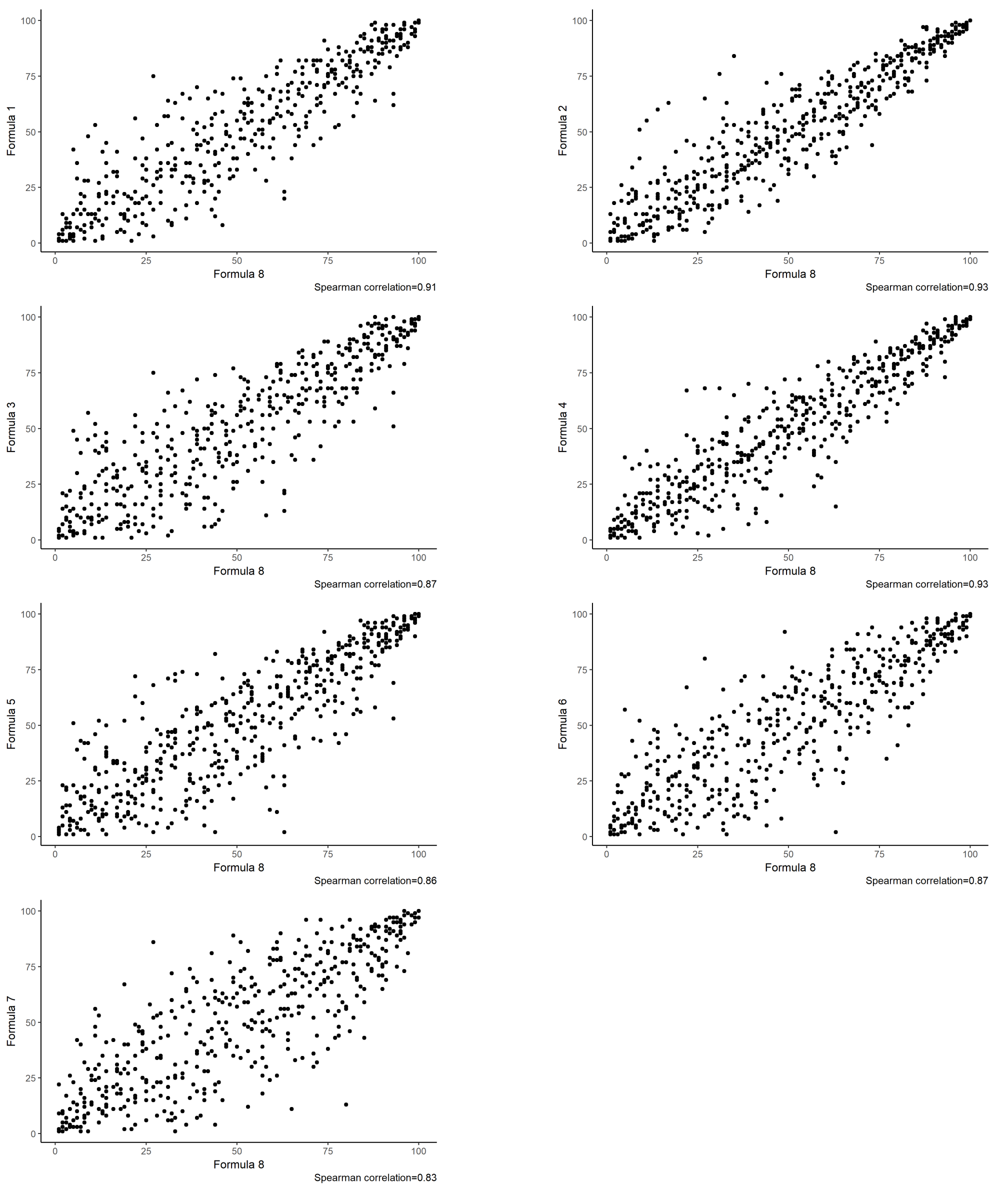
Measurement errors in consumption may be correlated with household characteristics. For instance, the number of adults in the household may affect the relative prevalence of measurement errors across modules (individual consumption from other adult household members may be missed in designs with a sole respondent). Table 1.5 explores the potential effects that interactions between key household characteristics (household size, number of adults, literacy, urban/rural location) and the formulas dummies could have on the TER. The poverty line is defined in absolute terms in Panel A and in relative terms in Panel B. Household size and the number of adult members do not seem to mediate the impact of measurement errors on targeting accuracy. No interaction term is significant for any formula except formula 5 (the usual month recall) in Panel A and formulas 4–6 in Panel B. Column 3 shows that literate households seem more vulnerable to targeting errors (due to measurement errors) vis-à-vis the benchmark (module 8). For six out of seven formulas in Panel A interaction terms are significantly different from zero. Finally, household vulnerability to targeting errors (due to measurement errors) does not seem to depend on urban/rural location.

Table 1.3: Targeting Performances, \$1.25 Poverty Line

	(1) TE_{ik}	(2) IE_{ik}	(3) EE_{ik}	(4) TE^2_{ik}	(5) IE^2_{ik}	(6) EE^2_{ik}
Formula 1	0.054**	0.200***	-0.266***	0.034***	0.072***	-0.049***
	(0.023)	(0.025)	(0.036)	(0.012)	(0.016)	(0.014)
Formula 2	0.040**	0.116***	-0.127***	0.021**	0.046***	-0.034**
	(0.017)	(0.020)	(0.032)	(0.010)	(0.013)	(0.014)
Formula 3	0.050**	0.136***	-0.139***	0.021**	0.044***	-0.029**
	(0.020)	(0.024)	(0.037)	(0.010)	(0.013)	(0.012)
Formula 4	0.058**	0.238***	-0.335***	0.037***	0.085***	-0.070***
	(0.024)	(0.026)	(0.038)	(0.014)	(0.017)	(0.018)
Formula 5	0.083***	0.261***	-0.304***	0.060***	0.120***	-0.070***
	(0.024)	(0.027)	(0.039)	(0.016)	(0.020)	(0.019)
Formula 6	0.024	0.130***	-0.209***	0.007	0.040***	-0.065***
	(0.021)	(0.024)	(0.038)	(0.011)	(0.012)	(0.023)
Formula 7	0.044**	0.104***	-0.089**	0.014	0.041***	-0.043*
	(0.019)	(0.022)	(0.043)	(0.011)	(0.010)	(0.026)
F-statistics	2.25**	17.75***	13.74***	2.68**	6.99***	3.77***
Observations	4024	2760	1264	4024	2760	1264
Number of Households	503	345	158	503	345	158
Mean in *Formula* 8	0.247	0.139	0.481	0.059	0.029	0.124

Notes: This table reports regressions of targeting performances of PMT by survey design. The dependent variable in column 1 is a dummy equal to 1 if household i with consumption derived from PMT Formula k is mistargeted, and 0 otherwise. Dependent variable in column 4 is equal to mean squared error if household i with consumption derived from PMT Formula k is mistargeted, and 0 otherwise. Columns 2–3 and 5–6 disaggregate the results by error type. Formula k (with $k = \{1, 8\}$) is a dummy variable taking the value of 1 if PMT Formula k is used to predict Y_{ik}. All coefficients are interpretable relative to formula 8, which is the omitted category and the benchmark to assess the impact of measurement error on the predictive performances by survey design. The mean of the dependent variable in formula 8 is shown in the bottom row. LPM is used for regressions 1–3. OLS is used for regressions 4–6. Standard errors in parentheses are clustered at the village level. F-test is performed on the null hypothesis that the coefficients of all controls are jointly zero. *** p<0.01, ** p<0.05, * p<0.1.

Figure 1.2: Correlation between PMT score's percentile predicted by the benchmark PMT formula (formula 8) and the seven other formulas



Notes: Each point in the graphs represent the percentile of the household in the consumption distribution when PMT formula 8 is used (x-axis) against the percentile in the consumption distribution for the same household when PMT formula k (with $k = \{1, 7\}$) is used (y-axis).

Table 1.4: Targeting Performances, 30% Poverty Threshold

	(1) TE_{ik}	(2) IE_{ik}	(3) EE_{ik}	(4) TE_{ik}^2	(5) IE_{ik}^2	(6) EE_{ik}^2
Formula 1	0.024	0.026	0.020	0.006	0.002	0.015
	(0.016)	(0.018)	(0.036)	(0.008)	(0.010)	(0.014)
Formula 2	0.004	0.003	0.007	0.004	0.007	-0.002
	(0.012)	(0.013)	(0.031)	(0.006)	(0.007)	(0.014)
Formula 3	0.024	0.017	0.040	0.002	0.000	0.008
	(0.016)	(0.020)	(0.040)	(0.008)	(0.009)	(0.015)
Formula 4	0.012	0.009	0.020	0.002	-0.003	0.013
	(0.015)	(0.018)	(0.035)	(0.005)	(0.005)	(0.014)
Formula 5	0.032*	0.023	0.053	0.008	0.003	0.018
	(0.016)	(0.020)	(0.038)	(0.007)	(0.008)	(0.014)
Formula 6	0.022	0.017	0.033	0.000	0.001	-0.001
	(0.019)	(0.022)	(0.047)	(0.010)	(0.010)	(0.026)
Formula 7	0.010	0.009	0.013	0.000	0.009	-0.019
	(0.018)	(0.021)	(0.048)	(0.010)	(0.010)	(0.026)
F-statistics	0.65	0.40	0.54	0.62	0.59	0.82
Observations	4024	2816	1208	4024	2816	1208
Number of Households	503	352	151	503	352	151
Mean in *Formula* 8	0.258	0.185	0.430	0.065	0.044	0.114

Notes: Inclusion threshold is adjusted to obtain 30% of the household targeted for each module. LPM is used for regressions 1–3. OLS is used for regressions 4–6. Standard errors in parentheses are clustered at the village level. *** p<0.01, ** p<0.05, * p<0.1. See notes to Table 1.3 for other details.

Table 1.5: Interaction of PMT formula and select household characteristics

	(1) Household size	(2) Number of adults	(3) Literacy	(4) Urban
Panel A:				
Targeting Error ($1.25 poverty line)				
Interaction 1	0.009	0.005	0.078*	-0.038
	(0.006)	(0.011)	(0.041)	(0.043)
Interaction 2	0.002	-0.002	0.014	-0.017
	(0.004)	(0.010)	(0.035)	(0.030)
Interaction 3	0.010	0.010	0.080**	-0.014
	(0.006)	(0.013)	(0.040)	(0.039)
Interaction 4	0.004	-0.001	0.110**	0.009
	(0.008)	(0.014)	(0.047)	(0.046)
Interaction 5	0.014**	0.031**	0.101**	0.014
	(0.007)	(0.013)	(0.048)	(0.047)
Interaction 6	0.009	0.010	0.072*	0.017
	(0.007)	(0.014)	(0.043)	(0.043)
Interaction 7	0.004	-0.008	0.088**	0.013
	(0.006)	(0.012)	(0.043)	(0.036)
Panel B:				
Targeting Error (30% poverty threshold)				
Interaction 1	0.008	0.011	0.022	0.017
	(0.005)	(0.011)	(0.036)	(0.031)
Interaction 2	0.003	0.002	0.015	0.020
	(0.004)	(0.011)	(0.030)	(0.021)
Interaction 3	0.006	0.002	0.047	0.043
	(0.005)	(0.012)	(0.038)	(0.033)
Interaction 4	0.012**	0.017*	0.036	0.043
	(0.005)	(0.009)	(0.031)	(0.031)
Interaction 5	0.017***	0.032***	0.026	0.013
	(0.006)	(0.012)	(0.035)	(0.034)
Interaction 6	0.013*	0.011	0.052	0.011
	(0.007)	(0.014)	(0.039)	(0.040)
Interaction 7	0.002	-0.010	0.058	0.029
	(0.006)	(0.012)	(0.042)	(0.036)
Observations	4024	4024	4024	4024
Number of Households	503	503	503	503

Notes: This table represents the results of (separate) LPM estimates of a selected measure of targeting performances (mentioned in panels' title) on PMT formula dummies, a single selected household characteristic (mentioned in the column headings) and their interactions. Only the interaction terms are reported due to space limitations. Interaction k (with $k = \{1, 8\}$) is an interactive variable between the characteristic mentioned in the column heading and formula k dummy. Standard errors in parentheses are clustered at the village level. *** p<0.01, ** p<0.05, * p<0.1. See notes to Table 1.3 and 1.4 for other details.

1.6 Conclusion

In this paper, I have investigated the impact of non-random measurement error on PMT performances. Assessments of PMT performances rely on the assumption that consumption data underlying PMT regressions are error-free or measured with random error, even though this assumption has been challenged by recent literature. Using a unique survey experiment in Tanzania, I show that the presence of non-random measurement error in consumption reduces the predictive performances of PMT by a magnitude ranging from 5 to 27 percent, which in turn induces a 10 to 34 percent increase in the incidence of targeting errors (using the typical PPP $1.25 poverty line). More reassuringly, when poverty is defined in relative terms, impacts on the relative distribution of households are small and non-significant, meaning that measurement errors in consumption have weak implications on the distribution of PMT scores.

Some unresolved questions remain. First, I only discussed one dimension of PMT, i.e. its predictive and targeting performances, and more attention on cost-efficiency, transparency, fairness and acceptance would be welcome. Second, I focused on measurement errors in the dependent variable, while measurement errors in the independent variables could also impact PMT performances. Third, I do not take into account PMT vulnerability to data fraud or data fabrication by interviewers. I chose instead to focus on measurement errors due to survey design, which are likely to have more external validity. However, the problem of data fabrication in surveys has been shown to be prevalent (Finn and Ranchhod, 2017). Finally, recent studies such as McBride and Nichols (2016) have shown that new tools from machine learning applied to poverty prediction outperform PMT. These new tools are typically trained on survey-based data and documenting whether they are also vulnerable to measurement errors is a potential avenue for future research.

Nevertheless, the results presented in this paper provide empirical evidence on one largely ignored aspect of PMT, namely its vulnerability to non-random errors due to survey design. The results may be of relevant interest to researchers in their assessments of PMT performances or in their comparisons of the different targeting mechanisms available. It also has implications for development practitioners and governments designing the targeting devices of the many SSNP implemented in developing countries. If the objective is to target people below the poverty line, then PMT performances are quite vulnerable to measurement errors. However, if the objective is to target a fixed share of the population (regardless of whether they are above or below the poverty line), PMT performances are quite robust to the presence of measurement errors.

Bibliography

Ahmed, N., Brzozowski, M., and Crossley, T. F. (2006). Measurement errors in recall food consumption data. Working Paper 06/21, Institute for Fiscal Studies (IFS).

Alatas, V., Banerjee, A., Hanna, R., Olken, B. A., and Tobias, J. (2012). Targeting the poor: Evidence from a field experiment in Indonesia. *The American Economic Review*, 102(4):1206–1240.

Basurto, P. M., Dupas, P., and Robinson, J. (2017). Decentralization and efficiency of subsidy targeting: Evidence from chiefs in rural Malawi. Working Paper 23383, National Bureau of Economic Research.

Beegle, K., Coudouel, A., and Monsalve, E. (2018). Realizing the full potential of social safety nets in Africa. Washington, DC World Bank.

Beegle, K., De Weerdt, J., Friedman, J., and Gibson, J. (2012). Methods of household consumption measurement through surveys: Experimental results from Tanzania. *Journal of Development Economics*, 98(1):3–18.

Bound, J., Brown, C., and Mathiowetz, N. (2001). Measurement error in survey data. *Handbook of Econometrics*, 5:3705–3843.

Bound, J. and Krueger, A. B. (1991). The extent of measurement error in longitudinal earnings data: Do two wrongs make a right? *Journal of Labor Economics*, 9(1):1–24.

Brown, C., Ravallion, M., and van de Walle, D. (2018). A poor means test? Econometric targeting in Africa. *Journal of Development Economics*, 134:109–124.

Cameron, L. and Shah, M. (2013). Can mistargeting destroy social capital and stimulate crime? Evidence from a cash transfer program in Indonesia. *Economic Development and Cultural Change*, 62(2):381–415.

Cirillo, C. and Tebaldi, R. (2016). Social protection in Africa: inventory of non-contributory programmes. *IPC-IG*.

De Weerdt, J., Beegle, K., Friedman, J., and Gibson, J. (2016). The challenge of measuring hunger through survey. *Economic Development and Cultural Change*, 64(4):727–758.

Deaton, A. (1997). *The analysis of household surveys: a microeconometric approach to development policy*. World Bank Publications.

Deaton, A. and Grosh, M. (2000). Consumption. In *Designing Household Questionnaires for Developing Countries: Lessons from Fifteen Years of the Living Standard Measurement Study*, volume 15, pages 91–133.

Del Ninno, C. and Mills, B. (2015). *Safety Nets in Africa: Effective Mechanisms to Reach the Poor and Most Vulnerable*. World Bank and Agence Française de Développement Publication.

Devereux, S., Masset, E., Sabates-Wheeler, R., Samson, M., Rivas, A.-M., and te Lintelo, D. (2017). The targeting effectiveness of social transfers. *Journal of Development Effectiveness*, 9(2):162–211.

Finn, A. and Ranchhod, V. (2017). Genuine fakes: The prevalence and implications of data fabrication in a large South African survey. *The World Bank Economic Review*, 31(1):129–157.

Fiszbein, A. and Schady, N. R. (2009). *Conditional cash transfers: reducing present and future poverty*. The World Bank.

Friedman, J., Beegle, K., De Weerdt, J., and Gibson, J. (2017). Decomposing response errors in food consumption measurement. *Food Policy*, 72:94–111.

Gibson, J., Beegle, K., De Weerdt, J., and Friedman, J. (2015). What does variation in survey design reveal about the nature of measurement errors in household consumption? *Oxford Bulletin of Economics and Statistics*, 77(3):466–474.

Gibson, J. and Kim, B. (2007). Measurement error in recall surveys and the relationship between household size and food demand. *American Journal of Agricultural Economics*, 89(2):473–489.

Griliches, Z. and Hausman, J. A. (1986). Errors in variables in panel data. *Journal of Econometrics*, 31(1):93–118.

Grosh, M., Del Ninno, C., Tesliuc, E., and Ouerghi, A. (2008). *For protection and promotion: The design and implementation of effective safety nets*. World Bank Publications.

Grosh, M. E. (1994). *Administering targeted social programs in Latin America: From platitudes to practice*, volume 94. World Bank Publications.

Hanna, R. and Olken, B. A. (2018). Universal basic incomes vs. targeted transfers: Anti-poverty programs in developing countries. *Journal of Economic Perspectives*, 32(4):201–226.

Hausman, J. (2001). Mismeasured variables in econometric analysis: problems from the right and problems from the left. *The Journal of Economic Perspectives*, 15(4):57–67.

Houssou, N., Asante-Addo, C., Andam, K. S., and Ragasa, C. (2018). How can African governments reach poor farmers with fertiliser subsidies? Exploring a targeting approach in Ghana. *The Journal of Development Studies*. Forthcoming.

Jolliffe, D. (2001). Measuring absolute and relative poverty: the sensitivity of estimated household consumption to survey design. *Journal of Economic and Social Measurement*, 27(1, 2):1–23.

Karlan, D. and Thuysbaert, B. (2016). Targeting ultra-poor households in Honduras and Peru. *The World Bank Economic Review, Lhw036*.

Kidd, S. and Wylde, E. (2011). Targeting the poorest: An assessment of the proxy means test methodology. Ausaid research paper, Australian Agency for International Development, Canberra, Australia.

Lanjouw, P. and Ravallion, M. (1996). How should we assess poverty using data from different surveys? *Poverty Lines*, (3).

McBride, L. and Nichols, A. (2016). Retooling poverty targeting using out-of-sample validation and machine learning. *The World Bank Economic Review, Lhw056*.

Premand, P. and Schnitzer, P. (2018). Efficiency, legitimacy and impacts of targeting methods: Evidence from an experiment in Niger. Policy Research Working Paper No. 8412. Washington, DC: World Bank.

Scott, C. and Amenuvegbe, B. (1991). Recall loss and recall duration: an experimental study in Ghana. *Inter-Stat*, 4(1):31–55.

Sebastian, A., Shivakumaran, S., Silwal, A. R., Newhouse, D., Walker, T., and Yoshida, N. (2018). A Proxy Means Test for Sri Lanka. Policy Research Working Paper No. 8605. Washington, DC: World Bank.

Stoeffler, Q., Mills, B., and Del Ninno, C. (2016). Reaching the poor: Cash transfer program targeting in Cameroon. *World Development*, 83:244–263.

Sulla, V., Zikhali, P., Schuler, P., and Jellema, J. (2017). Does fiscal policy benefit the poor and reduce inequality in Namibia? World Bank Document, World Bank.

Chapter 2

Cash Transfers and Migration: Experimental Evidence from Comoros

This chapter is a joint work with Eric Mvukiyehe (DIME, The World Bank) and Olivier Sterck (ODID, Oxford University).

2.1 Introduction

Given the widespread promotion of cash transfers to foster development, understanding how income shocks affect international migration is critical, not only for academics who work on related topics, but also for policy-makers who have preferences over migration outcomes. The link between income and international migration is surprisingly complex (McKenzie and Rapoport, 2007; Clemens et al., 2014; Dustmann and Okatenko, 2014). On one hand, migrants need to finance their journey to the destination country. This upfront cost can be very high (Adhikari and Gentilini, 2018), especially for illegal migrants, who often need to pay smugglers and face important risks (Chiswick, 1988; Hanson, 2006). Aspiring migrants facing liquidity and credit constraints are often unable to afford this cost (Bazzi, 2017), despite the very high expected returns to migration (Yang, 2008; McKenzie et al., 2010; Clemens, 2011; Gibson and McKenzie, 2012; Bryan et al., 2014). On the other hand, the opportunity cost of migration increases with income at home (Sjaastad, 1962), which itself depends on human capital. Human capital, in turn, not only affect the expected returns to emigration but also increases as a result of migration (Gibson and McKenzie, 2012). The sum of these opposite effects is theoretically ambiguous, and has been the subject of empirical investigation since Zelinsky (1971) hypothesized the existence of an inverted U-shaped relationship between development and migration.

In line with Zelinsky's hypothesis, researchers using macro-level data have identified a clear inverted U-shaped relationship between income and migration rates (Clemens et al., 2014; Dao et al., 2018). Micro-level evidence is, however, far less conclusive, and mostly focusing on middle-income countries. A few recent micro-level studies explored the relationship between income and migration by exploiting exogenous variation in cash transfer programmes (Adhikari and Gentilini, 2018). These micro-level studies offer mixed results, suggesting that different mechanisms are operating in different contexts. For example, the effect of Mexico's Opportunitades programme on migration to the U.S. seems to depend on which type of migration is considered: while Stecklov et al. (2005) find that the programme reduced overall migration to the U.S., Angelucci (2015) shows that the programme increased labor-induced migration to the U.S. by relieving the credit constraints of eligible households. In India, the NREGA cash-for-work programme reduced short-term migration by increasing the opportunity cost of migrating (Imbert and Papp, 2019). In Bangladesh, small cash transfers increased seasonal migration during the lean season especially for households close to subsistence (Bryan et al., 2014).

In this paper, we study the impact on international migration of a randomized cash transfer intervention targeted at very poor households in Comoros.

We model the decision to migrate and identify four channels through which a cash transfer intervention could affect migration. First, cash transfers relax the budget constraint and can therefore facilitate the migration of households facing a liquidity constraint (liquidity channel). Second, cash transfers that are conditional on remaining in the origin country increase the opportunity cost of migrating and can therefore reduce migration (opportunity-cost channel). Third, cash transfers can facilitate access to credit and thereby increase migration of credit constrained households as soon as they are selected to benefit from cash transfers (credit-constraint channel). Finally, as migration is a risky investment, cash transfers can encourage the migration of individuals whose preferences are characterized by decreasing absolute risk aversion (DARA) while restraining those characterized by increasing absolute risk aversion (IARA) (risk-aversion channel).

In the empirical analysis, we assess the effects of a randomized cash-for-work program in Comoros on international migration. The Comoros Social Safety Net Program (SSNP) was initiated in 2015 by the Government of Comoros and the World Bank. The main component of the SSNP provided temporary cash-for-work (CFW) opportunities to selected poor households. Between the baseline and endline surveys, beneficiary households received up to the equivalent of US$320 in cash conditional on their participation to public work activities.[1]

Migration patterns are salient in Comoros, especially towards Mayotte – the neighboring French Island. A mix of geographic proximity and economic disparities has caused many Comorians to migrate to Mayotte. While Mayotte is located about 70 kilometers to the south-east of Comoros, the GDP per capita in Mayotte is 10 times that of Comoros,[2] and Mayotte has much better public infrastructures. Comorian migrants typically use small fishing boats called *kwassa-kwassa* to reach Mayotte. The journey is both risky and costly, especially since 1995, after France established visa requirements for Comorians traveling to Mayotte, forcing aspiring migrants to use smugglers and illegal sea routes. Thousands of Comorians have died on this often overlooked migration route. The cost of a trip is currently between US$230 and US$1150 depending on the number of persons on the kwassa.[3]

We find that cash windfalls had a sizable and positive impact on migration to Mayotte. The migration rate of beneficiary households increased by about 38 percent (from 7.8% to 10.8%). We rule out alternative causes for the observed increase, including selective attrition and asymmetric indirect treatment effects. We find suggestive evidence that the liquidity and the risk-aversion channels drive the results. In the control group, migration to Mayotte is significantly larger for households with high levels of savings at baseline. In line with the liquidity channel, the effect of the cash transfers on migration to Mayotte is mainly visible for households with low levels of savings at baseline. We also find that control households reporting higher degree of risk-aversion at baseline are less likely to migrate to Mayotte between the survey waves, and that the effect of the cash transfers on migration is only visible for households that are more risk-averse at baseline, which suggests that cash transfers can ease risk-bearing and thereby risky migrations.

By contrast, the opportunity cost channel seems irrelevant in this study. According to the opportunity cost channel, the impact of cash transfers on migration should be negative, which is not what we observe. The opportunity cost channel is only relevant if the cash transfers were conditional upon the whole household staying in Comoros. In practice, the cash-for-work program was very flexible: beneficiary households were entitled to send one adult of their choice to public works. The program was not conditional upon other house-

[1] Throughout the paper, we use an exchange rate 430 KMF (Comorian Franc) for one dollar.

[2] In 2017, the GDP per capita of Comoros in current US$ was US$1312 (World Bank data), while the GDP per capita of Mayotte was US$13,050 USD (authors' calculation based on EUROSTAT, INSEE, and OECD data).

[3] As a comparison, the median annual consumption per capita in our sample is US$460.

hold members staying in Comoros. In fact, we find that migrants and workers are very different. Compared to migrants, household members who participated to the cash-for-work activities are more likely to identify as a woman, are older, less educated, and less likely to have migration experience.

The credit-constraint channel seems negligible in our study. If this channel was operating, the effect of the cash transfer program should have appeared soon after households learned that they were selected into the program. In contrast with this prediction, we find that the effect of transfers on migration takes time to appear, in line with the liquidity channel.

2.2 Conceptual framework

We propose a simple model to disentangle four channels through which a cash transfer intervention could affect migration. First, cash transfers ease the budget constraint and can therefore increase the migration of households facing a liquidity constraint. Second, cash transfers that are conditional on presence in the origin country increase the opportunity cost of migrating and can therefore reduce migration. Third, cash transfers can facilitate access to credit and thereby increase migration of credit constrained households as soon as they are selected to benefit from cash transfers (and even before they receive any cash). Finally, as migration is a risky investment, cash transfers can encourage the migration of individuals whose preferences are characterized by decreasing absolute risk aversion (DARA) while restraining those characterized by increasing absolute risk aversion (IARA).

We study the decision process of a household that can send one of its member abroad to work. The model has two periods, denoted t_1 and t_2. In both periods, the household first decides whether to finance the migration of one of its member. If the member migrates, the household needs to pay the upfront migration costs c using savings s_{t-1} (a credit market will be added in Section 2.2.3). Then, the household earns an income, which is denoted w_o if all members are living in the origin country, and w_d if one member has migrated to the destination country. We assume that migration increases household income ($w_d > w_o$). Migration is therefore seen as an investment. Finally, the household decides how much of the income and savings to consume and to save for the next period. Household savings are denoted s_t ($s_t \geq 0$). Without loss of generality, we assume that the household stays if it is indifferent between staying or migrating.

We abstract from the decision to smooth consumption over time by assuming that the utility function of the household is a function of lifetime wealth

($u' > 0$, $u'' \leq 0$). Without this assumption, there is no closed-form solution when risk is included in the model.

The household has to compare three options: investing in migration in t_1 (Case 1), investing in migration in t_2 (Case 2), or not investing in migration (Case 3). The lifetime utilities associated with these cases are:

$$
\begin{aligned}
U^{Case1} &= u(s_0 - c + 2w_d) \\
U^{Case2} &= u(s_0 - c + w_o + w_d) \\
U^{Case3} &= u(s_0 + 2w_o)
\end{aligned}
$$

The following proposition characterizes the decision to finance the migration of a household member.

Proposition 1. *A household member migrates in t_1 if and only if migration can be financed in t_1 (inequality (2.1)) and if the benefit of migrating in t_1 is larger than the cost (inequality (2.2)):*

$$
\begin{cases}
s_0 \geq c. & (2.1) \\
2(w_d - w_o) > c & (2.2)
\end{cases}
$$

A household member migrates in t_2 if and only if migration can be financed in t_2 but not in t_1 (inequality (2.3)) and if the benefit of migrating in t_2 is larger than the cost (inequality (2.4)):

$$
\begin{cases}
c - w_o \leq s_0 < c. & (2.3) \\
w_d - w_o > c & (2.4)
\end{cases}
$$

Proof in Appendix B1.

The possible outcomes are represented in Figure 2.1 as a function of the wage differential $w_d - w_o$ and of initial savings s_0. In words, a member migrates in t_1 if savings are large and if the return to migration is intermediate or large. A member migrates in t_2 if savings are intermediate and if the return to migration is large.

In the next sections, we study how a cash transfer can affect the decision-making process, distinguishing four scenarios: an unconditional cash transfer (Section 2.2.1), a cash transfer conditional on not migrating (Section 2.2.2), an unconditional cash transfer with a functioning credit market (Section 2.2.3), and an unconditional cash transfer in the presence of risk and risk-aversion (Section 2.2.4).

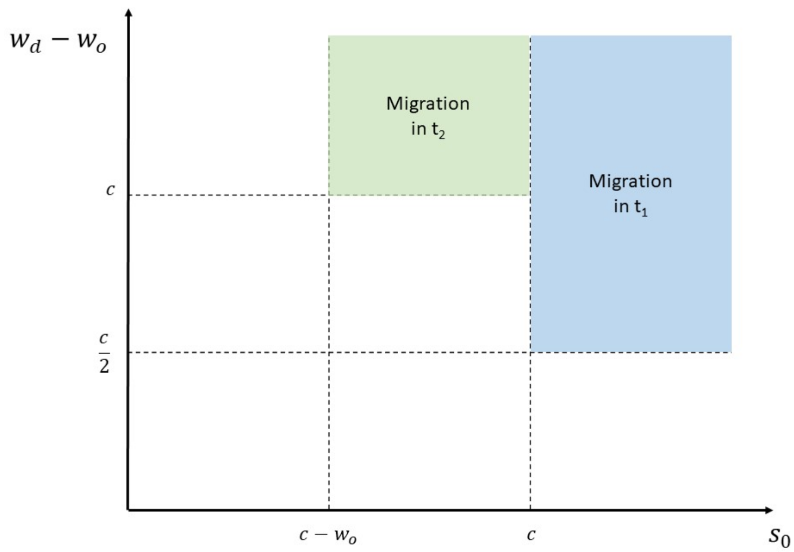


Figure 2.1: Outcomes of the benchmark model as a function of the wage differential $w_d - w_o$ and of initial savings s_0

2.2.1 The liquidity channel

In this first extension of the benchmark model, we assume that the household is selected to receive an unconditional cash transfer $\tau > 0$ at the end of t_1. This extra wealth can be consumed or saved. While the utility returns from migration are not affected by the cash transfer, as the cash transfer is unconditional, the budget constraint (2.3) is eased by the cash transfer. The cash transfer modifies the decision to migrate as follows.

Proposition 2. *While the unconditional cash transfer does not affect decision to migrate in t_1, it facilitates migration in t_2 by easing the budget constraint. In particular, a household member migrates in t_2 if:*

$$\begin{cases} c - w_o - \tau \leq s_0 < c. & (2.5) \\ w_d - w_o > c & (2.6) \end{cases}$$

Proof in Appendix B1.

It is clear that inequality (2.3) is more stringent that inequality (2.5): the amount τ eases the budget constraint of the household in t_2, as illustrated in Figure 2.2. The cash transfer allows the migration of households that would be liquidity constrained without the transfer but that are able to finance migration in t_2 thanks to the transfer.

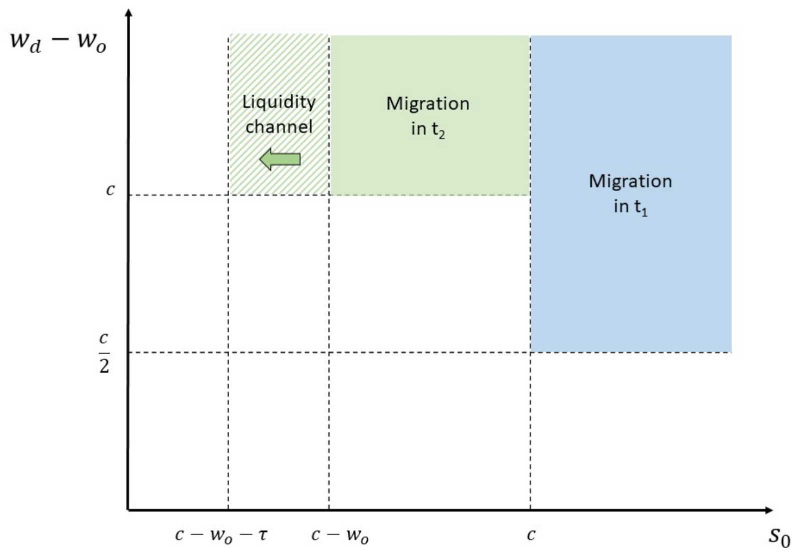


Figure 2.2: Effect of a cash transfer through the liquidity channel

2.2.2 Opportunity cost channel

We examine the effect of adding a conditionality to the cash transfer. If the cash transfer is conditional on all household members working in the origin country at t_1, households that would have been migrating in t_1 without the conditionality cancel or postpone migration if the value of the cash transfer is larger than the cost of canceling or delaying migration. Compared to the benchmark model, the conditional cash transfer does not affect the lifetime utility of migrating in t_1, but it increases the lifetime utility of migrating in t_2 and the lifetime utility of not migrating at all.

The following proposition describes when the household finances the migration of one of its member in the presence of a conditional cash transfer.

Proposition 3. *In the presence of a conditional cash transfer, a household member migrates in t_1 if and only if:*

$$\begin{cases} s_0 > c. & (2.7) \\ w_d - w_o > Max(\frac{c+\tau}{2}, \tau) & (2.8) \end{cases}$$

A household member migrates in t_2 if conditions (2.5) and (2.6) are satisfied, or if:

$$\begin{cases} s_0 > c. & (2.9) \\ c < (w_d - w_o) < \tau & (2.10) \end{cases}$$

Proof in Appendix B1.

The effect of the conditional cash transfer is illustrated in Figures 2.3a and 2.3b. On the one hand, the cash transfer increases households' ability to finance migration in t_2 (liquidity effect). On the other hand, the conditionality increases the opportunity cost of migrating in t_1. It affects the decision of households able to finance migration at t_1 ($s_0 > c$) and for which the wage differential $w_d - w_o$ is lower than the transfer τ. These households are either prevented from migrating (if the wage differential is low such that migration in t_2 is not optimal) or they postpone migration until t_2 (if the wage differential is large enough to cover the cost of migration in t_2).

2.2.3 Credit constraint channel

So far, we have assumed that credit markets are absent. The presence of an effective credit market would ease the budget constraint of households, who can borrow to finance migration in t_1 and pay back the loan in t_2 thanks to the wage differential $w_d - w_o$. We assume that households can borrow a maximum amount $B \geq 0$ at the beginning of t_1. The loan needs to be repaid at the end of t_2 with an interest rate r.

In the presence of such credit market (and in the absence of cash transfer), a liquidity-constrained household borrows in t_1 to finance the migration of one of its member in t_1 if the following conditions are jointly satisfied.

Proposition 4. *The household borrow to finance migration in t_1 if:*

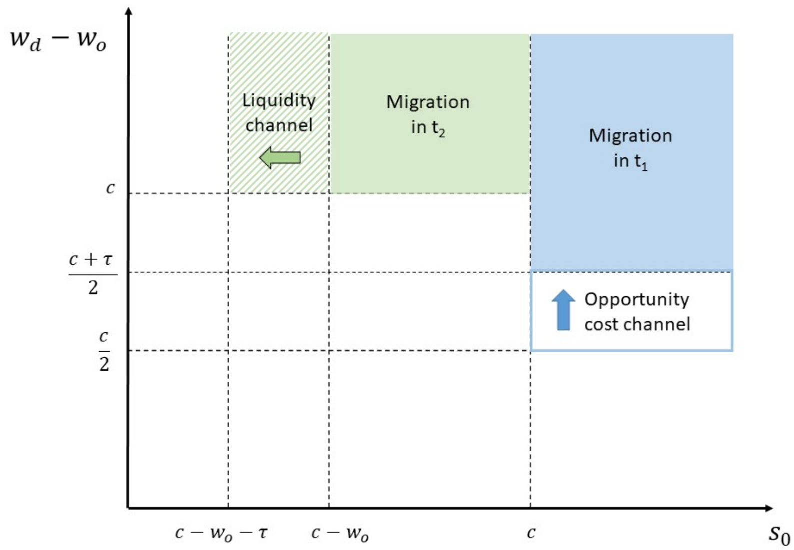
$$\begin{cases} c - B \leq s_0 < c. & (2.11) \\ w_d - w_o > Max(r(c - s_0), \dfrac{c + r(c - s_0)}{2}) & (2.12) \end{cases}$$

Proof in Appendix B1.

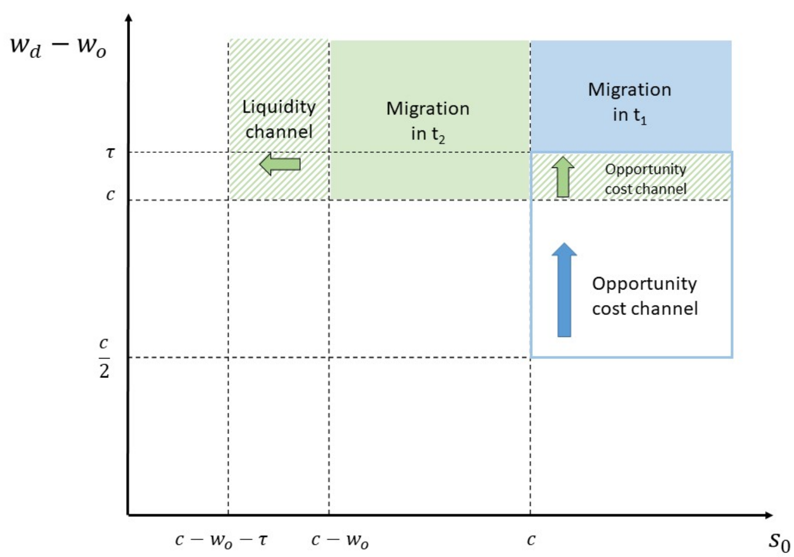
Thanks to the credit market, a household facing a liquidity constraint in t_1 can finance migration through borrowing if the maximum amount of the loan B and initial savings s_0 are large enough to cover the upfront cost of migration c (Figure 2.4a).[4] However, because borrowing is costly, inequality (2.12), which is represented by the yellow lines on Figure 2.4a, is more stringent that inequality (2.6).

[4] In Figure 2.4a, we assume that $B < w_o$. It is indeed very unlikely that a lender would provide loans that are larger than the net present value of future income in the origin country, $\frac{w_o}{1+r}$. Results are qualitatively similar if $B \geq w_o$: instead of self-financing migration in t_2, households with $c - B \leq s_0 < c$ borrow to finance migration in t_1.

Figure 2.3: Effect of a conditionality



(a) Small cash transfer $\tau < c$



(b) Large cash transfer $\tau > c$

With a functioning credit market, an unconditional cash transfer has three effects on the decision to finance migration, as illustrated in Figure 2.4b. In line with the liquidity channel described in Section 2.2.1, the direct effect of an unconditional cash transfer is to ease the budget constraint in t_2, which facilitates migration in t_2 for households with intermediate levels of savings ($c - w_o - \tau \leq s_o \leq c - w_o$). But in the presence of a functioning credit market, a cash transfer can have two other indirect effects. First, the maximum amount that households can borrow, B, is likely to increase as soon as households are selected to benefit from the unconditional cash transfer, as the guaranteed future income stream can play the role of a collateral. If the maximum amount of the loan, B, is increased, more households are able to finance migration in t_1 through borrowing. Second, the interest rate of the loan r is likely to be reduced because the risk of default is reduced by the increase in future income. If the interest rate r is reduced, more households find it optimal to borrow to finance migration in t_1.

2.2.4 Risk-aversion channel

In this section, we modify the benchmark model and assume that migration is risky. With a probability $p \in]0,1[$, the migrant reaches its destination and the household income is w_d. With a probability $1-p$, the migrant's journey is unsuccessful, and the household income is w_o. In the presence of risk, the degree of risk aversion of the household will influence the decision-making process. Risk aversion means that the utility function is concave ($u'' < 0$), which implies that households dislike zero-mean risks.

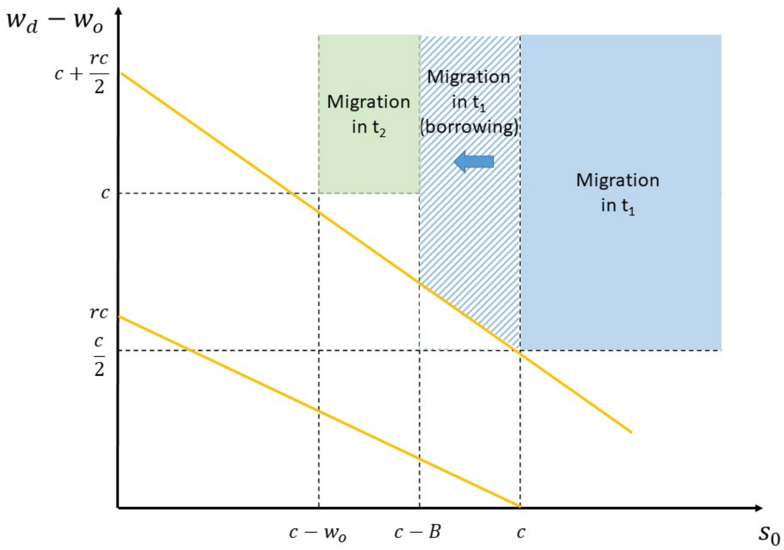
We introduce risk aversion in the model using the concepts of certainty equivalent and risk premium (Eeckhoudt et al., 2005; Myerson, 2005). The certainty equivalent of a gamble for a decision-maker is the sure amount of money that the decision-maker would be willing to accept instead of the gamble. The difference between the expected monetary value of the gamble and the certainty equivalent of the gamble is called the risk premium.[5] In the presence of risk and risk aversion, the household finances the migration attempt of one of its members if the following conditions are satisfied.

Proposition 5. *If migration is risky, a household member attempts to migrate in t_1 if:*

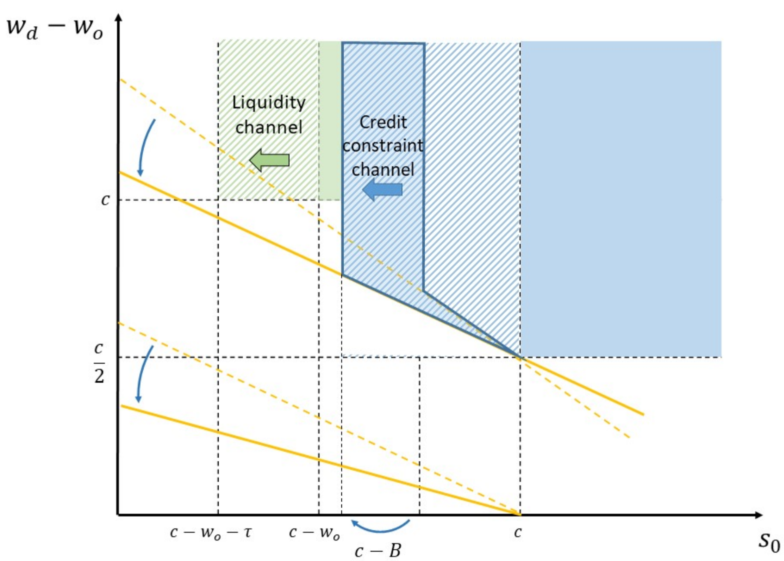
$$\begin{cases} s_0 > c. & (2.13) \\ 2p(w_d - w_o) > c + \pi_1 & (2.14) \end{cases}$$

[5] For a small risk, the risk premium π can be approximated as: $\pi \approx 1/2\sigma^2 A(w)$ where σ is the variance of the gamble, and $A(w) = -u''/u'$ is the degree of absolute risk aversion of the decision-maker, which is a function of wealth w. For a large risk, the risk premium also depends upon the other moments of the distribution of the risk, not just its mean and variance.

Figure 2.4: Decision to finance migration using credit



(a) Without cash transfer



(b) With a cash transfer τ

where π_1 is the risk premium associated with financing migration in t_1.[6]

A household member attempts to migrates in t_2 if:

$$\begin{cases} c - w_o < s_0 < c. & (2.15) \\ p(w_d - w_o) > c + \pi_2 & (2.16) \end{cases}$$

where π_2 is the risk premium associated with financing migration in t_2.[7]

Proof in Appendix B1.

The presence of risk has two effects, which are illustrated in Figure 2.5a. First, risk reduces the expected benefit from migration by a factor p. Second, risk aversion reduces households' expected utility of migrating, as captured by the risk premiums π_1 and π_2.

In the presence of risk and risk-aversion, a cash transfer not only impacts the budget constraint (liquidity channel), but also the expected utility returns from migration. The cash transfer is an income shock. Therefore, the direction of the impact depends on how risk aversion varies with income, as summarized in the following proposition.

Proposition 6. *An unconditional cash transfer eases the budget constraint in t_2 (liquidity channel). Furthermore, if the utility function of the household is characterized by decreasing absolute risk aversion (DARA), the unconditional cash transfer increases the expected utility returns from investing in migration. By contrast, if the utility function is characterized by increasing absolute risk aversion (IARA), the unconditional cash transfer reduces the expected utility returns from investing in migration.*

Proof in Appendix B1.

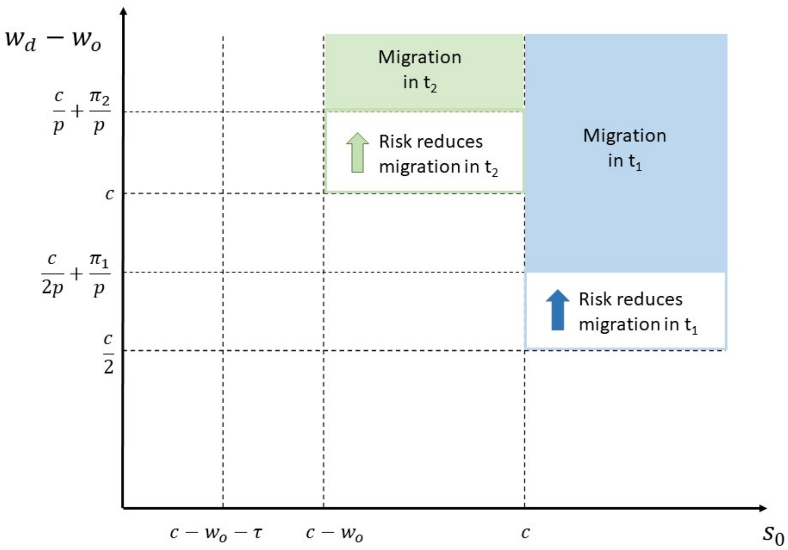
Experimental and empirical evidence supports the hypothesis of decreasing absolute risk aversion.[8] In Figure 2.5b, we illustrate the effect of the cash transfer when the utility function is DARA. The direct effect of the transfer is to ease the budget constraint of the household in t_2 (liquidity channel). But if migration is risky and if the utility function is DARA, the cash transfer also reduces risk aversion, thereby increasing the expected utility returns from financing migration in t_1 or in t_2.

[6] The risk premium π_1 is defined as: $u(s_o - c + 2pw_d + 2(1-p)w_o - \pi_1) = p[u(s_0 - c + 2w_d)] + (1-p)[u(s_0 - c + 2w_o)]$.

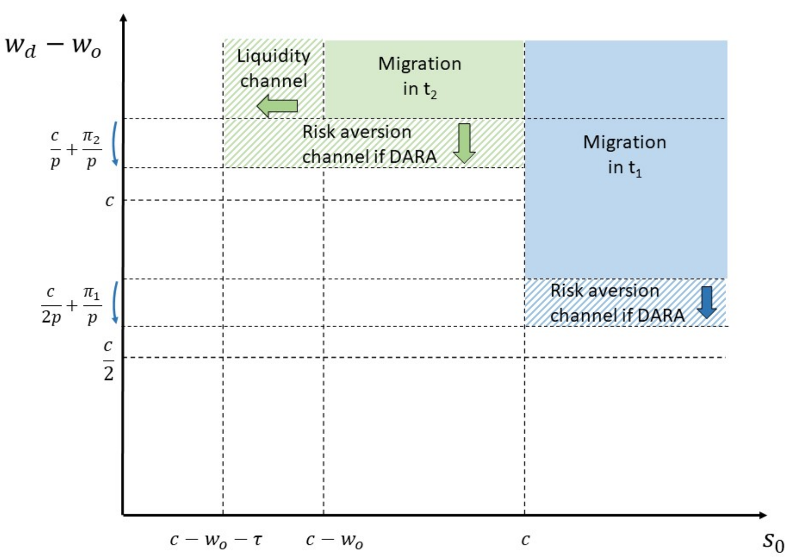
[7] The risk premium π_2 is defined as: $u(s_o - c + w_o + pw_d + (1-p)w_o - \pi_2) = p[u(s_0 - c + w_o + w_d)] + (1-p)[u(s_0 - c + 2w_o)]$.

[8] See e.g. (Dohmen et al., 2011, 2010; Yesuf and Bluffstone, 2009; Guiso and Paiella, 2008; Wik* et al., 2004; Levy, 1994).

Figure 2.5: Decision to migrate if migration is risky



(a) Effect of risk on the decision to migrate



(b) Effect of an unconditional cash transfer with DARA

2.3 Background of the cash-for-work program

2.3.1 Context

The Comoro archipelago consists of four islands located in the Mozambique Channel, between Mozambique and Madagascar (see Figure 2.6). Three islands belong to the Union of Comoros (Comoros henceforth), a poor country with a population of 820,000 people. The remaining island, Mayotte, is French. Mayotte is situated about 70 kilometers to the south-east of Comoros. Mayotte has a population of 260,000 people. The GDP per capita in Mayotte is more than 10 times that of Comoros.

Strong ties unite the four islands. During the French colonisation, the islands were unified under a single administration and placed under the authority of the French colonial governor of Madagascar. People share a similar language, *Shikomori*, and are predominantly Muslim.[9] They also have similar social structures such as a matrilineal system shaped by the informal institution of the *Grand mariage* – a determinant of social status whose completion greatly increases one's standing in society.

However, during the 1974 independence referendum, Mayotte voted to remain politically a part of France while other islands voted for independence and formed the Comoros nation.[10] Since then, Mayotte has been continuously administered by France and even became a French overseas department in 2011.[11] Socioeconomic conditions have steadily improved in Mayotte while stagnating in neighbouring Comorian islands. Since independence, Comoros has experienced recurring political crises and conflict between the islands. Comoros' low GDP per capita (US$1,312 in 2017) is stagnating because of relatively low GDP growth rates (between 2 and 3.5 percent) and high population growth (2.4 percent). Poverty is high with 42 percent of the population living with incomes below US$1.90 per day, and one-third of all children under five years of age suffering from chronic malnutrition. Although Mayotte is the poorest French department, its US$13,050 GDP per capita in 2017 is extremely attractive relative to Comorian standards.

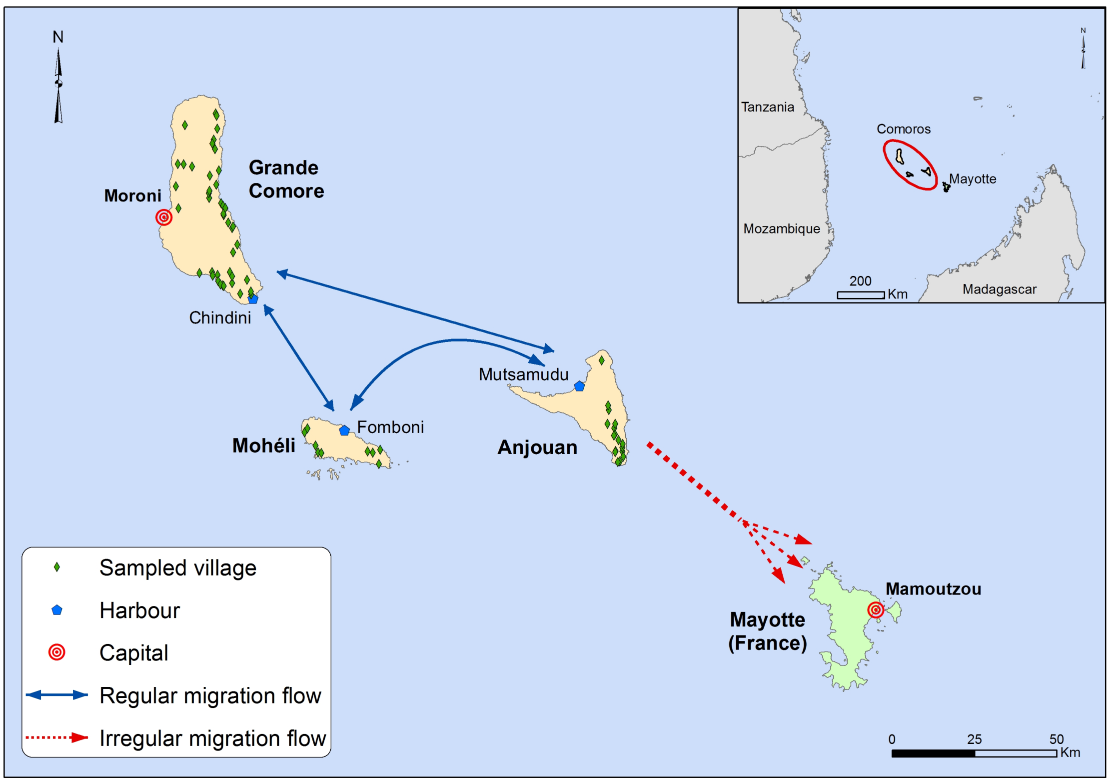
In order to control migration of Comorians to Mayotte, France issued strict visa requirements in 1995. However, illegal sea routes and people smuggling

[9] Slightly different variants of *Shikomori* are found on each of the four islands (*Shingazidja*, *Shimwali*, *Shinzwani* and *Shimaore*) but people can easily communicate.

[10] See Blanchy (2002) for a discussion on why the people of Mayotte decided to remain French.

[11] France has vetoed several United Nations Security Council resolutions that would affirm Comorian sovereignty over Mayotte.

Figure 2.6: Migration route to Mayotte



Source: Authors' elaboration

emerged such that the flow of Comorian migrants has never stopped. In 2015, it was estimated that 61 percent of the population in Mayotte had a connection to Comoros, with 42 percent born in Comoros, and an additional 19 percent having a Comorian mother (Marie et al., 2017). The routes used by Comorian migrants are depicted in Figure 2.6. Migrants typically converge to the south-east of Anjouan and then use small fishing boats (called *kwassa-kwassa*) to reach Mayotte. While this migration route has attracted scant international attention, tragic accidents of kwassas occasionally make the headlines of French newspapers (Le Monde, 2017). French Parliamentary reports usually cite figures ranging from several hundreds to one thousand deaths per year on this migration route (Sénat, 2001, 2008, 2012). However, because an important number of migrant deaths probably go unrecorded, there is no official record on the number of fatalities.

The qualitative survey provides sobering evidence on these migration flows. All respondents reported that using a kwassa from Anjouan to Mayotte is the only migration technology available for them. They perceive it as particularly risky and have many friends or relatives who died from a migration attempt to Mayotte (often in recent years). As a respondent put it: "*There is only one way to go to Mayotte, it is to take a kwassa. Only people with a normal situation can travel by plane or boat. The journey is so difficult and risky. I know many people who have lost their lives in this sea. The number of people in this village who died because of Mayotte is uncountable*". Although in theory legal migration pathways exist, in practice, respondents reported that the likelihood to get a visa to Mayotte is close to zero for poor Comorians. Migration costs are relatively high and typically depend on the number of persons in the kwassa (the more persons, the higher the price of the journey). Our qualitative evidence suggests that migration costs can go from a minimum of US$230 – if the kwassa is overloaded – to a maximum of US$1150 for a "VIP kwassa" (i.e. a kwassa with only a few persons). Migrants generally finance these costs through their savings, the sale of livestock, and the help of relatives. In addition to the risk of dying en route, migrants face substantial risks of being arrested and expelled by the French police. According to official French statistics, each year, more than 20,000 illegal migrants are deported to Comoros. Several respondents to the qualitative survey alluded to these risks and there consequences, as reflected in this quote: "*Sometimes, we sell high-value properties to pay transportation costs, and unfortunately we get arrested by the police and have to start again from scratch. In these cases, we are in a depressing situation with nameless regrets*".

2.3.2 The Comoros Social Safety Net Program (SSNP)

The SSNP was initiated in 2015 by the Government of Comoros and the World Bank. The main implementing agency was FADC (*Fonds d'Appui au Dveloppement Communautaire*) – renamed ANACEP (*Agence Nationale de Conception et d'Exécution de Projets*) in 2017. Prior to running this program, FADC had successfully implemented a variety of World Bank projects, including similar cash-for-work programs. The objectives of the SSNP were to improve poor communities' access to safety nets and nutrition services, smooth consumption, and support the development of productive activities. While there was no explicit targets on emigration to Mayotte due to the political sensitivity of the topic, there was implicitly the hope that the program would deter migration flows.

The main component of the program provided cash-for-work (CFW) opportunities to poor households, i.e. cash transfers conditional on their participation in public works such as reforestation, water management, and terracing. Beneficiary households were entitled to send one able-bodied adult of their choice to public works. Households with no able-bodied adults received unconditional cash transfers. Cash-for-work activities were implemented in periods of 20 days with payments made at the end of each period by a local micro-credit institution known as MECK. From 2016 to 2018, households have been provided with an average 60 days of work per year at the wage rate of US$2.3 for 4 hours of work per day.

A total of 69 rural villages were selected by FADC to receive the intervention. According to the national distribution formula, Grande Comore received 45 percent of the program funds, while Anjouan 42 percent, and Moheli 13 percent. Based on these percentages, FADC selected the poorest villages using the poverty map drawn up by the Comorian national institute of statistics (known as INSEED) in 2003/2004. In Table 2.1, we see that households of selected villages are much poorer than households of non-selected villages, with an overall poverty rate of 88.2% against 42.1%.

Table 2.1: Poverty rates in treated villages

	Non-CFW villages		CFW villages	
	Pop (hh)	Poverty rate	Pop (hh)	Poverty rate
Grande Comore	42,744	41.3%	5,435	80.6%
Anjouan	38,152	41.5%	4,778	95.6%
Moheli	4,987	55.0%	1,097	94.8%
Total	85,883	42.1%	11,310	88.2%

Notes: Authors' calculations based on the 2003/04 poverty mapping.

Within villages, the selection of beneficiaries relied on several steps and mechanisms. First, self-targeting was expected because of the labor requirement, the (non-monetary) front costs of applying, and the low wage rate for the public works. Second, village committees, in collaboration with project staff, applied specific selection criteria. There were 4 criteria and each could give one point to the household: (i) the household head attended primary school at most; (ii) the household has at least 4 children below 15 years of age; (iii) the household has children aged between 6 and 14 who are not enrolled in school; (iv) the household has no agricultural field. Based on these criteria, committees pre-selected the poorest 60 percent households in their villages. As there were more pre-selected households than CFW opportunities, the selection of beneficiaries lastly relied on a public lottery organized by committees and FADC's staff in each village.

2.4 Experimental design and data

2.4.1 Empirical strategy

The impact evaluation has been designed as a multi-level randomized control trial. At the household level, beneficiaries were randomly selected from the group of 60% households who had been pre-selected by local committees (see Section 2.3.2 above). At the cluster level, in order to assess indirect effects, villages were randomly assigned to receive a low or high intensity of the treatment. Specifically, in each village, one third or two thirds of the pre-selected were randomly assigned to the treatment. This means that overall 20% or 40% of eligible households were selected.[12]

These two levels of random assignment are core to the empirical strategy. Because of the random assignments, households and villages with different treatment conditions are similar in expectation in every respect except for their treatment status. Any difference in outcome between treatment and control groups after the program can thus be attributed to the difference in treatment. Below, we provide more details on how we estimate the direct, indirect and heterogeneous intention-to-treat (ITT) effects of the SSNP on migration, as outlined in our pre-analysis plan.[13]

[12] The evaluation design also had a gender component. Households with both male and female potential workers chose one individual of each gender to be the potential worker. Then for these households, the gender of the main worker was randomly selected. In practice, however, the rule that the main worker should participate to the works was never enforced: households had eventually lots of flexibility to send the person of their choice, as they could replace the main worker on their own. Ultimately, the majority of households sent a female worker as the daily wage rate was mostly attractive attractive for them. For this reason, the analysis of the gender randomization face power issues and is mostly inconclusive. Results, available upon request, are not reported in this paper due to space limitation.

[13] Our pre-analysis plan is available here: http://egap.org/registration/5302. Section B2

Direct effects

First we estimate a regression equation of the following form to derive direct effects of the program:

$$y_{iv} = \beta_0 + \beta_1 CFW_{iv} + \beta_2 \boldsymbol{X}_{iv} + \varepsilon_{iv} \tag{2.17}$$

Where y_{iv} is the outcome of interest for household i in village v; CFW_{iv} is a dummy indicating whether household i in village v was assigned to the treatment group or not; $\boldsymbol{X}_{iv}$ is a vector of imbalanced covariates at baseline;[14] and ε_{iv} is the disturbance term. The direct effects of the program on the outcomes of beneficiaries are given by the coefficient β_1.

Indirect effects

Indirect average treatment effects (ITE) of the SSNP are ascertained by comparing the outcomes of households in high intensity villages with those of households in low intensity villages. Specifically, we estimate an equation of the following form:

$$y_{iv} = \beta_0 + \beta_1 CFW_{iv} + \beta_2 P_{40v} + \beta_3 CFW_{iv} * P_{40v} + \beta_4 \boldsymbol{X}_{iv} + \varepsilon_{iv} \tag{2.18}$$

Where y_{iv} is the outcome of interest for household i in village v; CFW_{iv} is a dummy indicating whether household i in village v was assigned to the treatment group or not; P_{40v} is a dummy variable at the village level indicating an assignment rate of 40% in village v; $CFW_{iv} * P_{40v}$ is thus a dummy for being assigned to treatment in a village with an assignment rate of 40%; $\boldsymbol{X}_{iv}$ is a vector of imbalanced covariates at baseline (and balanced covariates if specified); and ε_{iv} is the disturbance term.

Equation 2.18 provides an estimation of ITE on both beneficiary and non-beneficiary households. ITE among non-beneficiary households are estimated by the parameter β_2, that is the effect of being assigned to the control group in a village where 40% of the eligible population was assigned to treatment, compared to being assigned to the control group in a village where only 20% of the eligible population was assigned to treatment. ITE among beneficiary

presents sub-analysis that were specified in the PAP but are not incorporated in the paper due to space limitation.

[14] In further specifications, $\boldsymbol{X}_{iv}$ will also include a set of balanced covariates at baseline and island fixed effects to improve precision.

households are given by $\beta_2 + \beta_3$, that is the effect of being assigned to treatment in a village where 40% of the eligible population was assigned to treatment, compared to being assigned to treatment in a village where only 20% of the eligible population was assigned to treatment.

Heterogeneous effects

Finally, we estimate heterogeneous effects with an equation of the following form:

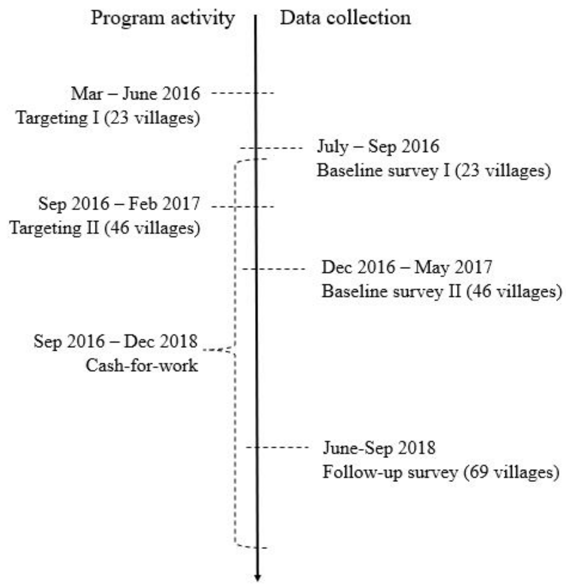
$$\begin{aligned} y_{iv} = \beta_0 + \beta_1 CFW_{iv} + \beta_2 CHARACTERISTIC_{iv} + \\ \beta_3 CFW_{iv} * CHARACTERISTIC_{iv} + \beta_4 \boldsymbol{X}_{iv} + \varepsilon_{iv} \end{aligned} \tag{2.19}$$

Where y_{iv} is the outcome of interest for household i in village v; CFW_{iv} is a dummy indicating whether household i in village v was assigned to the treatment group or not; $CHARACTERISTIC_{iv}$ corresponds to the dimension of heterogeneity studied for household i in village v; $CFW_{iv} * CHARACTERISTIC_{iv}$ is their interaction; $\boldsymbol{X}_{iv}$ is a vector of imbalanced covariates at baseline (and balanced covariates if specified); and ε_{iv} is the disturbance term. This equation tests whether the effects of the program is conditional on baseline characteristics. Because these baseline characteristics were of course not randomly allocated across households, this analysis of heterogeneous effects should be considered as exploratory and results should not be interpreted as causal. In order to limit omitted-variable concerns, we will include interaction terms of the dimension studied with other baseline characteristics: $\boldsymbol{X}_{iv} * CHARACTERISTIC_{iv}$.

2.4.2 Data

The sample is composed of the villages benefiting from the SSNP, with each village considered as statistical domains. Villages with population below 30 households were excluded from the experimental design because the number of beneficiaries would have been too small to conduct the public works. In these villages, 100% of the eligible households participated in the public works. The final sample is composed of 62 villages, including 37 villages from Grande Comore, 16 villages from Anjouan and 9 villages from Moheli. We performed power calculation exercises to determine the optimal number of households to include in the sample in order to both measure the impacts of CFW activities and minimize survey budget. In each village, we sampled 25 beneficiary households and 15 pre-selected but non-beneficiary households. Each households within a given village and category had the same probability of being sampled.

Figure 2.7: Timeline diagram



Source: Authors' elaboration

A baseline survey was conducted after household randomization and before the launch of CFW activities. The baseline survey took place in two phases to mirror program implementation timeline:[15] (i) from July to September 2016 in one third of the villages and (ii) from December 2016 to May 2017 in the remaining two thirds. A follow-up survey was conducted between July and September 2018, while treated households received between US$140-320. Household attrition was low (about 4 percent of the baseline sample) and balanced across treatment and control groups.[16] INSEED, the national institute of statistics, was responsible for data collection and worked under the supervision of the authors. Enumerators were not informed of the treatment status of households prior to the interviews, and could thus only infer this information from questions related to CFW activities in the last module of the follow-up survey.

We implemented a qualitative survey as a complement to the quantitative survey. While the quantitative survey can provide rigorous estimates of impact, it is limited in its explanatory power to determine the mechanisms through which that impact occurred. Qualitative research is also useful to study perceptions, norms, and narratives, which are complex and difficult to quantify. About one hundred qualitative interviews have been conducted by local research assistants under the supervision of the authors. The sample in-

[15]The sampling frame required the completion of the targeting process, which was implemented by FADC in two phases due to capacity constraints.

[16]Attrition will be discussed in more details in Section 2.5.3.

cluded a broad range of actors, including (i) participants and non-participants in project activities, (ii) government officials and local community leaders, and (iii) NGOs and local firms in charge of the execution of CFW activities. In particular, the sample frame included 10 villages (4 in Grande Comore, 4 in Anjouan, and 2 in Moheli), with 6 beneficiaries (2 males, 2 females, and 2 persons belonging to migrant households), 2 non-beneficiaries, and 1 community leader in each village.

Table 2.2 summarizes key baseline variables and tests for balance between treatment and control groups. The first four columns report subsample means and standard deviations, and the last two columns report the difference and associated p-values. Migration experience corresponds to the total number of attempts made by household members. We follow De Brauw and Carletto (2012) and proxy migration network using a dummy equals to one if the household head has one children residing in Mayotte. Only one of the 20 variables tested has an imbalance significant at the 10% level. Household heads assigned to treatment are slightly less likely to have completed primary school only (19% vs. 22%).

Table 2.2: Household characteristics at baseline

	Control		Treatment			
	Mean	SD	Mean	SD	Diff	p-value
Household size	6.55	2.80	6.57	2.82	-0.01	0.91
Consumption (PAE)	7.17	1.02	7.14	0.97	0.03	0.55
Has a bank account	0.28	0.45	0.27	0.45	0.01	0.64
Has an income generating activity (other than agriculture)	0.48	0.50	0.45	0.50	0.03	0.17
Owns fields	0.76	0.43	0.75	0.43	0.01	0.72
Livestock (tropical unit)	0.49	0.93	0.52	0.99	-0.03	0.48
Has electricity	0.59	0.49	0.60	0.49	-0.01	0.50
Has a private water access	0.63	0.48	0.63	0.48	0.01	0.74
Head is male	0.77	0.42	0.76	0.43	0.01	0.59
Head age	48.66	16.03	48.34	15.20	0.32	0.63
Head education						
Did not complete primary	0.56	0.50	0.58	0.49	-0.02	0.39
Primary	0.22	0.41	0.19	0.39	0.03	0.06*
Secondary	0.18	0.38	0.19	0.39	-0.01	0.48
Tertiary	0.04	0.20	0.04	0.20	-0.00	0.83
Willingness to migrate to Mayotte	0.21	0.41	0.23	0.42	-0.02	0.31
Migration experience to Mayotte	0.54	1.33	0.51	1.53	0.03	0.61
Migrant network in Mayotte	0.13	0.34	0.14	0.35	-0.01	0.49
Island of residence						
Ngazidja	0.58	0.49	0.58	0.49	0.00	0.84
Ndzuani	0.27	0.44	0.28	0.45	-0.01	0.58
Mwali	0.15	0.36	0.15	0.35	0.01	0.67
Debts	8.34	5.51	8.35	5.50	-0.01	0.96
F-test joint significance						0.78
Observations	900	900	1372	1372	2272	2272

Notes: This table reports subsample means with standard deviations. The last column reports the pvalue of a ttest of mean equality across subsamples. The F-test corresponds to a regression of the treatment on baseline characteristics using the same specification as in the subsequent analysis (*omnibus test*). An inverse hyperbolic sine transformation has been applied to consumption and debts. PAE denotes per adult equivalent.

2.5 Results

2.5.1 Program take-up

In Table 2.3, we check that households assigned to treatment were indeed more likely to perform CFW activities and test whether they saw an improvement of their levels of employment and income. On one hand, access to CFW opportunities should directly increase employment and income levels of beneficiaries. On the other hand, substitution effects could undermine these direct effects (e.g. if beneficiaries gave up on other profitable activities because of the labor requirement of the program). Our main outcome variables aggregate individual measures of employment and incomes at the household level. A 30 days recall period has been used in order to limit the scope for measurement errors.

Table 2.3: Treatment effects on labor market outcomes

	Employment			Income		
	(1) CFW	(2) Total (excl. CFW)	(3) Total (incl. CFW)	(4) CFW	(5) Total (excl. CFW)	(6) Total (incl. CFW)
Panel A						
Treatment	4.969***	0.490	5.459***	1.281***	-0.248**	1.033***
	(0.317)	(1.622)	(1.668)	(0.074)	(0.115)	(0.138)
Extended controls						
Island FE						
Panel B						
Treatment	4.905***	0.472	5.377***	1.265***	-0.235**	1.029***
	(0.315)	(1.540)	(1.583)	(0.074)	(0.109)	(0.131)
Extended controls	✓	✓	✓	✓	✓	✓
Island FE						
Panel C						
Treatment	4.918***	0.379	5.297***	1.268***	-0.239**	1.029***
	(0.313)	(1.510)	(1.556)	(0.073)	(0.107)	(0.129)
Extended controls	✓	✓	✓	✓	✓	✓
Island FE	✓	✓	✓	✓	✓	✓
Control mean	1.881	51.924	53.805	0.489	3.098	3.587
Observations	2181	2181	2181	2181	2181	2181

Notes: This table shows estimates of equation (2.17) using various employment and income variables as outcome variables. Employment variables are expressed as number of days worked. Total employment includes farming, livestock rearing, fishing, and other activities (and CFW if specified in the column header). An inverse hyperbolic sine transformation has been applied to all income variables. All estimates control for unbalanced covariates. Standard errors in parentheses are clustered at the level of the treatment (household). *** p<0.01, ** p<0.05, * p<0.1.

In column 1 and 4, we see that the randomization was effective at driving treated households to participate in CFW activities. Households randomly assigned to the treatment worked significantly more likely to participate in public works than control households (p<0.001). Some evidence of substitution effects can be seen from column 5. Excluding cash-for-work income, treated households earned a lower total income than their control counterparts. This substitution effect is only visible for income, and not sufficient to remove CFW positive direct effects. Overall, the total treatment effects on employment and income are substantial and positive (columns 3 and 6), such that the program can be considered as a large positive income shock. The estimates are similar when extended controls and island fixed effects are included (panels B and C).

The control group appears to have been slightly contaminated by the treatment. Control households reported an average of 1.88 days spent in public works. We further explore program take-up by looking at the treatment status reported by endline respondents themselves.[17] We find a non-compliance rate of 19.6% overall (14.7% in the treatment group; 27.2% in the control group). The main explanation for non-compliance is related to the replacements of beneficiaries dropping out. For example, a respondent to the qualitative survey reported that *"after a month, I received the 20,000 KMF [US$46] and decided to go back to my own farming because it was more profitable. My wife also didn't want to go to the public works. Then, another person took our place. I saw that the program was not going to help me on much"*.

As mentioned in Section 2.4.1, we will use ITT estimates in order to avoid biasing our evaluation of program effects. In robustness analysis, we will use the treatment randomly assigned as an IV for the treatment status actually observed in the survey to obtain local average treatment effects (LATE) of the program, i.e. the impact of the program on compliers.

2.5.2 Impact on migration

The main results of the paper are presented in Table 2.4, where we report the ITT effect of the SSNP on migration to Mayotte. We were concerned about the sensitivity of the topic because migration of Comorians to Mayotte is usually illegal, especially for the study population which is poorer than the average Comorian and has a tiny probability of getting visas. In addition, many people have died in the last few decades trying to reach Mayotte and development agencies are increasingly concerned by the phenomenon. In terms of identification, experimenter demand effects and socially desirable answers could induce beneficiary households to be more reluctant to reveal they sent migrants to Mayotte, which would bias the treatment effects downwards. In order to avoid respondents discomfort and biased responses, we collected information as indirectly as possible, by leveraging data on household composition collected at

[17] Questions on the program were asked in the last module of the survey in order to avoid influencing the behaviors of respondents and interviewers in other modules.

baseline. In particular, our main measure of migration relies on questions asking whether each baseline household member is still residing in the household at follow-up, and if not, where he or she is currently residing with Mayotte as one of the choices. Because it does not make salient that the purpose of the questions is to assess migration to Mayotte, we believe that this design limits the risks of respondents unease and biased responses.

Table 2.4: Treatment effects on migration to Mayotte

	Migration (excl. returns)			Migration (incl. returns)		
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	0.030**	0.028**	0.028**	0.037**	0.034**	0.033**
	(0.013)	(0.012)	(0.012)	(0.015)	(0.015)	(0.015)
Extended controls		✓	✓		✓	✓
Island FE			✓			✓
Control mean	0.078	0.078	0.078	0.128	0.128	0.128
Observations	2181	2181	2181	2181	2181	2181

Notes: This table reports LPM estimates of treatment effects on migration using equation (2.17). The dependent variable in columns 1 to 3 is a dummy equal to one if at least one household member migrated to Mayotte after the baseline survey and is still in Mayotte during the follow-up survey. In columns 4 to 6, the dependent variable also equals one if at least one household member migrated to Mayotte after the baseline survey but returned to his household of origin (voluntarily or not). All estimates control for unbalanced covariates. Extended controls include the following variables (measured at baseline): household willingness to migrate; migration experience; network in Mayotte; household head's gender, age, and schooling; household size, consumption, and livestock; dummy variables equal to one if the household has a bank account, income-generation activities (other than agriculture), fields, electricity, and a private water access. Standard errors in parentheses are clustered at the level of the treatment (household). *** p<0.01, ** p<0.05, * p<0.1.

Because the French police expels a large number of illegal Comorians each year, migration is often short-term.[18] Therefore, we also collected information on return migrants, by inquiring whether any household member took a *kwassa* for Mayotte in the 24 months prior to the follow-up survey. This measure is not without caveats and could bias the estimates, given that (i) it is more direct and thus exposed to the reporting bias mentioned above, (ii) the 24 months recall period may include pre-program migrations because of program's progressive roll-out and recall errors, and (iii) it does not inquire about household members who have died (some of which may have died en route to Mayotte), or household members who have left the household and are not currently in Mayotte, but could still have been in Mayotte in between.[19] These caveats are likely to

[18]Each year, about 20,000 migrants are deported to Comoros (Sénat, 2008). This corresponds to roughly 8 percent of Mayotte population or 2.5 percent of Comoros population.

[19]Comorian migrants are always deported to Anjouan (Mayotte's closest neighbor), even

attenuate our estimates of treatment effects.

We find that the program had a sizable and positive impact on migration to Mayotte. Column 1 shows that the treatment increased migration to Mayotte by three percentage points (significant at the 5% level), which represents a 38 percent increase relative to the control group. Including returnees does not alter the results. Estimates of treatment effects are larger in absolute terms but smaller in relative terms (consistent with the attenuation bias highlighted above). As can be seen in column 4, the program increased migration by 3.7 percentage points, equivalent to a 29 percent increase relative to the control group. Results are stable when extended controls and island fixed effects are included (columns 2, 3, 5 and 6). Table B1 shows LATE effects of the program. Not surprisingly, LATE estimates are consistently larger than ITT estimates suggesting that the program particularly increased migration among the sample of compliers.

2.5.3 Threats to our interpretation

These results are consistent with the idea that the cash-for-work program increased migration to Mayotte. However, this interpretation is exposed to various threats that could produce a similar pattern in the data. We explore three alternative explanations for the observed effects: (i) selective attrition; (ii) selective household dissolution; (iii) asymmetric indirect effects.

Selective attrition Because attrition can sometimes be explained by whole household migration, a typical concern with impact evaluation looking at migration is related to differential attrition rates between experimental groups. In our case, if households in the control group were more affected by whole household migration than households in the treatment group, our estimates would be biased upwards. A few observations help to mitigate this concern. First, the attrition rate is very low (about 4%) and similar across experimental groups (Table B2). Moreover, qualitative interviews indicate that whole household migration to Mayotte is uncommon. Because migration to Mayotte is both risky and costly, households typically send one migrant, two at most (e.g. when a parent migrates with his or her child). Finally, even if we considered an unlikely scenario in which all attritors migrated to Mayotte, we would still observe a positive impact on migration.

Selective household dissolution A similar concern is related to household dissolution and migration. As shown by Bertoli and Murard (2019), the mi-

though they are from Grande Comore or Moheli. Then, they either return to their island of origin, settle in a new location, or try to get back to Mayotte.

gration of an individual increases the probability that his or her household of origin dissolves subsequently. Because the program was targeted at the household level, beneficiary households may have had an incentive to preserve their living arrangements after the migration of a household member, thus being relatively less likely to dissolve. Again, this would lead to a relatively higher attrition rate in the control group and would bias our results upwards. In Table B3, we check whether beneficiary households are less likely to dissolve by analyzing attrition reasons given by enumerators. Reassuringly, household dissolution was similar in the control and treatment groups. About two percent of households in both experimental groups could not be followed-up because they dissolved.[20]

Asymmetric indirect effects A number of recent studies highlight the importance to estimate not just direct effects of anti-poverty programs but also their indirect effects (Angelucci and De Giorgi, 2009; Beegle et al., 2017). Indirect negative (resp. positive) effects would bias our results upwards if they were more (resp. less) prevalent for control households. For instance, control households could be hurt by price spikes or increased competition for scarce investment opportunities. We estimate indirect effects using equation (2.18). Table B4 reports the sign and magnitude of indirect effects for both experimental groups. We see no evidence of significant indirect treatment effects. If anything, these affects are small and similar across treatment and control groups.

2.6 Channels

We now turn to the investigation of the various channels that could explain why the cash-for-work program increased migration to Mayotte.[21] We explore the four channels highlighted in our conceptual framework (Section 2.2): (i) the liquidity channel; (ii) the opportunity cost channel; (iii) the credit constraint channel; and (iv) the risk-aversion channel.[22] The evidence suggests that the

[20]Two ingredients of the project implementation may explain this pattern. First, payments were made to individuals performing the work rather than to household heads. Second, formal and informal arrangements to replace workers were possible both within and across households. Drop-out workers were supposed to be replaced by another adult household member, but in practice, FADC did not keep track of the exact initial household composition, meaning that the replacements could incorporate endogenous household changes. The qualitative interviews with beneficiaries reveal that replacements by extended family members or relatives were quite common. Taken together, these observations support the idea that incentives for beneficiaries to preserve the household structure were likely weak in practice.

[21]This section should be regarded as exploratory since it was not included in our pre-analysis plan.

[22]In this section, we focus on our first definition of migration (i.e. excluding return migrants) due to space limitation. Results including return migrants are similar (available upon request).

increase in migration is explained by the alleviation of liquidity and risk constraints on one hand, and by the fact that the program did not increase the opportunity cost of the persons who were the most likely to migrate on the other hand.

2.6.1 Liquidity channel

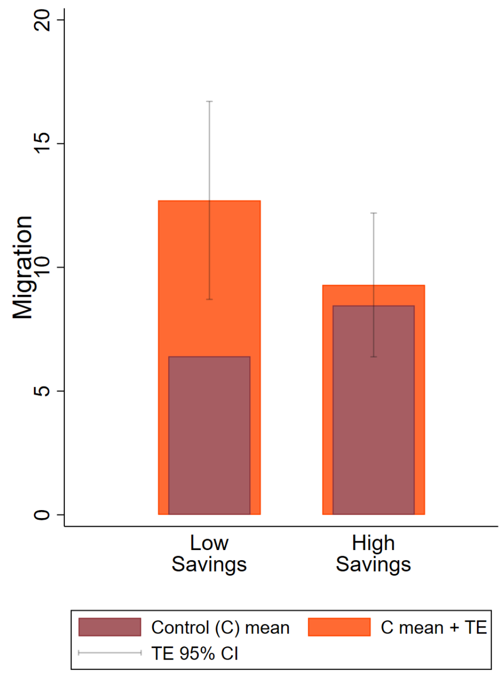
According to the liquidity channel, cash transfers may have relaxed budget constraints, thus facilitating the migration of households facing liquidity constraints. In order to check whether this channel is relevant in our setting, we estimate program effects conditional on baseline savings using equation (2.19).[23] In line with the liquidity channel, Figure 2.8 shows that positive effects on migration were mostly visible for the group of households with low baseline savings. Importantly, the migration rate in the control group was relatively higher for households with high baseline savings, suggesting that financial constraints are binding in our setting. Overall, it seems that cash transfers allowed some households with low baseline savings to overcome otherwise binding financial constraints. Table B5 shows that these results are qualitatively unchanged using more parsimonious specifications of equation (2.19) and a continuous variable for savings.

Our second approach to investigate the liquidity channel is to look at program effects on migration to other destinations. If the increase in migration to Mayotte is due to relaxed financial constraints, we should not detect similar effects on migration to cheaper, previously unconstrained destinations. As can be seen from Table 2.5, the program had small and non-significant effects on domestic migration. We do not observe effects on migration to mainland France either, most likely because the binding constraint for this destination is administrative rather than financial.[24]

[23] It is often challenging to measure savings, especially in low-income settings where it can take various forms. In Comoros, households typically save using livestock and tontines. In addition, many households take on debts from various operators (friends, shop owners, etc.) such that their savings can actually be negative. In order to capture household net savings, we derive a variable combining the value stored in these various vehicles. Specifically, the money saved in livestock and tontines enter positively in the variable, whereas the amount of debts enter negatively.

[24] The absence of impact on migration to other destinations could also be due to the fact that Comorians typically migrate legally to these destinations, implying that there is less uncertainty in the migration outcome and that the risk-aversion channel could be inactive (see Section 2.6.4 below).

Figure 2.8: Liquidity channel



Notes: This figure shows follow-up households' migration rates conditional on baseline savings. Households are divided in two groups depending on their levels of net savings at baseline. Low (resp. high) savings correspond to savings below (resp. above) mean savings. An IHS transformation was applied to savings in order to limit the influence of outliers. Treatment effects and 95% confidence intervals are derived from the estimate of equation (2.19) including all controls (balanced covariates, island fixed effects, and their interactions with savings). N = 2181.

2.6.2 Opportunity cost channel

As shown is Section 2.2.2, cash transfers that are conditional on remaining in the origin country increase the opportunity cost of migrating and could therefore reduce migration. We argue that this channel has not been operating in our setting because the cash-for-work program was very flexible. Beneficiary households were entitled to send one adult of their choice to public works and, most importantly, the cash transfers were not conditional upon other household members staying in Comoros. Beneficiary households could therefore select one household member to participate in public works, and, in the meantime, use the cash transfers to finance the migration of another household member. This conjecture is reinforced by the fact that there is qualitative evidence

Table 2.5: Other migration patterns

	(1) Domestic Mig. (intra-island)	(2) Domestic Mig. (inter-island)	(3) Migration France	(4) Migration Other
Treatment	-0.023	0.007	-0.001	0.002
	(0.018)	(0.010)	(0.007)	(0.007)
Extended controls	✓	✓	✓	✓
Island FE	✓	✓	✓	✓
Control mean	0.236	0.057	0.029	0.030
Observations	2181	2181	2181	2181

Notes: All estimates control for unbalanced covariates. Standard errors in parentheses are clustered at the level of the treatment (household). *** p<0.01, ** p<0.05, * p<0.1. See notes to Table 2.4 for more details.

suggesting that people are financing the migration of others: "*I gave 40,000 KMF [US$92] to my son for his trip to Mayotte. Life is hard. We had no one to ask for help. My son decided alone to leave in the hope of helping us. I didn't have much. But to encourage him, I gave this small amount*".

Although in theory the cash-for-work program could have still increased the opportunity cost of migrating of the persons participating in CFW activities, in practice these persons were pretty far from the profile presented by typical migrants. As can be seen in Table 2.6, workers were on average older and less educated than migrants, and most workers were females with no migration experience while a majority of migrants were males. This suggests that the program primarily increased the opportunity cost of persons who were unlikely to migrate (i.e. relatively old and lowly educated females with no previous migration experience).[25]

[25] Table 2.6 could actually suggest that participating in CFW activities did not deter migration at all, since treated and control migrants were very similar. If CFW participation had reduced the migration rate of workers, we may expect the differences between control migrants and workers to be smaller than the differences between treated migrants and workers. We find little evidence to suggest this pattern except for the dummy indicating whether individuals had an income-generating activity (other than agriculture) at baseline.

Table 2.6: Summary statistics on project workers and migrants

	Treated			Controls
	Non-migrants		Migrants	Migrants
	Worker=1	Worker=0		
Age	39.56	30.30	28.73	29.18
Male	0.22	0.60	0.59	0.56
Education				
Did not complete primary	0.56	0.26	0.28	0.25
Primary	0.25	0.19	0.28	0.35
Secondary	0.17	0.44	0.38	0.33
Tertiary	0.03	0.11	0.06	0.07
IGA	0.24	0.20	0.14	0.22
Migration experience	0.07	0.06	0.29	0.30
Observations	991	3166	196	105

Notes: The sample is restricted to adults (15-65 at baseline).

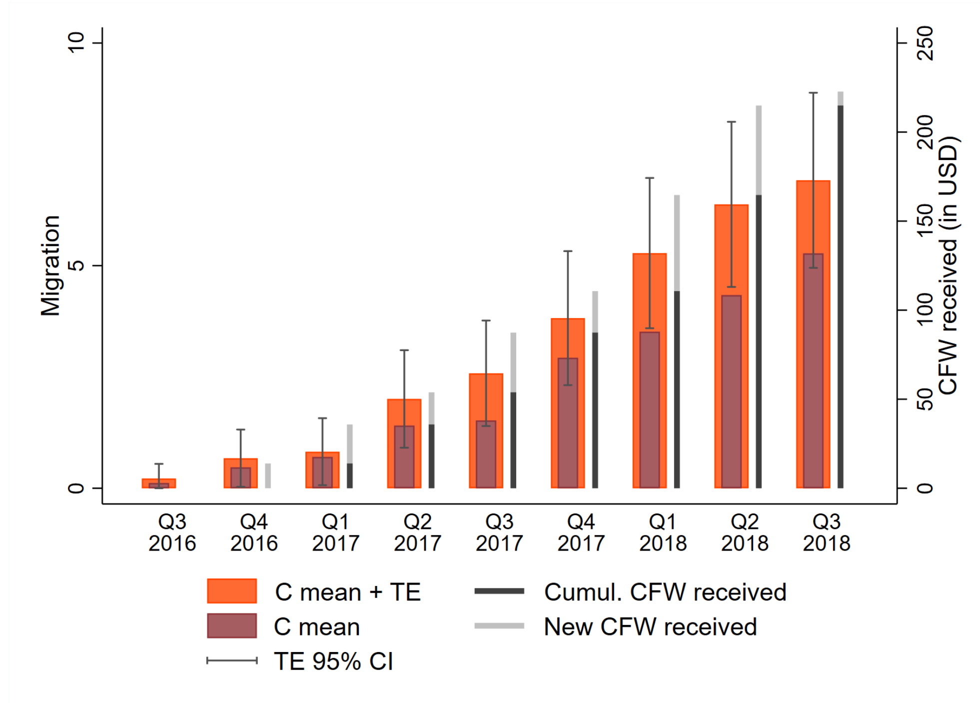
2.6.3 Credit constraint channel

According to the credit constraint channel, cash transfers could have facilitated access to credit and thereby increase migration of credit constrained households as soon as they got selected to benefit from the program. Our evidence suggests that the credit constraint channel was negligible in this study. First, Table 3.3 indicates that control and treated households had similar baseline levels of debts, meaning that beneficiary households did not alter their financial behaviors although they already knew they would benefit from streams of cash transfers. Second, when respondents reported a migrant, we further inquired about the month and year of migration. This retrospective data allows us to explore the evolution of the treatment effect over time. Figure 2.9 shows, for each quarter between July 2016 and September 2018, the treatment effect along with the migration rate in the control group.[26] Figure 2.9 also shows the timing of cash transfers (measured using administrative data). Overall, we see that the treatment effect increased over time, and that the correlation with cash transfers disbursement is rather strong: increases in treatment effect follow closely the disbursement of cash transfers.

In order to have a better understanding of these dynamics, Table 2.7 investigates in a more systematic way the timing of cash transfers and migration. We assemble a panel with detailed information on migration history and cash transfers received. We are particularly interested to check (i) whether migration

[26] The treatment effect and the control mean in the last period (Q3 2018) are lower than in Table 2.4 because 25 percent of the respondents only recalled the year of migration and are thus excluded from the pool of migrant households for this analysis. As a robustness check, we replaced missing month by a randomly generated month. Results, available upon request, show that the dynamic of the treatment effect is the same though the estimates are more precise.

Figure 2.9: Treatment effect over time



Notes: This figure shows the evolution of follow-up households' migration rates over time. Treatment effects and 95% confidence intervals are derived from the estimate of equation (2.17) including all controls (unbalanced convariates, balanced covariates and island fixed effects). N = 2181.

decisions at time t are explained by the amount of cash received at time t, cash received at time t-1, or total cash received pre-t, and (ii) whether the impact of the cash received at time t is conditional on the total amount of cash received beforehand. In column 1, we see that most of the impact seems to come from cash received at time t, meaning that individuals reacted rather quickly to cash transfers. In contrast, cash transfers received at time t-1 did not seem to make much difference (column 2). However, it is interesting to see in column 3 that the impact of cash received at time t is actually conditional on the total amount received beforehand. Overall, it seems that migration occurred in time periods where households received cash conditional on having accumulated enough liquidity in the previous periods. This pattern suggests that liquidity constraints rather than credit constraints may have been alleviated by the program. If anything, this evidence reinforces the relevance of the liquidity channel (Section 2.6.1).

Table 2.7: Timing of cash transfers and migration

	Migration t		
	(1)	(2)	(3)
Cash t	0.0044***	0.0049***	0.0002
	(0.002)	(0.002)	(0.002)
Cash Tot. t-1	0.0007		-0.0003
	(0.000)		(0.000)
Cash t-1		-0.0033	
		(0.002)	
Cash Tot. t-2		0.0014*	
		(0.001)	
Cash t **x** Cash Tot. t-1			0.0027***
			(0.001)
Migration t-1	0.982***	0.981***	0.982***
	(0.002)	(0.003)	(0.002)
Extended controls	✓	✓	✓
Island FE	✓	✓	✓
Control mean	0.023	0.023	0.023
Observations	17304	15141	17304

Notes: All estimates control for unbalanced covariates. Standard errors in parentheses are clustered at the level of the treatment (household). *** p<0.01, ** p<0.05, * p<0.1.

2.6.4 Risk-aversion channel

As shown in our simple theoretical model, if migration is risky and if households have DARA preferences, the cash transfers can reduce risk-aversion and thereby increase the expected utility returns from migration. This risk-aversion channel could be particularly relevant to explain our results given that Comorians migrating to Mayotte face considerable risks of death or expulsion. As emphasized in Section 2.3.1, thousands of Comorian migrants have died trying to reach Mayotte, and even more have been arrested and deported to Comoros. Qualitative interviews suggest that these risks have a strong influence on migration decisions, as illustrated in the following quote: *"There are two things that automatically get inside the minds of the person who wants to migrate and his family: the risk of dying in the sea which is very common; the risk of being arrested by the police which can be really painful considering the expenses incurred"*. In addition, we believe that it is reasonable to assume DARA preferences in a setting suffering from widespread poverty and a lack of formal social safety nets. In this context, many households could face what Bryan et al. (2014) called a *"subsistence constraint"*, that is a situation where poverty is so strong that failed investments would lead to unbearable welfare losses.

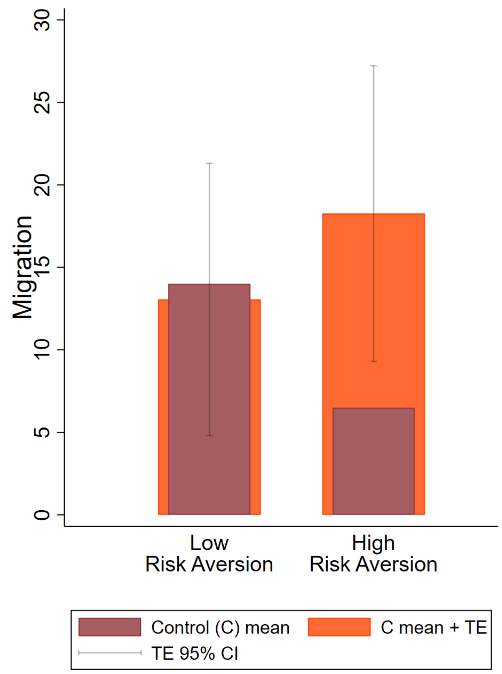
To investigate this channel, we estimate program effects conditional on baseline risk-aversion. Our measure of risk-aversion is derived from a simple discrete choice experiment in which respondents willing to migrate to Mayotte were asked to make a choice about the number of persons in the kwassa. Our qualitative evidence indicates that aspiring migrants typically face this choice in the real world and trade-off migration costs and migration risks. The more persons in a kwassa, the lower the price of the journey but the higher the risks of accident or arrest. Respondents were presented with three choices: (i) an overloaded kwassa (the less expensive but most risky technology); (ii) a properly loaded kwassa; (iii) what is often called a VIP kwassa, i.e. a kwassa with very few people (the most expensive but less risky technology). The exact question was as follow:[27]

> *Imagine that you should take a small kwassa to migrate to Mayotte. The maximum capacity of the kwassa is 10 persons. You have the choice between three prices:*
>
> 1. *You pay 100,000 KMF [US$230] and there are more than 10 persons on the kwassa*
> 2. *You pay 250,000 KMF [US$575] and there are between 5 to 10 persons on the kwassa*
> 3. *You pay 500,000 KMF [US$1150] and there are less than 5 persons on the kwassa*
>
> *Which option would you choose?*

Overall, 50.1% of the respondents selected choice (i), 20.6% choice (ii), and 28.3% choice (iii). We estimate a simple regression of the choice on baseline consumption and use the residuals as a proxy of risk aversion. In other words, risk-aversion is derived from the part of the choice that is not explained by household consumption. In line with the risk-aversion channel, Figure 2.10 shows that positive effects are only visible for the group of households with high levels of risk-aversion at baseline. In the control group, migration is lower among the highly risk-averse, which is consistent with DARA preferences and the theoretical result that risk adversity is a barrier to migration. Table B6 shows that these results are qualitatively unchanged using more parsimonious specifications of equation (2.19) and a continuous variable for risk-aversion.

[27] These choices have been calibrated during enumerators' training and the pilot survey to reflect real world choices in as much as possible.

Figure 2.10: Risk-aversion channel



Notes: This figure shows follow-up households' migration rates conditional on baseline risk-aversion. Households are divided in two groups depending on their levels of risk-aversion at baseline. Low (resp. high) risk-aversion corresponds to risk-aversion below (resp. above) mean risk-aversion. Treatment effects and 95% confidence intervals are derived from the estimate of equation (2.19) including all controls (balanced covariates, island fixed effects, and their interactions with savings). N = 476.

2.7 Conclusion

Although international migration can lead to large income gains, existing migration flows remain relatively limited. In this paper, we show that cash transfers targeted to very poor households in Comoros increased migration to the neighboring and richer French island of Mayotte. This increase in migration can be explained by the alleviation of liquidity and risk constraints on one hand, and by the fact that the program did not increase the opportunity cost of the persons who were the most likely to migrate on the other hand. The effect of the cash transfers on migration to Mayotte is significantly larger for households with low levels of savings at baseline, or high levels of risk-aversion at baseline. This suggests that cash transfers can ease liquidity and risk constraints and thereby increase (costly and risky) migrations. Although in theory the labor requirement of the program could have reversed these effects by increasing the opportunity cost of migrating, in practice the persons that participated in the program were pretty far from the profile presented by typical migrants.

These findings are in line with the idea that many households do not migrate because of binding liquidity constraints (Angelucci, 2015; Bazzi, 2017; Mahajan and Yang, 2017). It also adds to a nascent literature showing that risk is an important deterrant of migration decisions. For example, Batista and McKenzie (2018) find that adding a risk of unemployment to migration outcomes plays a crucial role to explain the low levels of migration typically observed. In the same vein, Bah and Batista (2018) conducted a Lab-in-the-Field experiment in Gambia showing that the willingness to migrate illegally to Europe is heavily shaped by the risk of dying en route and by the probability of obtaining a legal residency status. While our findings confirm that risk is an important barrier to migration, it also suggests that a social protection program such as the SSNP can ease risk bearing and thereby risky migrations. In other words, the program provided a safety net to beneficiaries allowing them to invest in risky migration opportunities.

Bibliography

Abadie, A., Chingos, M. M., and West, M. R. (2018). Endogenous stratification in randomized experiments. *Review of Economics and Statistics*, 100(4):567–580.

Adhikari, S. and Gentilini, U. (2018). Should I stay or should I go: do cash transfers affect migration? Working Paper. The World Bank.

Angelucci, M. (2015). Migration and financial constraints: Evidence from Mexico. *Review of Economics and Statistics*, 97(1):224–228.

Angelucci, M. and De Giorgi, G. (2009). Indirect effects of an aid program: How do cash transfers affect ineligibles' consumption? *American Economic Review*, 99(1):486–508.

Bah, T. L. and Batista, C. (2018). Understanding willingness to migrate illegally: Evidence from a lab in the field experiment. Mimeo, Universidade Nova de Lisboa, NOVAFRICA.

Batista, C. and McKenzie, D. (2018). Testing classic theories of migration in the lab. Mimeo, Universidade Nova de Lisboa, NOVAFRICA and World Bank.

Bazzi, S. (2017). Wealth heterogeneity and the income elasticity of migration. *American Economic Journal: Applied Economics*, 9(2):219–55.

Beegle, K., Galasso, E., and Goldberg, J. (2017). Direct and indirect effects of Malawi's public works program on food security. *Journal of Development Economics*, 128:1–23.

Bertoli, S. and Murard, E. (2019). Migration and co-residence choices: Evidence from Mexico. *Journal of Development Economics*.

Blanchy, S. (2002). Mayotte: française à tout prix. *Ethnologie Française*, 32(4):677–687.

Bryan, G., Chowdhury, S., and Mobarak, A. M. (2014). Underinvestment in a profitable technology: The case of seasonal migration in Bangladesh. *Econometrica*, 82(5):1671–1748.

Chiswick, B. R. (1988). Illegal immigration and immigration control. *Journal of Economic Perspectives*, 2(3):101–115.

Clemens, M. A. (2011). Economics and emigration: Trillion-dollar bills on the sidewalk? *Journal of Economic perspectives*, 25(3):83–106.

Clemens, M. A. et al. (2014). Does development reduce migration? *International Handbook on Migration and Economic development*, pages 152–185.

Dao, T. H., Docquier, F., Parsons, C., and Peri, G. (2018). Migration and development: Dissecting the anatomy of the mobility transition. *Journal of Development Economics*, 132:88–101.

De Brauw, A. and Carletto, C. (2012). Improving the measurement and policy relevance of migration information in multi-topic household surveys. *Living Standards Measurement Study Working Paper*.

Dohmen, T., Falk, A., Huffman, D., and Sunde, U. (2010). Are risk aversion and impatience related to cognitive ability? *American Economic Review*, 100(3):1238–60.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., and Wagner, G. G. (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association*, 9(3):522–550.

Dustmann, C. and Okatenko, A. (2014). Out-migration, wealth constraints, and the quality of local amenities. *Journal of Development Economics*, 110:52–63.

Eeckhoudt, L., Gollier, C., and Schlesinger, H. (2005). *Economic and financial decisions under risk*. Princeton University Press.

Gibson, J. and McKenzie, D. (2012). The economic consequences of 'brain drain' of the best and brightest: Microeconomic evidence from five countries. *The Economic Journal*, 122(560):339–375.

Guiso, L. and Paiella, M. (2008). Risk aversion, wealth, and background risk. *Journal of the European Economic association*, 6(6):1109–1150.

Hanson, G. H. (2006). Illegal migration from Mexico to the United States. *Journal of Economic Literature*, 44(4):869–924.

Imbert, C. and Papp, J. (2019). Short-term migration, rural public works and urban labor markets: Evidence from India. *Journal of the European Economic Association*. Forthcoming.

Le Monde (2017). "M. Macron, les kwassa-kwassa ont fait plus de 10 000 morts". *Opinion column - Hachim Saïd Hassane*.

Levy, H. (1994). Absolute and relative risk aversion: An experimental study. *Journal of Risk and Uncertainty*, 8(3):289–307.

Mahajan, P. and Yang, D. (2017). Taken by storm: Hurricanes, migrant networks, and us immigration. National Bureau of Economic Research Working Paper.

Marie, C.-V., Breton, D., Crouzet, M., Fabre, E., Merceron, S., et al. (2017). Migrations, natalité et solidarités familiales. La société de Mayotte en pleine mutation. *Insee La Réunion-Mayotte*.

McKenzie, D. and Rapoport, H. (2007). Network effects and the dynamics of migration and inequality: Theory and evidence from Mexico. *Journal of Development Economics*, 84(1):1–24.

McKenzie, D., Stillman, S., and Gibson, J. (2010). How important is selection? Experimental vs. non-experimental measures of the income gains from migration. *Journal of the European Economic Association*, 8(4):913–945.

Myerson, R. B. (2005). *Probability models for economic decisions*. Duxbury Press, Pacific Grove, CA.

Sénat (2001). Projet de loi relatif à Mayotte. Rapport Sénat 361.

Sénat (2008). L'immigration clandestine à Mayotte. Rapport Sénat 461.

Sénat (2012). Rapport d'information. Rapport Sénat 675.

Sjaastad, L. A. (1962). The costs and returns of human migration. *Journal of Political Economy*, 70(5, Part 2):80–93.

Stecklov, G., Winters, P., Stampini, M., and Davis, B. (2005). Do conditional cash transfers influence migration? A study using experimental data from the Mexican PROGRESA program. *Demography*, 42(4):769–790.

Wik, M., Aragie Kebede, T., Bergland, O., and Holden, S. T. (2004). On the measurement of risk aversion from experimental data. *Applied Economics*, 36(21):2443–2451.

Yang, D. (2008). International migration, remittances and household investment: Evidence from Philippine migrants' exchange rate shocks. *The Economic Journal*, 118(528):591–630.

Yesuf, M. and Bluffstone, R. A. (2009). Poverty, risk aversion, and path dependence in low-income countries: Experimental evidence from Ethiopia. *American Journal of Agricultural Economics*, 91(4):1022–1037.

Zelinsky, W. (1971). The hypothesis of the mobility transition. *Geographical Review*, pages 219–249.

Chapter 3

The (lack of) Value of Public Works: Evidence from Ethiopia

This chapter is a joint work with Victor Stéphane (GATE, Université Jean Monnet, Saint-Etienne).

3.1 Introduction

Public Works Programs (PWP) are popular poverty alleviation tools in developing countries. These programs provide short-term employment opportunities to poor, underemployed individuals in labor-intensive infrastructure projects (Alderman and Yemtsov, 2013; Subbarao et al., 2012). They follow the twin goals of reducing the poverty of participants and generating infrastructures to enhance development at a broader level. This premise of killing two birds with one stone has made PWP extremely appealing for policy makers. They have been implemented for decades in numerous low and middle-income countries such as Argentina, Ethiopia, India and South Africa, among others. Recently, they triggered attention as tools for countries that suffer from environmental degradation or need to adapt to climate change (Subbarao et al., 2012).

Despite the popularity of PWP, there is surprisingly little evidence on the productive value of the infrastructures they generate. This lack of evidence is problematic because it prevents comprehensive cost-effectiveness exercises and comparisons with other poverty alleviation programs. PWP are usually more expensive to run due to higher administrative costs. For instance, Gehrke and Hartwig (2018) estimate that for each dollar spent in cash transfer programs an average of 42 cents reaches beneficiaries while this amount falls to 31 cents with PWP. In addition, because of the labor requirement of PWP, forgone income can be considerably higher in PWP than in other poverty alleviation programs (Murgai et al., 2015). While these drawbacks are generally justified by the assumption that PWP infrastructures generate important productive effects,[1] empirical grounds remain scant.

In this paper, we attempt to provide crucial evidence on the infrastructures generated by public works in the context of the Productive Safety Net Program (PSNP). The PSNP is a flagship PWP implemented in Ethiopia since 2005. It provides cash or food transfers to 8 million beneficiaries in chronically food insecure woredas (districts) in exchange for their participation in labor-intensive activities. Because most of the activities are focused on land management and environmental projects such as soil and water conservation activities (afforestation, construction of terraces and flood control structures, renovation of traditional water bodies, etc.), and soil fertility measures (agroforestry, gully control, compost generation, etc.), the PSNP is sometimes considered as Africa's largest climate change adaptation program (Subbarao et al., 2012). The reasoning behind these activities is to address the underlying causes of food insecurity by increasing agricultural productivity and resilience to climate shocks.

A substantial amount of literature has investigated PSNP direct effects on beneficiaries welfare. The evidence suggests that cash and food transfers had positive effects on food security (Berhane et al., 2014; Gilligan et al., 2009), children nutritional status and human capital accumulation (Debela et al., 2015; Porter and Goyal, 2016; Favara et al., 2019; Mendola and Negasi, 2019), livestock holding (Berhane et al., 2014), and tree planting (Andersson et al., 2011). In addition, the literature indicates that the PSNP did not divert children from schooling or increased child labor (Hoddinott et al., 2010) – two typical concerns with PWP. While appealing, these positive effects are regularly observed – generally at lower costs – in alternative poverty alleviation programs such as unconditional cash transfers (Baird et al., 2014; Haushofer and Shapiro, 2016). A question with high policy relevance is therefore whether PWP such as the PSNP provides any additional, specific welfare gain that could justify higher operational costs. While it has been argued that PSNP works increased land productivity by three to four times (European Commission, 2015), or that they improved land and water management technologies of an estimated 901,654 hectares (World Bank, 2016), it is hard to know how these figures have been derived and whether they are reliable. The objective of this paper is to provide rigorous evidence on the productive effects of PSNP works.

In theory, the works could have had two main effects. First, at the intensive margin, they may have increased agricultural productivity by promoting sus-

[1]For instance, Subbarao et al. (2012) argue that "*there is no reason to do public works if the public goods generated do not have a positive impact on the community*".

tainable land management measures on existing plots. Second, at the extensive margin, they may have lead to the rehabilitation of degraded or infertile lands that were not previously cultivated. These effects could vary depending on ecological contexts. In particular, Ethiopian highlands face acute environmental degradation and are therefore particularly likely to benefit from the public works. For this reason, we will analyze PSNP effects separately for highlands and lowlands.

To assess rigorously the productive effects of PSNP works, we face two main challenges: (i) a dearth of data to gauge the evolution of agricultural productivity over a sufficient period of time; (ii) a lack of obvious counterfactual because the PSNP was targeted and not randomly allocated. To overcome the first challenge, we use satellite, geo-referenced data. We build a proxy for crop productivity by combining the Normalized Difference Vegetation Index (NDVI) with highly disaggregated information on land use, crop types, and crop calendars. We show that this variable is a good predictor of agricultural output. Then, we assemble data on climatic conditions (rainfall and temperature), topographic characteristics, night-time lights, and population density to control for a maximum of potential confounding factors. This results into an original dataset covering whole Ethiopia over the 2000-2013 period. We mitigate the second concern by using difference-in-differences estimates and the inverse probability weighting method (Hirano et al., 2003). We show that this empirical strategy allows us to get rid of pre-treatment trend differences between the treated and control groups.

This paper contributes to the literature by providing rigorous estimates of the productive value of PWP in the Ethiopian context.[2] In contrast with existing narratives, we find no evidence to support that public works had measurable impacts on agricultural productivity and resilience to climate shocks. We conduct several robustness checks to assess the validity of our results. In all cases, the effects of the PSNP remain quantitatively small and non-significant. These results point out that PWP infrastructures will not always generate significant and measurable productive effects.

To the best of our knowledge, there are only two other studies evaluating quantitatively the effects of PWP infrastructures. Using a randomized control trial of the Labor Intensive Works Program (LIWP) in Yemen, Christian et al. (2015) show that water-related projects (e.g. water storage tanks and cisterns, rainwater harvesting tanks, and improvement of shallow wells) had large and positive effects on water accessibility especially in villages with poor baseline access. However, a clear concern here is that the results were derived from the

[2] A minor, supplementary contribution of our paper is to provide another application of Geospatial Impact Evaluation (GIE) (BenYishay et al., 2017). Our study further substantiate the potential of GIEs to answer important questions at low costs and in data-scarce settings.

subset of villages with completed projects at the time of the follow-up survey (only 8 out of the 82 projects were completed at follow-up) and could therefore reflect convergence in water access rather than the effects of the LIWP. Using a panel GMM (generalized method of moments) estimator to evaluate the productive effects of the PSNP, Filipski et al. (2016) find that soil and water conservation measures increased the average yields of grain crops by about 2.8 percentage points but had no effects on non-grain crops. However, these results are subject to the typical concerns about the use of GMM estimators to achieve causal inference (Roodman, 2009). In addition, as mentioned by the authors themselves, some results could reflect a lack of statistical power.

Naturally, our own study is not without caveats. First, it provides only reduced form estimates of the impact of the PSNP on agricultural productivity and resilience, and while these estimates are well-tailored to the current policy debate (PWP's presumed double dividend), we note that they could partly reflect negative effects of PSNP transfers on agricultural productivity (we provide suggestive evidence that this is not the case). Second, we observe a lack of impact of PWP infrastructures but there is little we can say on the reasons that may explain it.[3] Third, while PSNP works could have had effects at both the intensive and extensive margins (see above), this study is only informative of the lack of effects at the intensive margin. We tried to design a satellite-based outcome to capture effects at the extensive margin but were admittedly unsuccessful. Fourth, context obviously matters, and our study does not mean that all PWP infrastructures have limited effects. Further research is required to see whether this result is specific to this setting or has some external validity.

Nonetheless, we believe that this study provides a useful piece to the debate surrounding the design and implementation of efficient social safety net programs in developing countries. The Ethiopian PSNP is Africa's flagship PWP and has probably been playing a crucial role in the decision of 38 other African countries to implement government-supported PWP (World Bank, 2015). If anything, our results suggest that PWP infrastructures do not always generate measurable effects, and thus call for a more attentive examination of the double dividend that development practitioners typically attribute to public works.

[3]One potential reason for the lack of impact could be related to the low quality or lack of durability of the infrastructures. For example, there is some evidence from other contexts that the infrastructures generated in PWP are often undermined by climate shocks such as droughts or floods (Kaur et al., 2019; Gazeaud et al., 2019b). Further research along these lines would be particularly welcome.

3.2 Background

With more than 100 million inhabitants, Ethiopia has currently the second largest population in Africa after Nigeria. The country is administratively divided into ethnically based regions, which are themselves subdivided into zones, and woredas (districts). It is composed of a vast territory made of mountains and plateaus lying at elevations above 1500m, divided by the Rift Valley, and surrounded by lowlands. Livelihoods predominately depend on crop production in highlands and on agro-pastoralism in lowlands. Over the last decades, Ethiopia has faced severe droughts which often resulted in large-scale food crisis.[4] Despite some progress, it is still one of the poorest country in the world with a per capita income of US$783 in 2017 (World Bank data). Poverty is especially widespread in rural areas where most people are engaged in subsistence agriculture and face important environmental degradation. Environmental degradation not only reduces land productivity, it is also reducing the capacity to effectively manage water resources. According to a World Bank report, "*[Ethiopian] land base has been damaged through erosion and degradation, land productivity has declined, and rainfall infiltration has reduced such that many spring and stream sources have disappeared or are no longer perennial*" (World Bank, 2006, p.2). Environmental degradation is particularly prevailing in Ethiopian highlands because of the steep slopes and widespread deforestation.[5]

The Productive Safety Net Program (PSNP) was launched in 2005 by the Government of Ethiopia in an attempt to provide a long-term solution to the chronic food insecurity found in rural parts of the country. The PSNP replaced an old system where food aid depended on emergency humanitarian appeals for international assistance. This system proved inefficient as assistance was unpredictable both for planners and local populations (Jayne et al., 2002; Kehler, 2004). In contrast, the PSNP aimed to provide a predictable and reliable safety net to address chronic food insecurity and mitigate recurrent climate shocks. The PSNP was quickly scaled up to reach approximately 8 million beneficiaries in 2006, thereby becoming the largest workfare program in Africa. Today, it operates with an annual budget of more than US$500 million.

The main component of the PSNP consists of cash or food transfers to selected poor households conditional on their participation in labor-intensive public works projects. Using a mix of geographic and community-based tar-

[4] The most prominent example is probably the 1983-1984 famine from which up to one million people are estimated to have died (Devereux, 2000).

[5] For a rather old but enlightening examination of environmental degradation in Ethiopian highlands, see a study commissioned by the FAO arguing that "*the highlands of Ethiopia contain what is probably one of the largest areas of ecological degradation in Africa, if not in the world*" (Hurni, 1983, p.ii).

geting devices, the program targets chronically food insecure households in chronically food insecure woredas. The government identified chronically food insecure woredas based on the number of years they had required food assistance prior to 2005. Then, in each eligible woredas, local community councils known as food security task forces (FSTF) identified food insecure households, namely households that (i) have repeatedly faced food gaps or received food aid in the past three years, (ii) have suffered from a severe loss of assets due to a severe shock, and (iii) have no other source of support (family or social protection programs). These targeting instructions were intended as a broad national framework, but in practice the program implementation manual allowed for regional and local adaptation by FSTF (Sharp et al., 2006). To avoid interference with farming and other income-generating activities, able-bodied adults of beneficiary households could participate in public works only during the agricultural off-season.[6] The wage rate was initially set at 6 birr per day – approximately US$0.70 using the 2005 official exchange rate – but it gradually increased to reflect inflation patterns. According to administrative data, PSNP activities generated 227 millions person-days of employment in 2008 (World Bank, 2016).

In the PSNP, most public works focus on watershed development, with the objective to achieve environmental rehabilitation and increase agricultural productivity. Projects were selected locally through a community-based participatory approach and integrated into woredas development plans. Importantly, the peculiar conditions found in pastoral regions (Afar and Somali) caused implementation delays and required some tweaking in terms of program design. In particular, beneficiary households received unconditional transfers until 2009-2010, implying that these regions only started to benefit from public works activities in 2010. Our empirical strategy will need to incorporate this specificity.

3.3 Data

To estimate the impact of the PSNP infrastructures on crop production, we assemble an original database covering whole Ethiopia over the 2000-2013 period. We rely on high resolution satellite data to conduct our study at the woreda-year level. This section describes in details how we built this dataset.

[6]Beneficiary households with no able-bodied adult members were included in the direct support component of the PSNP (i.e. unconditional cash or food transfers of the same amount). Direct support beneficiaries represent about 16 per cent of total beneficiaries.

3.3.1 Crop Production

We use the the Normalized Difference Vegetation Index (NDVI) provided by MODIS as a proxy of crop production. This variable is available bi-monthly at a resolution of 250m and has been often used to measure vegetation cover (Ali et al., 2018) and crop productivity (Pettorelli et al., 2005; Wang et al., 2005). Since our study is conducted at the woreda-year level, the NDVI needs to be aggregated at this scale. Doing so could lead to an "aggregate-out" problem. That is, if the treatment has a spatially localized impact, averaging the NDVI over the whole woreda could dilute its effect and lead us to misconclude to an absence of effect of the PSNP. A similar concern arises regarding the time dimension. Indeed, because we do not expect PSNP effects outside of the growing season, we could worry, as above, that averaging the NDVI over the full year would dilute the effect of the program. Again, this would lead us to misconclude that public works implemented through the PSNP have no effect on crop productivity.

To tackle these issues, we impose spatial and time constraints when aggregating the NDVI. Regarding the spatial dimension, we average the NDVI using pixels covering cultivated areas only. To do so, we rely on the Land Use database provided by MODIS. This database, available on an annual basis, provides information on soil occupation (forest, savannas, grasslands, croplands, etc.) at a resolution of 500m.[7] One may worry that land occupation could itself be affected by the program, as the PSNP may lead farmers to cultivate new plots (extensive margin effect). In that case, using yearly data on soil occupation to compute the NDVI could also lead to a bias in our estimates.[8] For this reason, we decide to focus only on plots that were cultivated at the the beginning of the program. Then, we compute the NDVI using pixels covering cultivated areas in 2005. Regarding the time dimension, we aggregate the NDVI over months covering the growing season of the main crop in each woreda. To do so, we use the MIRCA 2000 database which provides information on the type of crops, the size cultivated, and the period of the growing season at a grid resolution of 5 arc-minute (10km).

In sum, for each woreda and each year of the 2000-2013 period, we compute the average NDVI using pixels covering cultivated areas in 2005 and months corresponding to the growing season of the main crop cultivated.

We use the 2013 and 2015 LSMS-ISA survey rounds to investigate whether our proxy is indeed a good predictor of crop production and crop productivity.[9]

[7] A pixel is considered to be cultivated if at least 60 percent of its surface is cultivated.

[8] For instance, if pixels newly identified as cultivated areas have a lower NDVI than older cultivated areas, our estimates would be biased downward.

[9] An additional LSMS-ISA survey round was implemented in 2011. However, data on crop

We derive both the total production and the average productivity of land in 2013 and 2015, and test whether these measures are well correlated with our proxy. The results, presented in Table 3.1, support the idea that our proxy is a good predictor of agricultural output. Our NDVI proxy is positively and significantly correlated with both measures of agricultural output. Importantly, these relationships hold when woreda fixed effects are included, meaning that our proxy not only predicts levels of agricultural outputs (columns 1 and 4), but also their variations over time (columns 2 and 5). Overall, the NDVI proxy seems perfectly suited to capture PSNP works effects at the intensive margin, i.e., productivity gains on parcels already cultivated when the program was launched in 2005.

Table 3.1: Correlation between NDVI and survey-based agricultural output

	Production (LSMS-ISA)			Productivity (LSMS-ISA)		
	(1)	(2)	(3)	(4)	(5)	(6)
NDVI (MODIS)	0.606***	0.581***	0.739***	0.396***	0.599***	0.774***
	(0.109)	(0.202)	(0.212)	(0.084)	(0.185)	(0.188)
Woredas FE		✓	✓		✓	✓
Time FE			✓			✓
Observations	480	478	478	476	470	470
R-squared	0.14	0.84	0.86	0.11	0.73	0.76

Notes: Data on agricultural output comes from the Ethiopian 2013 and 2015 LSMS-ISA surveys. In columns (1)-(3), the dependant variable corresponds to the overall production in woreda w at time t (with $t = 2013 \mid 2015$). In columns (4)-(6), the dependant variable corresponds to the average production per hectare in woreda w at time t. An inverse hyperbolic sine (IHS) transformation has been applied to all dependant variables. OLS estimator is used for all regressions. Standard errors in parentheses are clustered at woreda level. *** p<0.01, ** p<0.05, * p<0.1.

While these results are encouraging, it is worth noting that the correlation between our satellite-based variable and the true (unobserved) agricultural output may be even stronger than suggested in Table 3.1 due to some shortcomings in the LSMS-ISA measures of crop production. First, the GPS coordinates included in LSMS-ISA datasets have been slightly modified (with a random noise of 0-10km) to preserve anonymity, such that we cannot exclude that some plots have been assigned to neighboring woredas. Second, because LSMS-ISA surveys are typically not designed to be representative at the woreda level, agricultural production may lack precision.[10] Finally, agricultural data on harvest and cultivated area are typically recalled with errors (Beegle et al., 2012; Carletto et al., 2015), thereby further obscuring the precision of survey-based measures of agricultural output. The fact that our NDVI variable is still

production are missing for most of the plots because of implementation issues. We are therefore not able to incorporate these data in our investigation of the NDVI proxy.

[10]The sampling frame included 15 households per enumeration areas. A majority of woredas have only one enumeration area at most.

correlated with these survey-based measures is pretty reassuring. If anything, the true (unobserved) predictive power of our satellite-based proxy should be magnified.

As mentioned above, PSNP works may also have had effects at the extensive margin. Likewise, we tried to design a satellite-based outcome to capture these effects but were admittedly unsuccessful. Specifically, we computed for each years of the 2001-2013 period the share of cultivated areas by woredas using MODIS Land Use database, and checked the predictive power of this variable using cultivated areas derived from the 2013 and 2015 LSMS-ISA surveys. While satellite-based cultivated area appear to be a good predictor in levels, it does not appear to be capturing variations properly.[11] This signals that the use of this proxy is not suited for our purpose. We therefore prefer to leave investigations on the extensive margin to future research.

3.3.2 Treatment variables

Data on program implementation are drawn from UN Office for the Coordination of Humanitarian Affaires (UNOCHA). To investigate the impact of the PSNP, we use three definitions of the treatment. We first use a basic dummy variable taking the value one if the woreda received the treatment and zero otherwise. Figure 3.1a below represents the woredas that received the PSNP. Second, we use a treatment intensity variable, defined as the percentage of population targeted by the program in 2006 (Figure 3.1b). Data on treatment intensity present two noteworthy shortcomings. First, it provides intervals of treatment intensity instead of precise percentages. The intervals are the following: (i) 2-13%; (ii) 14-25%; (iii) 26-42%; (iv) 43-65%; and (v) 66-90%. We choose to use an ordered categorical variable taking the value 0 (if the woreda is not treated) to 5 (if the woreda received the highest treatment intensity). Second, because we could only found these data for the year 2006, we have to assume that treatment intensity remained stable over time.[12,13] Last, we use a treatment density variable, defined as the number of beneficiaries per square kilometers in each woreda (Figure 3.1c). In particular, we rely on population density estimates from 2005, and approximate exact treatment intensity using the median value

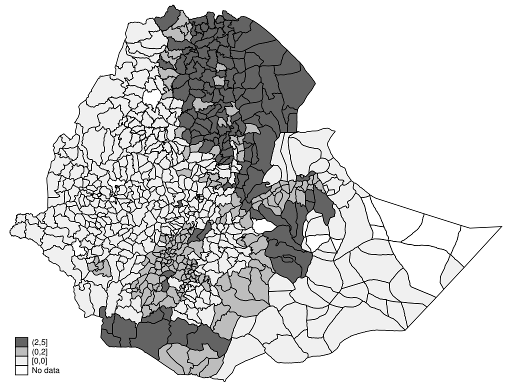
[11] Results are presented in Table C1.

[12] While there is some qualitative evidence suggesting that treatment intensity varied in absolute terms, this should not be problematic as long as relative treatment intensity remained stable across woredas – an assumption which is more likely to hold.

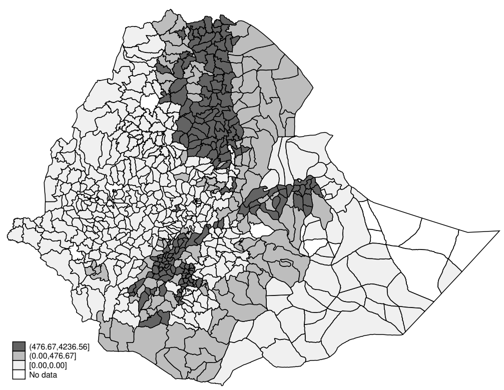
[13] An additional issue is that a few woredas were not yet beneficiaries of the program in 2006, and join it later. For these woredas, we are not able to observe the treatment intensity. One possibility would be to drop these observations from the analysis. Alternatively, we could assign the average treatment intensity derived from the sample of woredas with positive values. To limit power losses, and because it concerns only 15 woredas, we prefer the latter option.

Figure 3.1: Woredas covered by the PSNP

(a) Treatment dummy



(b) Treatment intensity



(c) Treatment density

Notes: Authors' elaboration from UNOCHA data.

over each of the five intervals (e.g. for the interval 2-13% we derive a treatment intensity of 7.5%).

3.4 Empirical analysis

In order to investigate the impact of the PSNP on crop production, we estimate the following difference-in-differences (DID) model using the OLS estimator:

$$Y_{wt} = \beta_0 + \beta_1 Treated_w \times Post_t + \boldsymbol{X}' \alpha + \nu_w + \gamma_t + \varepsilon_{wt} \tag{3.1}$$

where β_1 gives the average treatment effect of interest; Y_{wt} is the average NDVI for woreda w at time t; *Treated* corresponds to one of the three treatment variables defined in Section 3.3 (i.e. the treatment dummy, intensity, or density);

Post is a dummy variable taking the value one for post-program years (2005-2013 in most woredas), and zero otherwise;[14] $\boldsymbol{X}$ is a vector of time varying control variables including rainfall, temperature, and their respective quadratic terms drawn from CHIRPS database;[15] ν_w is a vector of woreda fixed effects controlling for time-invariant factors; γ_t is a vector of year fixed effects controlling for common shocks; and ε_{wt} is the error term.

As mentioned above, Ethiopia is known for its large ecological disparities, especially between highlands and lowlands. Because these heterogeneities could be an important factor mediating PSNP effects, we conduct the analysis separately for highlands and lowlands. We follow Hurni (1983) and define as highlands all woredas with mean elevation above 1500m.[16]

The crucial assumption underlying DID models is the parallel trend assumption. That is, in the absence of treatment, the difference between the treated and controls would have remained constant over time. This assumption seems rather strong in our setting because the program was targeted towards chronically food insecure woredas, i.e., woredas that required frequent food assistance prior to 2005 and could therefore present unobserved time-varying peculiarities. One of the main advantage of our dataset is that it includes multiple time periods prior to PSNP roll-out so that we can actually check whether the parallel trend assumption holds prior to the program. Figures 3.2a and 3.2d plot the evolution of the NDVI for the treated and control groups. While both figures are purely descriptive in nature, they tend to suggest that NDVI did not follow similar paths in the two groups before the program. To test in a more systematic and comprehensive way whether there were specific pre-trends in NDVI across treatment groups, we use a regression model similar to equation (3.1), but incorporating interactions between the treatment and each of the pre-program year dummies. Results are presented in Table 3.2. The significant interactions in columns 1-2 and 5-6 confirm the intuitions from Figures 3.2a and 3.2d. The two groups were already following distinct paths prior to public works implementation, and it would therefore be pretty implausible to assume parallel trends post-program.

[14] For some treated woredas who started to receive public works only in 2009-10 the variable *Post* takes the value one only from 2010 onwards.

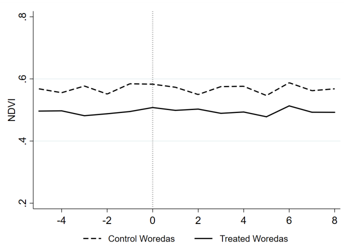
[15] We may be tempted to include additional variables such as nighttime lights or population density to control for economic and demographic dynamics. However, because these variables could be themselves affected by the treatment, we prefer to exclude them from the model. In robustness analysis, we will check whether including these variables affect the main results.

[16] As a robustness check, we will present the main results by elevation deciles instead of using an arbitrary cut-off (Figure C1).

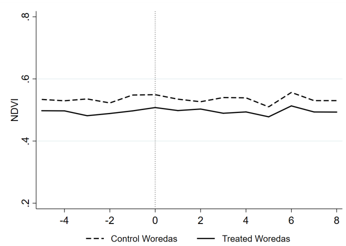
Figure 3.2: NDVI trends in treatment and control woredas

Panel A: Highlands

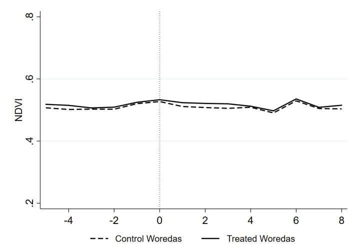
(a) Raw means



(b) Raw means
(common support)

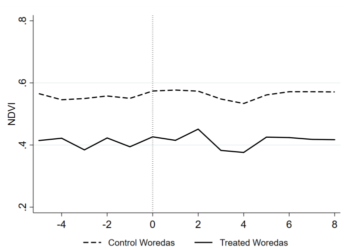


(c) IP-weighted means
(common support)

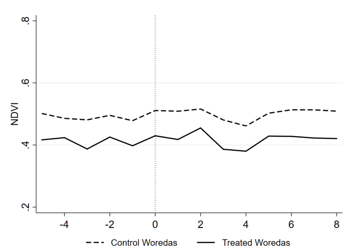


Panel B: Lowlands

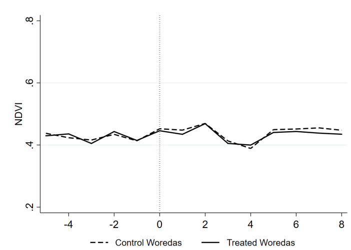
(d) Raw means



(e) Raw means
(common support)



(f) IP-weighted means
(common support)



Notes: Authors' elaboration from MODIS data. Each sub-figure compares NDVI trends between treated and control woredas. Dotted lines display PSNP rollout.

Table 3.2: Pre-treatment trends

	Highlands				Lowlands			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	NDVI	NDVI	NDVI	NDVI	NDVI	NDVI	NDVI	NDVI
Treatment × 2001	0.013***	0.015***	0.007*	-0.001	0.027***	0.032***	0.028***	0.020***
	(0.001)	(0.003)	(0.004)	(0.004)	(0.004)	(0.006)	(0.006)	(0.008)
Treatment × 2002	-0.003**	-0.005*	-0.005	-0.001	0.003	-0.002	0.000	-0.002
	(0.001)	(0.003)	(0.003)	(0.006)	(0.004)	(0.006)	(0.007)	(0.008)
Treatment × 2003	-0.003***	-0.003	0.000	0.005	-0.011***	-0.005	0.001	0.000
	(0.001)	(0.003)	(0.003)	(0.006)	(0.003)	(0.005)	(0.006)	(0.007)
Treatment × 2004	0.025***	0.016***	0.007*	0.003	0.029***	0.019***	0.018**	0.007
	(0.002)	(0.003)	(0.004)	(0.004)	(0.006)	(0.007)	(0.008)	(0.008)
Treatment × 2005	-0.014***	-0.015***	-0.014***	0.005	-0.014***	0.001	0.006	0.009
	(0.001)	(0.003)	(0.003)	(0.006)	(0.003)	(0.005)	(0.006)	(0.006)
Woredas FE	✓	✓	✓	✓	✓	✓	✓	✓
Time FE	✓	✓	✓	✓	✓	✓	✓	✓
Time-varying controls		✓	✓	✓		✓	✓	✓
Common support			✓	✓			✓	✓
IP-weights				✓				✓
Observations	6384	6384	4606	4606	2478	2478	2044	2044
R-squared	0.92	0.94	0.91	0.91	0.95	0.97	0.95	0.95

Notes: This table tests for the presence of specific pre-program trends in NDVI between treated and control woredas. The outcome variable is observed at the woreda-year level. Time varying controls include climatic variables (i.e. rainfall, temperature, and their respective quadratic terms). OLS estimator is used for all regressions, except regressions (4) and (8) where a WLS estimator with IP-weights is used. Only the interaction terms are reported due to space limitation. Standard errors in parentheses are clustered at the level of the treatment (*woredas*). *** p<0.01, ** p<0.05, * p<0.1.

One potential avenue to enrich our baseline DID model and remove these specific trends is to employ the Inverse Probability Weighting (IPW) method. This method, first pioneered by Horvitz and Thompson (1952), has been widely used in recent years to recover unbiased estimates of average treatment effects in observational studies (Austin and Stuart, 2015). The idea is to compute propensity scores (defined as the probability of treatment assignment conditional on observed covariates) to weight each observations in the empirical model.In particular, the propensity score corresponds to $e_w = Pr(T_w = 1 \mid X_w)$, where T_w is a dummy variable indicating whether woreda w is treated, and X_w is a vector of observed baseline covariates. Weights P_w are then derived as the inverse of the probability of receiving the treatment that the woreda received; in math: $P_w = \dfrac{T_w}{e_w} + \dfrac{1 - T_w}{1 - e_w}$. In sum, the IPW method gives more (resp. less) weight to (i) treated woredas with low (resp. high) propensity scores, and to (ii) control woredas with high (resp. low) propensity scores.

In practice, we first estimate the following equation using a logit estimator:

$$Treated_w = \alpha + \boldsymbol{X}'_{\boldsymbol{w}}\beta + \varepsilon_w \tag{3.2}$$

$\boldsymbol{X_w}$ is a vector of baseline covariates including climatic, geological, agricultural, demographic and economic determinants of treatment assignment. More specifically, $\boldsymbol{X}$ includes rainfall, temperature, elevation, slope, start and end months of the growing season, total area, share of cultivated area, population density, nighttime lights, and NDVI.[17] Each of these variables is averaged by woreda over the whole pre-treatment period (2000-2005). Then, we use estimates from equation (3.2) to predict propensity scores:

$$e_w = \hat{\alpha} + \boldsymbol{X}_{\boldsymbol{w}}'\hat{\beta} \tag{3.3}$$

Finally, we derive weights P_w and incorporate them in estimates of equation (3.1) using Weighted Least Squares (WLS).

The distribution of propensity scores in treated and control woredas are reported in Figure 3.3.[18] Clearly, distributions among these two groups are very different. In particular, there are large spikes of (i) control woredas with low probabilities of treatment, and of (ii) treated woredas with high probabilities of treatment. These spikes imply that the model specified in equation (3.2) is quite successful at predicting assignment to treatment, and that using IPW techniques to estimate equation (3.1) has the potential to improve estimates. However, as can also be seen from the figure, there is a non-negligible share of woredas falling outside of the common support region (i.e. no treatment or control woredas can be found for values of propensity scores close to the extremities of the distributions). To avoid that these woredas affect our estimates, we exclude them from the main regressions.[19]

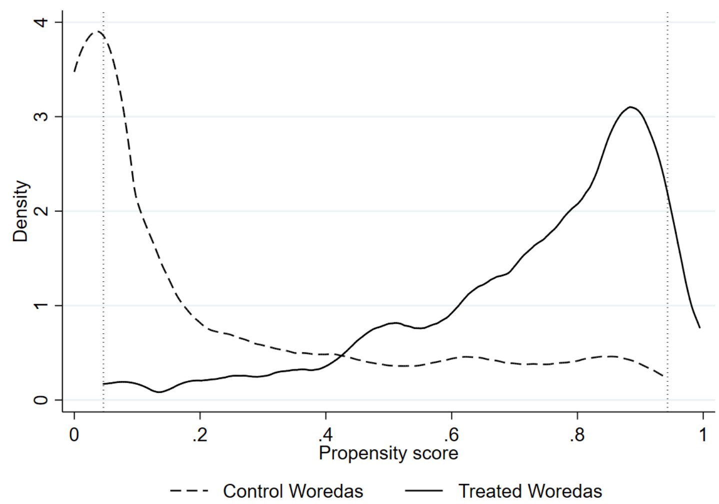
We test the validity of our common support restriction and IPW procedure in a variety of ways. First, by checking descriptively whether differences in pre-program NDVI trends between treated and controls are attenuated in Figure 3.2. As can be seen from sub-figures 3.2b and 3.2e, restricting the sample to the common support region reduces the gap between the two curves and seems to slightly improve their parallelism. In addition, IPW procedure further reduces differences between the two curves. Pre-program trends become relatively difficult to distinguish (sub-figures 3.2c and 3.2f). Second, we test more formally for the existence of significant differences in NDVI pre-trends between the two groups in Table 3.2. Results confirm that the procedures described above successfully removed specific pre-trends. As shown in Table 3.2, both the common

[17] Rainfall and temperature are drawn from CHIRPS database, elevation and slope are from the Shuttle Radar Topography Mission dataset (v4.1), start and end months of the growing season are from the MIRCA 2000 database, total area and share of cultivated area are from the MODIS land use database, population density is from Gridded Population of the World v4, nighttime lights is from DMSP-OLS (v4)

[18] Estimates of equation (3.2) are presented in Table C2.

[19] IPW techniques typically allocate low weights to woredas outside the common support region. However, because of the relatively large number of woredas concerned in our setting, we prefer to go one step further and exclude these woredas to prevent them from having any influence on the estimates. In robustness analysis, we will replicate our main analysis keeping these woredas.

Figure 3.3: Propensity scores distribution by treatment groups



Notes: This figure reports the distribution of the treatment assignment probabilities derived from equation (3.3) among treatment and control groups. Dotted lines display limits of the common support region.

support restriction and the use of IPW reduce the magnitude and significance of interaction terms. Only one out of ten interactions is statistically significant (columns 4 and 8).[20] Finally, we conduct balance tests on the common support sample. Table 3.3 clearly indicates that the IPW procedure allows to balance pre-program characteristics in the two groups. Most of the 11 characteristics tested show significant differences using unweighted means, whereas none of these differences are significant using IP-weighted means. Most importantly, the F-stat for joint significance is very low and non-significant in the IPW case. Overall, these tests provide some reassurance on the validity of our procedures and on our ability to recover unbiased estimates of public works effects. The next section presents the main results.

[20] Lowlands treated woredas experienced a significant improvement of their productivity in 2001 that is not captured by our model. However, this improvement is not too worrying because it does not persist over time.

Table 3.3: Pre-treatment characteristics by sub-samples

	Raw means			IP-weighted means		
	(1) Treated	(2) Controls	(3) Diff	(4) Treated	(5) Controls	(6) Diff
Propensity score	0.71	0.35	0.36***	0.55	0.59	-0.04
			(0.02)			(0.05)
NDVI	0.47	0.52	-0.06***	0.49	0.49	0.00
			(0.01)			(0.01)
Rainfall	532.88	706.54	-173.66***	598.60	639.30	-40.70
			(26.11)			(33.00)
Temperature	29.51	26.96	2.55***	28.37	28.47	-0.10
			(0.52)			(0.67)
Total area	0.14	0.13	0.01	0.13	0.14	0.00
			(0.02)			(0.02)
Cultivated area (% total area)	0.22	0.19	0.03**	0.23	0.23	0.00
			(0.02)			(0.03)
Start growing season	6.02	6.10	-0.08	6.03	6.05	-0.02
			(0.08)			(0.06)
End growing season	10.03	9.81	0.21**	9.99	10.00	-0.01
			(0.10)			(0.10)
Elevation	1701.57	1819.34	-117.77**	1808.47	1758.30	50.17
			(58.00)			(80.77)
Slope	5.57	4.82	0.76***	5.27	5.64	-0.37
			(0.28)			(0.54)
Population density	30.70	33.55	-2.85	28.21	30.93	-2.71
			(7.21)			(4.50)
Night time lights	0.10	0.23	-0.13	0.09	0.13	-0.04
			(0.15)			(0.08)
Observations	261	214	475	261	214	475
F-test joint significance			29.01***			0.39

Notes: Sample trimmed to common support region. The F-test corresponds to a regression of the treatment on baseline characteristics (using the same specification as in subsequent analysis). Standard errors in parentheses are clustered at the level of the treatment (*woredas*). *** p<0.01, ** p<0.05, * p<0.1.

3.5 Results

3.5.1 Treatment effects on crop productivity

We find no evidence to suggest that public works increased agricultural productivity in beneficiary woredas. Estimates of equation (3.1) using IPW and the common support restriction are reported in Table 3.4. Columns 1-3 report the results for each of the treatment variable on the sample of highland woredas. Columns 4-6 report results on the sample of lowland woredas. All estimates control for woreda fixed effects, time fixed effects, and time-varying controls including climatic variables, i.e., rainfall, temperature and their respective quadratic terms. Point estimates suggest that benefiting from public works had small and non-significant effects in both highlands and lowlands.

Table 3.4: Impacts on agricultural output

	Highlands			Lowlands		
	NDVI	NDVI	NDVI	NDVI	NDVI	NDVI
	(1)	(2)	(3)	(4)	(5)	(6)
Post × Treatment (dummy)	0.003			-0.004		
	(0.002)			(0.004)		
Post × Treatment (intensity)		0.000			-0.002	
		(0.001)			(0.001)	
Post × Treatment (density)			0.001			0.000
			(0.002)			(0.003)
Woredas FE	✓	✓	✓	✓	✓	✓
Time FE	✓	✓	✓	✓	✓	✓
Time-varying controls	✓	✓	✓	✓	✓	✓
IP-weights	✓	✓	✓	✓	✓	✓
Control mean	0.51	0.51	0.51	0.44	0.44	0.44
Observations	4606	4606	4606	2044	2044	2044
R-squared	0.91	0.91	0.91	0.95	0.95	0.95

Notes: Sample trimmed to common support region. Time varying controls include climatic variables (i.e. rainfall, temperature, and their respective quadratic terms). WLS estimator is used for all regressions. Standard errors in parentheses are clustered at the level of the treatment (*woredas*). *** p<0.01, ** p<0.05, * p<0.1.

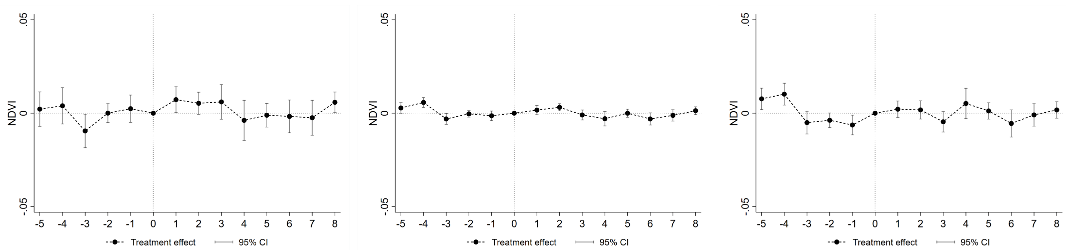
One typical concern with non-significant results is that they can reflect a lack of statistical power rather than a lack of positive effects. We argue that this is not the case in this study for at least two reasons. First, estimates on the sample of lowland woredas actually present a negative sign – a pattern inconsistent with positive effects hidden by insufficient power. Second, for highland beneficiary woredas, we can rule out even small positive effects. The upper bound of the 95 percent confidence interval on estimates for the treatment dummy is 0.007 (corresponding to a 1.5% increase relative to the control group average). Magnitude of impacts are often compared across studies in terms of standard deviations (SD). In Table C4, we replicate the results using a standardized NDVI outcome variable. We find point estimates of no more than 0.034 SD, with small standard errors. Following Ioannidis et al. (2017), we can derive the minimum detectable effect size at conventional power (80%) and statistical significance (5%) by multiplying standard errors by 2.8. Using the highest standard errors of Table C4 (0.032 in column 4), we find that our study is powered to detect effects above 0.09 SD. Such effects are generally considered as negligible.

The evidence presented above is indicative of null effects over the whole 2005-2013 period. Nevertheless, null effects could mask subtle temporal patterns. In Figure 3.4, we explore the evolution of treatment effects over time. Because positive effects of public works typically take time to manifest, and because public works received by beneficiary woredas naturally accumulate over time, we might expect to see a sustained increase in observed impact over the course of the program. In practice, we find no evidence of such an increase.

Figure 3.4: Treatment effects over time

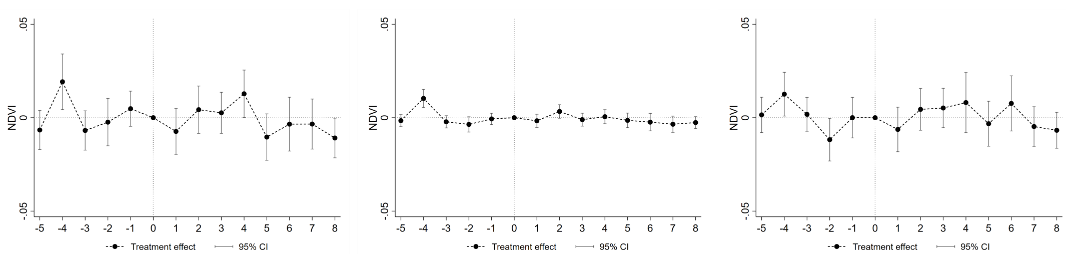
Panel A: Highlands

(a) Treatment (dummy) (b) Treatment (intensity) (c) Treatment (density)



Panel B: Lowlands

(d) Treatment (dummy) (e) Treatment (intensity) (f) Treatment (density)



Notes: Figures represent the evolution of treatment effects over time. Dotted vertical lines display PSNP rollout.

Treatment effects in late years are not particularly bigger nor more significant than treatment effects in early years. Most importantly, there is no obvious linear upward trend over the period considered in the analysis.

Finally, the impact of public works could be conditional on climatic conditions. In particular, the nature of PSNP works (e.g. land improvements, soil and water conservation measures) could help to mitigate adverse climate shocks such as droughts. For instance, a World Bank official document argues that "*the works have been found to bring demonstrable benefits to farmers from the conservation of moisture, which not only leads to visibly improved plant growth close to the bunds, but also to an increase in ground water recharge such that dry springs have started to flow again and local stream flows have increased*" (World Bank, 2006). This improvement in water resources availability and management could make beneficiary woredas more resilient to rainfall deviations. To explore these potential effects, we incorporate a triple-interaction $Treated_w \times Post_t \times Rainfall_{wt}$ in our main model. As can be seen from the signs of the triple-interactions reported in Table 3.5, crop productivity in beneficiary woredas seems to be less sensitive to high levels of rainfall. However, evidence on these effects remain limited as only one of the six interactions is significant at conventional levels.

Table 3.5: Triple difference

	Highlands			Lowlands		
	NDVI	NDVI	NDVI	NDVI	NDVI	NDVI
	(1)	(2)	(3)	(4)	(5)	(6)
Post × Treatment (dummy) × Rainfall	-0.004			-0.005		
	(0.005)			(0.005)		
Post × Treatment (dummy)	0.008			0.003		
	(0.007)			(0.007)		
Post × Treatment (intensity) × Rainfall		-0.002			-0.004**	
		(0.002)			(0.002)	
Post × Treatment (intensity)		0.002			0.002	
		(0.002)			(0.002)	
Post × Treatment (density) × Rainfall			-0.006			0.004
			(0.004)			(0.011)
Post × Treatment (density)			0.008			-0.002
			(0.005)			(0.010)
Woredas FE	✓	✓	✓	✓	✓	✓
Time FE	✓	✓	✓	✓	✓	✓
Time-varying controls	✓	✓	✓	✓	✓	✓
IP-weights	✓	✓	✓	✓	✓	✓
Control mean	0.51	0.51	0.51	0.44	0.44	0.44
Observations	4606	4606	4606	2044	2044	2044
R-squared	0.91	0.91	0.91	0.95	0.95	0.95

Notes: Annual rainfall expressed in meters. Standard errors in parentheses are clustered at the level of the treatment (*woredas*). *** p<0.01, ** p<0.05, * p<0.1. See notes to Table 3.4 for other details.

3.5.2 Robustness checks

We first explore the robustness of our findings to three variations to the main specification: (i) adding control variables to account for economic and demographic trends; (ii) estimating the main effects keeping woredas outside of the common support region; and (iii) estimating the main effects by elevation deciles instead of using an ad hoc cut-off to distinguish lowlands and highlands. Results, presented respectively in Table C5, Table C6, and Figure C1, are qualitatively unchanged. Adding control variables or keeping woredas outside of the common support region, only one coefficient becomes significant (column (5) in Table C6). In addition, no clear pattern is visible from estimates by elevation deciles.

Then, we investigate alternative explanations for the observed effects. Two main stories could threaten our interpretation: (i) a negative effect of PSNP transfers on agricultural productivity; (ii) an increase in net emigration from beneficiary woredas. Regarding the first threat, as mentioned in Section 3.2, beneficiary households received cash or food transfers in exchange for their participation in public works. This could be problematic for our estimates if these transfers had a negative effect on agricultural productivity, creating a downward bias in our estimates. We argue that it is unlikely to be the case for at least two reasons. First, the literature actually suggests that if anything PSNP transfers slightly increased crop productivity through the use of better technology such as fertilizers (Hoddinott et al., 2012). Second, we investigate whether PSNP transfers could have diverted beneficiaries from agriculture by looking at whether the share of pixels cultivated in 2005 and still cultivated in 2013 are affected by the treatment. Results, presented in Table C7, suggest that the program had no such effects.

An increase in net emigration from beneficiary woredas could also introduce a downward bias in our estimates if, for instance, it reduced the availability of labor for agriculture. Theoretically, the impact of a social protection program such as the PSNP on net emigration is ambiguous. On the one hand, it could increase emigration by relaxing financial and risk constraints typically faced by poor households (Angelucci, 2015; Gazeaud et al., 2019a). On the other hand, it could reduce emigration through increased opportunity costs (Imbert and Papp, 2016), or increase immigration by making beneficiary woredas more attractive to aspiring migrants. Because of a lack of data on migration flows, especially at relatively disaggregated levels, it is empirically challenging to investigate program effects on net emigration. We use data from the 2007 census to provide suggestive evidence that the program did not increase net emigration from beneficiary woredas.[21] We first compute domestic immigration rates (per 1,000 individuals) for each woreda over the 2000-2007 period, and then check whether the program had any effect using the main specification from Section 3.4. As can be seen from Table C8, the program did not seem to impact significantly immigration to beneficiary woredas. Because domestic immigration and domestic emigration are two sides of the same coin, and international migration flows are negligible in rural Ethiopia,[22] we argue that measuring the effect of the program on domestic immigration is actually the reverse of measuring the effect of the program on net emigration. We conclude from this exercise that the lack of impact on agricultural productivity does not appear to be explained by an increase in net emigration from beneficiary woredas.

[21] A census was conducted in 2017 but data still have to be released. We are not aware of other datasets allowing to explore PSNP effects on migration.

[22] For example, in 2007, international immigrants represent only 0.1% of all Ethiopians and 0.8% of all immigrants (author's estimates using data from the 2007 census).

3.6 Conclusion

This paper contributes to the literature by providing rigorous estimates of the productive value of PWP in the Ethiopian context. We find no evidence to support that public works had measurable impacts on agricultural productivity and resilience to climate shocks. We conduct several robustness checks to assess the validity of our results. First, since our main specification remains parsimonious in terms of control variables, we controlled for additional factors. Second, we included woredas lying outside the common support region. Last, to overcome the discretionary cutoff for highlands and lowlands woredas, we estimated the main model by elevation deciles. In all cases, the effects of the PSNP remained quantitatively small and non-significant.

These results point out that PWP infrastructures do not always generate significant and measurable productive effects. Further research is required to see whether the lack of effects of PWP infrastructures is specific to this particular setting or has some external validity. More research to understand the mechanisms would also be particularly welcome. Still, we believe that this study provides a useful piece to the debate surrounding the design and implementation of efficient social safety net programs in developing countries. The Ethiopian PSNP is Africa's flagship PWP and has probably been playing a crucial role in the decision of 38 other African countries to implement government-supported PWP (World Bank, 2015). If anything, our results call for more attention to the benefits that development practitioners typically attribute to PWP infrastructures.

Bibliography

Alderman, H. and Yemtsov, R. (2013). How Can Safety Nets Contribute to Economic Growth? *The World Bank Economic Review*, 28(1):1–20.

Ali, D. A., Deininger, K., and Monchuk, D. (2018). Using satellite imagery to assess impacts of soil and water conservation measures: Evidence from Ethiopia's Tana-Beles Watershed. The World Bank.

Andersson, C., Mekonnen, A., and Stage, J. (2011). Impacts of the productive safety net program in Ethiopia on livestock and tree holdings of rural households. *Journal of Development Economics*, 94(1):119–126.

Angelucci, M. (2015). Migration and financial constraints: Evidence from Mexico. *Review of Economics and Statistics*, 97(1):224–228.

Austin, P. C. and Stuart, E. A. (2015). Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. *Statistics in Medicine*, 34(28):3661–3679.

Baird, S., Ferreira, F. H., Özler, B., and Woolcock, M. (2014). Conditional, unconditional and everything in between: A systematic review of the effects of cash transfer programmes on schooling outcomes. *Journal of Development Effectiveness*, 6(1):1–43.

Beegle, K., Carletto, C., and Himelein, K. (2012). Reliability of recall in agricultural data. *Journal of Development Economics*, 1(98):34–41.

BenYishay, A., Runfola, D., Trichler, R., Dolan, C., Goodman, S., Parks, B., Tanner, J., Heuser, S., Batra, G., and Anand, A. (2017). A primer on geospatial impact evaluation methods, tools, and applications. AidData Working Paper#44.

Berhane, G., Gilligan, D. O., Hoddinott, J., Kumar, N., and Taffesse, A. S. (2014). Can social protection work in Africa? The impact of Ethiopia's productive safety net programme. *Economic Development and Cultural Change*, 63(1):1–26.

Carletto, C., Gourlay, S., and Winters, P. (2015). From guesstimates to GPStimates: Land area measurement and implications for agricultural analysis. *Journal of African Economies*, 24(5):593–628.

Christian, S., de Janvry, A., Egel, D., and Sadoulet, E. (2015). Quantitative Evaluation of the Social Fund for Development Labor Intensive Works Program (LIWP).

Debela, B. L., Shively, G., and Holden, S. T. (2015). Does Ethiopia's productive safety net program improve child nutrition? *Food Security*, 7(6):1273–1289.

Devereux, S. (2000). Famine in the twentieth century. IDS Working Paper, no. 105.

European Commission (2015). Ethiopia's safety net programme enhances climate change resilience of vulnerable populations. https://ec.europa.eu/europeaid/case-studies/ethiopias-safety-net-programme-enhances-climate-change-resilience-vulnerable_en. Online; accessed 2019-07-30.

Favara, M., Porter, C., and Woldehanna, T. (2019). Smarter through social protection? Evaluating the impact of Ethiopia's safety-net on child cognitive abilities. *Oxford Development Studies*, 47(1):79–96.

Filipski, M., Taylor, J. E., Abegaz, G. A., Ferede, T., Taffesse, A. S., and Diao, X. (2016). General equilibrium impact assessment of the Productive Safety Net Program in Ethiopia, 3ie Grantee Final Report. *International Initiative for Impact Evaluation (3ie)*.

Gazeaud, J., , Mvukiyehe, E., and Sterck, O. (2019a). Cash transfers and migration: Experimental evidence from Comoros. mimeo.

Gazeaud, J., , Mvukiyehe, E., and Sterck, O. (2019b). Public works and welfare: A randomized control trial of the Comoros Social Safety Net Project. Endline Report.

Gehrke, E. and Hartwig, R. (2018). Productive effects of public works programs: What do we know? What should we know? *World Development*, 107:111–124.

Gilligan, D. O., Hoddinott, J., and Taffesse, A. S. (2009). The impact of Ethiopia's productive safety net programme and its linkages. *Journal of Development Studies*, 45(10):1684–1706.

Haushofer, J. and Shapiro, J. (2016). The short-term impact of unconditional cash transfers to the poor: Experimental evidence from Kenya. *The Quarterly Journal of Economics*, 131(4):1973–2042.

Hirano, K., Imbens, G. W., and Ridder, G. (2003). Efficient estimation of average treatment effects using the estimated propensity score. *Econometrica*, 71(4):1161–1189.

Hoddinott, J., Berhane, G., Gilligan, D. O., Kumar, N., and Seyoum Taffesse, A. (2012). The impact of Ethiopia's productive safety net programme and related transfers on agricultural productivity. *Journal of African Economies*, 21(5):761–786.

Hoddinott, J., Gilligan, D. O., and Taffesse, A. S. (2010). The impact of Ethiopia's productive safety net program on schooling and child labor. *Social protection for Africa's children*.

Horvitz, D. and Thompson, D. (1952). A generalization of sampling without replacement from a finite population. *Journal of the American Statistical Association*, 47:663–685.

Hurni, H. (1983). Ethiopian highlands reclamation study. Food and Agriculture Organization of the United Nations (FAO).

Imbert, C. and Papp, J. (2016). Short-term migration rural workfare programs and urban labor markets-evidence from India. mimeo.

Ioannidis, J. P., Stanley, T., and Doucouliagos, H. (2017). The power of bias in economics research. *The Economic Journal*, 127(605):F236–F265.

Jayne, T. S., Strauss, J., Yamano, T., and Molla, D. (2002). Targeting of food aid in rural Ethiopia: Chronic need or inertia? *Journal of Development Economics*, 68(2):247–288.

Kaur, N., Agrawal, A., Steinbach, D., Panjiyar, A., Saigal, S., Manuel, C., Barnwal, A., Shakya, C., Norton, A., Kumar, N., Soanes, M., and Venkataramani, V. (2019). Building resilience to climate change through social protection: Lessons from MGNREGS, India. IIED Working Paper.

Kehler, A. (2004). When will Ethiopia stop asking for food aid. *Humanitarian Exchange Magazine*, 27:1–4.

Mendola, M. and Negasi, M. Y. (2019). Nutritional and schooling impact of a cash transfer program in Ethiopia: A retrospective analysis of childhood experience. mimeo.

Murgai, R., Ravallion, M., and Van de Walle, D. (2015). Is workfare cost-effective against poverty in a poor labor-surplus economy? *The World Bank Economic Review*, 30(3):413–445.

Pettorelli, N., Vik, J. O., Mysterud, A., Gaillard, J.-M., Tucker, C. J., and Stenseth, N. C. (2005). Using the satellite-derived NDVI to assess ecological responses to environmental change. *Trends in Ecology & Evolution*, 20(9):503–510.

Porter, C. and Goyal, R. (2016). Social protection for all ages? Impacts of Ethiopia's productive safety net program on child nutrition. *Social Science & Medicine*, 159:92–99.

Roodman, D. (2009). A note on the theme of too many instruments. *Oxford Bulletin of Economics and Statistics*, 71(1):135–158.

Sharp, K., Brown, T., and Teshome, A. (2006). *Targeting Ethiopia's productive safety net programme (PSNP)*. London and Bristol, UK: Overseas Development Institute and the IDL Group.

Subbarao, K., Del Ninno, C., Andrews, C., and Rodríguez-Alas, C. (2012). *Public works as a safety net: Design, evidence, and implementation*. The World Bank.

Wang, J., Rich, P. M., Price, K. P., and Kettle, W. D. (2005). Relations between NDVI, grassland production, and crop yield in the central great plains. *Geocarto International*, 20(3):5–11.

World Bank (2006). Ethiopia Second Productive Safety Nets Project. Project Apraisal Document. Washington, DC: World Bank.

World Bank (2015). The state of social safety nets 2015. Washington, DC: World Bank.

World Bank (2016). Ethiopia Third Productive Safety Nets Project. Implementation Completion and Results Report. Washington, DC: World Bank.

Conclusion

Over the last few decades, social safety nets have emerged as the public policy of choice to fight poverty and vulnerability in developing countries. According to World Bank estimates, around 2.5 billion people are covered by safety net programs, and an average of US$ 106 per citizen is spent annually in low and middle-income countries on these programs. A large evidence base has shown that social safety nets can achieve significant welfare gains, but many unresolved questions remain, especially regarding the many details conditioning successful programs.

This thesis contributes to the literature by providing empirical essays on the design and evaluation of social safety nets. In particular, it pays attention to three questions that have remained largely unexplored: (i) What are the implications of non-random errors in consumption data on PMT targeting performances? (ii) What are the effects of cash transfers on migration? (iii) What are the productive effects of cash-for-work activities?

Of course, this thesis does not provide definitive answers. However, it does present novel evidence showing that: (i) non-random errors in consumption can lead to a bias in PMT estimates of absolute poverty; (ii) cash transfers can ease liquidity constraints and risk bearing, thereby increasing risky migrations; (iii) cash-for-work activities do not always have measurable productive effects. Hopefully, these pieces will contribute to improve our understanding of social safety nets, and, ultimately, to design more effective public policies to fight poverty and vulnerability.

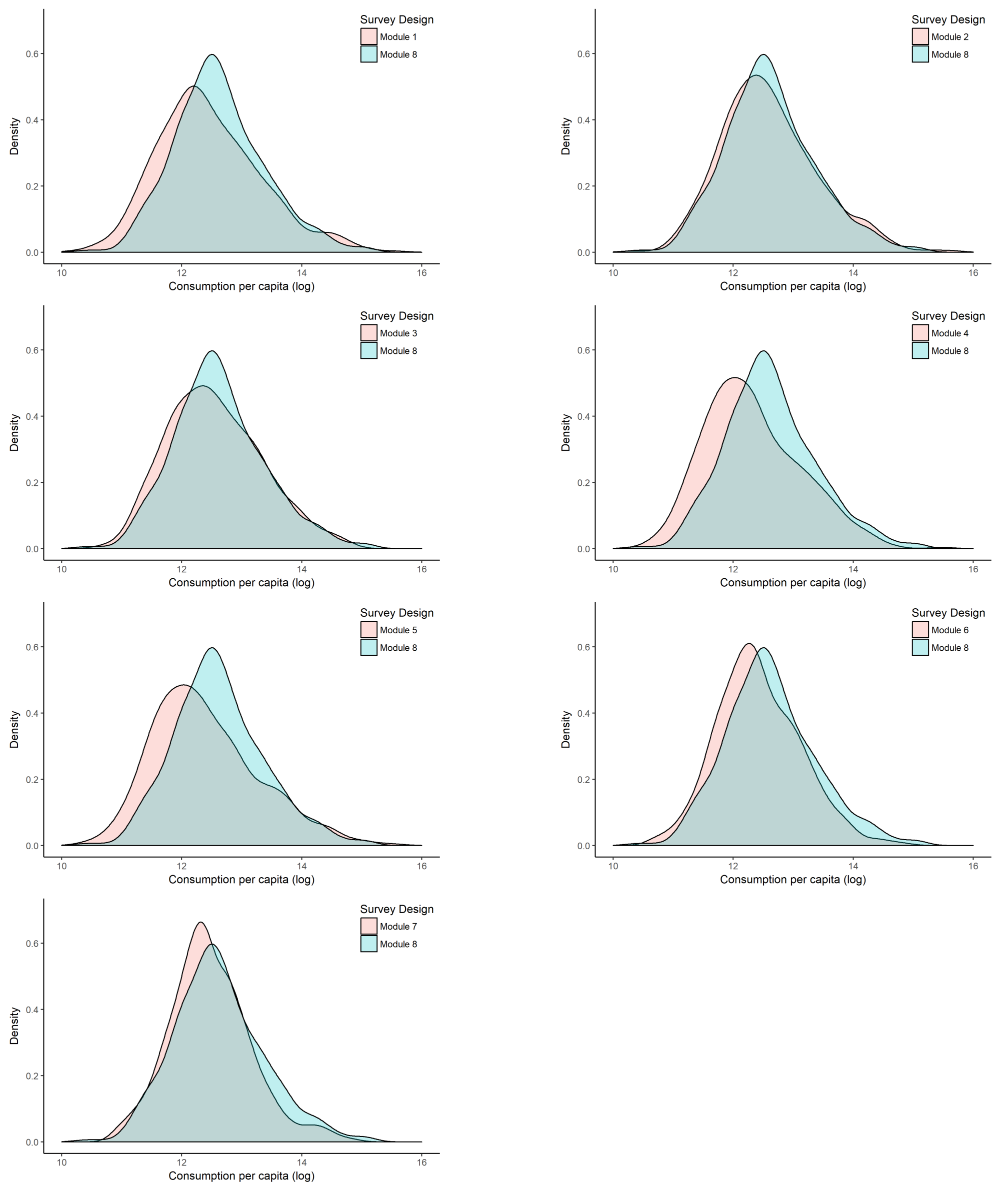
This thesis also raises a number of questions: What is the external validity of results obtained from various settings such as a survey experiment in Tanzania, a randomized cash-for-work intervention in Comoros, and a targeted public works program in Ethiopia? What are the long-run effects of cash transfers on migration? Why cash-for-work activities do not always yield measurable productive effects? How vulnerable to the presence of measurement errors are the performances of machine learning tools trained on survey data? Further research around these areas would be particularly welcome.

Appendix

Appendix A

Appendix to Chapter 1

Figure A1: Comparing distributions of consumption by survey design



Notes: Each figure compares the distribution of consumption of households assigned to Module 8 (the benchmark) with distributions of households assigned to module k (with $k = \{1,7\}$).

Table A1: Summary statistics

Variable	Mean	Std. Dev.	Min.	Max.	N
Ln conso	12.48	0.79	10.21	15.65	4025
hhsize	5.28	2.88	1	23	4025
children5	1.08	1.1	0	8	4025
children14	1.42	1.41	0	12	4025
elderly	0.33	0.59	0	3	4025
age	46.65	16.3	17	96	4025
male	0.8	0.4	0	1	4025
litteracy	0.65	0.48	0	1	4025
primary	0.72	0.45	0	1	4025
secondary	0.09	0.29	0	1	4025
primarymax	0.92	0.26	0	1	4025
secondarymax	0.2	0.4	0	1	4025
married	0.74	0.44	0	1	4025
widowed	0.13	0.34	0	1	4025
floor	0.57	0.49	0	1	4025
roof	0.34	0.47	0	1	4025
wall	0.72	0.45	0	1	4025
room	3.57	1.8	1	18	4025
electricity	0.14	0.34	0	1	4025
water	0.27	0.44	0	1	4025
flushedtoilet	0.1	0.31	0	1	4025
cooking	0.22	0.42	0	1	4025
urban	0.34	0.48	0	1	4025
mobile	0.3	0.46	0	1	4025
tv	0.1	0.29	0	1	4025
radio	0.6	0.49	0	1	4025
watch	0.43	0.5	0	1	4025
bicycle	0.44	0.5	0	1	4025
iron	0.25	0.43	0	1	4025
refrigirator	0.05	0.22	0	1	4025
mattress	0.83	0.38	0	1	4025
sewing_machine	0.07	0.26	0	1	4025
improved_stove	0.12	0.33	0	1	4025
motorcycle	0.02	0.14	0	1	4025
car	0.02	0.14	0	1	4025
wheelbarrow	0.04	0.21	0	1	4025
cattle	0.15	0.36	0	1	4025
sheep	0.05	0.22	0	1	4025
goat	0.25	0.43	0	1	4025
chicken	0.5	0.5	0	1	4025
land_owned	0.8	0.4	0	1	4025
land_rented	0.28	0.45	0	1	4025

Table A2: Description of variables

Dependent variables	
lnconso	Consumption per capita (in log)
Demographic characteristics	
hhsize	Household size
hhsize^2	Squared household size
young children	Number of children (0-5)
children	Number of children (6-14)
elderly	Number of elderly (65+)
primary max	=1 if at least one household member attended primary
secondary max	=1 if at least one household member attended secondary
Household head characteristics	
married	Household head is married
widowed	Household head is widowed
age	Household head's age
age^2	Squared household's head age
primary	Household head attended primary
secondary	Household head attended secondary
male	Household head is male
Dwelling characteristics	
mud/dirt floor	Floor is mud or dirt
thatch roof	Roof is thatched
mud walls	Walls are mud
n rooms	Number of rooms
electricity	Household has access to electricity
urban	Household is urban
water	Water is from piped or from covered well or from vendor
flushed toilet	Household has flushed toilet
cooking	Main fuel for cooking is not firewood
Assets	
mobile	Household has a mobile phone
TV	Household has a TV
radio	Household has a radio
watch	Household has a watch
bicycle	Household has a sewing machine
iron	Household has an improved stove
refrigerator	Household has a refrigerator
mattress	Household has a mattress
sewing machine	Household has a sewing machine
improved stove	Household has an improved stove
motorcycle	Household has a motorcycle
car	Household has a car
wheelbarrow	Household has a wheelbarrow
cattle	Household has cattle
sheep	Household has sheep
goat	Household has goat
chicken	Household has chicken
land ownership	Household owns land
land rented	Household rents land

Table A3: Balance Table

Variable	(1) Module 1 Mean/SE	(2) Module 2 Mean/SE	(3) Module 3 Mean/SE	(4) Module 4 Mean/SE	(5) Module 5 Mean/SE	(6) Module 6 Mean/SE	(7) Module 7 Mean/SE	(8) Module 8 Mean/SE	T-test Difference (8)-(1)	(8)-(2)	(8)-(3)	(8)-(4)	(8)-(5)	(8)-(6)	(8)-(7)
Hhsize	5.227	5.153	5.155	5.460	5.282	5.337	5.317	5.280	0.054	0.128	0.126	-0.180	-0.001	-0.056	-0.037
	(0.157)	(0.140)	(0.140)	(0.142)	(0.139)	(0.153)	(0.157)	(0.151)							
Young Children	1.083	1.016	1.065	1.141	1.069	1.068	1.070	1.093	0.010	0.078	0.028	-0.047	0.024	0.026	0.024
	(0.060)	(0.053)	(0.055)	(0.053)	(0.053)	(0.059)	(0.058)	(0.058)							
Children	1.429	1.333	1.440	1.498	1.444	1.444	1.355	1.433	0.004	0.100	-0.007	-0.065	-0.011	-0.011	0.078
	(0.071)	(0.062)	(0.071)	(0.070)	(0.069)	(0.067)	(0.075)	(0.071)							
Elderly	0.370	0.339	0.278	0.312	0.347	0.331	0.339	0.336	-0.034	-0.003	0.058	0.024	-0.011	0.005	-0.003
	(0.030)	(0.029)	(0.025)	(0.026)	(0.025)	(0.030)	(0.030)	(0.029)							
Married	0.742	0.732	0.720	0.730	0.718	0.747	0.737	0.763	0.022	0.031	0.043	0.033	0.045	0.016	0.027
	(0.022)	(0.021)	(0.022)	(0.021)	(0.021)	(0.020)	(0.021)	(0.021)							
Widowed	0.125	0.113	0.133	0.135	0.143	0.124	0.160	0.135	0.010	0.022	0.002	0.000	-0.008	0.012	-0.024
	(0.015)	(0.015)	(0.016)	(0.015)	(0.016)	(0.016)	(0.017)	(0.016)							
Age	47.628	46.192	46.048	46.419	46.532	46.629	46.988	46.803	-0.825	0.611	0.756	0.385	0.271	0.174	-0.185
	(0.838)	(0.756)	(0.762)	(0.751)	(0.717)	(0.765)	(0.809)	(0.809)							
Male	0.811	0.802	0.794	0.792	0.784	0.819	0.788	0.791	-0.020	-0.010	-0.002	-0.000	0.008	-0.027	0.003
	(0.019)	(0.019)	(0.019)	(0.018)	(0.019)	(0.018)	(0.019)	(0.019)							
Primary	0.710	0.726	0.738	0.710	0.708	0.731	0.715	0.706	-0.004	-0.020	-0.032	-0.005	-0.003	-0.025	-0.009
	(0.022)	(0.022)	(0.023)	(0.023)	(0.022)	(0.021)	(0.024)	(0.022)							
Secondary	0.082	0.101	0.087	0.097	0.091	0.088	0.102	0.099	0.018	-0.002	0.012	0.002	0.008	0.012	-0.002
	(0.014)	(0.014)	(0.015)	(0.016)	(0.015)	(0.014)	(0.016)	(0.017)							
Primary Max	0.913	0.919	0.927	0.933	0.940	0.914	0.942	0.913	0.000	-0.006	-0.014	-0.020	-0.028*	-0.002	-0.030*
	(0.013)	(0.013)	(0.011)	(0.012)	(0.010)	(0.014)	(0.011)	(0.012)							
Secondary Max	0.195	0.208	0.208	0.196	0.204	0.211	0.210	0.195	0.000	-0.014	-0.014	-0.002	-0.010	-0.016	-0.015
	(0.020)	(0.022)	(0.021)	(0.021)	(0.021)	(0.021)	(0.022)	(0.021)							
Mud/Dirt Floor	0.571	0.577	0.581	0.565	0.571	0.572	0.561	0.577	0.006	-0.001	-0.005	0.011	0.005	0.005	0.016
	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.034)	(0.034)	(0.033)							
Thatch Roof	0.328	0.321	0.331	0.345	0.323	0.339	0.357	0.340	0.012	0.019	0.009	-0.005	0.017	0.001	-0.017
	(0.028)	(0.027)	(0.028)	(0.029)	(0.029)	(0.030)	(0.028)	(0.028)							
Mud Walls	0.698	0.714	0.712	0.732	0.710	0.735	0.739	0.702	0.004	-0.012	-0.011	-0.030	-0.009	-0.033*	-0.037**
	(0.029)	(0.029)	(0.029)	(0.028)	(0.029)	(0.029)	(0.029)	(0.030)							
N Rooms	3.529	3.581	3.492	3.615	3.558	3.614	3.607	3.598	0.070	0.017	0.106	-0.017	0.041	-0.015	-0.008
	(0.103)	(0.100)	(0.087)	(0.095)	(0.089)	(0.103)	(0.095)	(0.095)							
Electricity	0.161	0.125	0.133	0.129	0.151	0.135	0.128	0.133	-0.028	0.008	0.000	0.004	-0.018	-0.002	0.005
	(0.022)	(0.020)	(0.021)	(0.021)	(0.022)	(0.021)	(0.020)	(0.021)							
Urban	0.203	0.202	0.202	0.202	0.202	0.203	0.200	0.201	-0.002	-0.002	-0.002	-0.002	-0.002	-0.002	0.001
	(0.031)	(0.031)	(0.031)	(0.031)	(0.031)	(0.031)	(0.031)	(0.031)							
Water	0.264	0.264	0.266	0.268	0.274	0.267	0.267	0.270	0.006	0.006	0.005	0.003	-0.003	0.003	0.003
	(0.031)	(0.031)	(0.032)	(0.031)	(0.032)	(0.032)	(0.031)	(0.031)							
Flushed Toilet	0.119	0.097	0.111	0.103	0.111	0.092	0.108	0.093	-0.026*	-0.004	-0.018	-0.010	-0.018	0.002	-0.014
	(0.021)	(0.017)	(0.020)	(0.019)	(0.020)	(0.018)	(0.020)	(0.018)							
Cooking	0.229	0.240	0.234	0.222	0.226	0.221	0.204	0.215	-0.014	-0.025**	-0.019	-0.008	-0.011	-0.006	0.011
	(0.031)	(0.031)	(0.030)	(0.029)	(0.030)	(0.030)	(0.029)	(0.030)							
N	503	504	504	504	504	502	501	503							
Clusters	168	168	168	168	168	168	168	168							

Notes: The value displayed for t-tests are the difference in means between households assigned to module 8 and households assigned to each of the other modules. Standard errors in parentheses are clustered at the village level. *** p<0.01, ** p<0.05, * p<0.1.

Table A4: Survey experiment consumption modules

Module	Consumption measurement	Food content	N households
1	Long list (58 food items) 14 day	Quantity from purchases, own-production, and gifts/other sources; Tshilling value of consumption from purchases	504
2	Long list (58 food items) 7 day	Quantity from purchases, own-production, and gifts/other sources; Tshilling value of consumption from purchases	504
3	Subset list (17 food items; subset of 58 foods) 7 day	Quantity from purchases, own-production, and gifts/other sources; Tshilling value of consumption from purchases	504
4	Collapsed list (11 food items covering universe of food categories) 7 day	Tshilling value of consumption	504
5	Long list (58 food items) Usual 12 month	Consumption from purchases: number of months consumed, quantity per month, Tshilling value per month Consumption from own-production: number of months consumed, quantity per month, Tshilling value per month Consumption from gifts/other sources: total estimated value for last 12 months	504
6	Household diary, frequent visits 14 day diary		503
7	Household diary, infrequent visits 14 day diary		503
8	Personal diary, frequent visits 14 day diary		503
			4029

Notes: Frequent visits entailed daily visits by the local assistant and visits every other day by the survey enumerator for the duration of the 2-week diary. Infrequent visits entail 3 visits: to deliver the diary (day 1), to pick up week 1 diary and drop off week 2 diary (day 8), and to pick up week 2 diary (day 15). Households assigned to the infrequent diary but who had no literate members (about 18% of the 503 households) were visited every other day by the local assistant and the enumerator.
Non-food items are divided into two groups based on frequency of purchase. Frequently purchased items (charcoal, firewood, kerosene/paraffin, matches, candles, lighters, laundry soap, toilet soap, cigarettes, tobacco, cell phone and internet, transport) were collected by 14-day recall for modules 1–5 and in the 14-day diary for modules 6–8. Non-frequent non-food items (utilities, durables, clothing, health, education, contributions, and other; housing is excluded) are collected by recall identically across all modules at the end of the interview (and at the end of the 2-week period for the diaries) and over the identical one or 12-month reference period, depending on the item in question.

Source: Gibson et al. (2015)

Table A5: PMT Regressions: food consumption only

	(1) Module 1	(2) Module 2	(3) Module 3	(4) Module 4	(5) Module 5	(6) Module 6	(7) Module 7	(8) Module 8	(9) All
Hhsize	-0.220***	-0.268***	-0.174***	-0.222***	-0.204***	-0.153***	-0.181***	-0.234***	-0.204***
	(0.032)	(0.032)	(0.039)	(0.030)	(0.035)	(0.019)	(0.022)	(0.027)	(0.011)
Hhsize^2	0.010***	0.013***	0.010***	0.009***	0.010***	0.007***	0.007***	0.009***	0.009***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.001)	(0.001)	(0.002)	(0.001)
Elderly		0.104*						-0.164***	
		(0.055)						(0.048)	
Young Children	-0.085***	-0.074**	-0.145***	-0.061**	-0.130***	-0.094***	-0.051**	-0.056*	-0.088***
	(0.029)	(0.029)	(0.030)	(0.029)	(0.029)	(0.026)	(0.025)	(0.030)	(0.009)
Children			-0.048*						
			(0.027)						
Mud/Dirt Floor						-0.132**			-0.049*
						(0.055)			(0.030)
Thatch Roof				-0.142**		-0.106**	-0.146***		-0.054**
				(0.060)		(0.053)	(0.054)		(0.026)
Mud Walls	-0.246***		-0.282***	-0.185***	-0.213***	-0.169**		-0.153**	-0.153***
	(0.078)		(0.056)	(0.069)	(0.064)	(0.073)		(0.074)	(0.030)
N Rooms		0.041**		0.037**			0.074***	0.086***	0.039***
		(0.019)		(0.018)			(0.016)	(0.017)	(0.007)
Electricity		0.227**	-0.153**	0.167*		0.198***	0.316***	0.401***	0.072*
		(0.089)	(0.066)	(0.085)		(0.071)	(0.091)	(0.087)	(0.041)
Urban		0.264***						0.287***	0.089**
		(0.071)						(0.076)	(0.037)
Water									
Flushed Toilet			0.157*			0.171**	0.291***		0.143***
			(0.091)			(0.083)	(0.083)		(0.045)
Cooking	0.529***	0.434***	0.448***	0.506***	0.744***				0.324***
	(0.087)	(0.082)	(0.062)	(0.092)	(0.074)				(0.046)
Married	-0.326***	-0.119*	-0.270***			-0.279***	-0.317***	0.168***	-0.174***
	(0.111)	(0.072)	(0.084)			(0.093)	(0.105)	(0.059)	(0.041)
Widowed	-0.186*	-0.186**							-0.107***
	(0.106)	(0.087)							(0.038)
Age		0.016*		0.017*					
		(0.009)		(0.009)					
Age^2		-0.000**		-0.000**	-0.000***				-0.000***
		(0.000)		(0.000)	(0.000)				(0.000)
Male	0.377***		0.351***	0.200***		0.377***	0.292***		0.193***
	(0.117)		(0.082)	(0.063)		(0.097)	(0.106)		(0.038)
Primary		0.140**				0.114**	0.177***		0.065**
		(0.058)				(0.054)	(0.052)		(0.026)
Secondary	0.233*		0.182*		0.377***		0.228**		0.118***
	(0.118)		(0.095)		(0.076)		(0.092)		(0.039)
Primary Max					0.250**				
					(0.104)				
Secondary Max	0.152**	0.232***	0.147**	0.177***				0.141**	0.100***
	(0.076)	(0.067)	(0.070)	(0.064)				(0.070)	(0.027)
Ajusted-R^2	0.533	0.553	0.569	0.525	0.591	0.415	0.412	0.421	0.480
Observations	503	504	504	504	504	502	501	503	4025

Notes: OLS estimator is used for all regressions. Standard errors in parentheses are clustered at the village level. *** p<0.01, ** p<0.05, * p<0.1. See notes to Table 1.1 for other details.

Table A6: PMT Regressions: consumption per adult equivalent

	(1) Module 1	(2) Module 2	(3) Module 3	(4) Module 4	(5) Module 5	(6) Module 6	(7) Module 7	(8) Module 8	(9) All
Hhsize	-0.232***	-0.230***	-0.162***	-0.241***	-0.240***	-0.111***	-0.152***	-0.194***	-0.179***
	(0.035)	(0.038)	(0.043)	(0.031)	(0.040)	(0.025)	(0.027)	(0.040)	(0.015)
Hhsize^2	0.013***	0.015***	0.012***	0.011***	0.013***	0.007***	0.008***	0.010***	0.010***
	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.001)	(0.001)	(0.002)	(0.001)
Elderly			-0.077*				-0.094**	-0.241***	-0.050**
			(0.044)				(0.039)	(0.054)	(0.024)
Young Children	-0.197***	-0.210***	-0.284***	-0.170***	-0.225***	-0.214***	-0.194***	-0.195***	-0.219***
	(0.031)	(0.033)	(0.033)	(0.031)	(0.029)	(0.031)	(0.032)	(0.039)	(0.013)
Children	-0.044*	-0.070***	-0.097***			-0.072**	-0.047*	-0.087**	-0.064***
	(0.027)	(0.027)	(0.030)			(0.030)	(0.028)	(0.035)	(0.012)
Mud/Dirt Floor		-0.112*				-0.158***			-0.072**
		(0.061)				(0.057)			(0.031)
Thatch Roof				-0.147**		-0.129**	-0.185***		-0.067**
				(0.061)		(0.054)	(0.056)		(0.026)
Mud Walls	-0.306***		-0.317***	-0.267***	-0.257***	-0.222**		-0.160*	-0.205***
	(0.081)		(0.058)	(0.075)	(0.068)	(0.088)		(0.083)	(0.032)
N Rooms		0.038*		0.045**			0.069***	0.092***	0.039***
		(0.020)		(0.018)			(0.017)	(0.018)	(0.008)
Electricity		0.198**	-0.135*	0.205**		0.323***	0.329***	0.414***	
		(0.091)	(0.069)	(0.093)		(0.076)	(0.100)	(0.094)	
Urban		0.294***				0.133*		0.260***	0.100**
		(0.073)				(0.072)		(0.086)	(0.042)
Water			0.123**						
			(0.062)						
Flushed Toilet					0.183*		0.349***		0.179***
					(0.098)		(0.093)		(0.050)
Cooking	0.673***	0.493***	0.558***	0.671***	0.786***			0.201*	0.432***
	(0.091)	(0.088)	(0.071)	(0.094)	(0.078)			(0.106)	(0.051)
Married	-0.260**	-0.262***	-0.318***			-0.336***	-0.462***	0.184***	-0.243***
	(0.110)	(0.091)	(0.086)			(0.095)	(0.118)	(0.063)	(0.044)
Widowed		-0.222**							-0.127***
		(0.092)							(0.040)
Age		0.022**		0.019**					0.007**
		(0.009)		(0.010)					(0.003)
Age^2		-0.000***		-0.000***	-0.000***				-0.000***
		(0.000)		(0.000)	(0.000)				(0.000)
Male	0.470***	0.164*	0.405***	0.236***	0.154**	0.468***	0.441***		0.277***
	(0.125)	(0.085)	(0.083)	(0.065)	(0.068)	(0.099)	(0.119)		(0.040)
Primary		0.123**				0.172***	0.172***		0.074***
		(0.057)				(0.057)	(0.055)		(0.026)
Secondary	0.219*	0.180*	0.291***		0.347***		0.277***		0.162***
	(0.120)	(0.103)	(0.109)		(0.118)		(0.092)		(0.043)
Primary Max	0.277***		0.172*		0.310***				
	(0.098)		(0.092)		(0.110)				
Secondary Max	0.166**	0.168**	0.158**	0.190***	0.144*			0.197***	0.114***
	(0.082)	(0.079)	(0.073)	(0.068)	(0.079)			(0.072)	(0.029)
Ajusted-R^2	0.622	0.640	0.662	0.629	0.672	0.526	0.506	0.526	0.581
Observations	503	504	504	504	504	502	501	503	4025

Notes: OLS estimator is used for all regressions. Standard errors in parentheses are clustered at the village level. *** p<0.01, ** p<0.05, * p<0.1. See notes to Table 1.1 for other details.

Table A7: PMT Regressions: extended list of covariates

	(1) Module 1	(2) Module 2	(3) Module 3	(4) Module 4	(5) Module 5	(6) Module 6	(7) Module 7	(8) Module 8	(9) All
Hhsize	-0.260***	-0.280***	-0.174***	-0.239***	-0.263***	-0.168***	-0.187***	-0.258***	-0.226***
	(0.030)	(0.032)	(0.039)	(0.029)	(0.033)	(0.021)	(0.022)	(0.023)	(0.011)
Hhsize^2	0.010***	0.014***	0.010***	0.011***	0.013***	0.006***	0.007***	0.010***	0.009***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Elderly		0.104**						-0.143***	
		(0.053)						(0.043)	
Young Children		-0.093***	-0.144***	-0.070***	-0.113***	-0.080***	-0.065**		-0.077***
		(0.025)	(0.027)	(0.025)	(0.026)	(0.028)	(0.026)		(0.009)
Children			-0.053*						
			(0.027)						
Mud/Dirt Floor						-0.185***			
						(0.050)			
Thatch Roof				-0.133**			-0.129**		
				(0.056)			(0.053)		
Mud Walls	-0.185**		-0.188***	-0.152**	-0.137**				-0.113***
	(0.082)		(0.066)	(0.068)	(0.062)				(0.028)
N Rooms							0.047***	0.042**	0.014**
							(0.016)	(0.017)	(0.007)
Electricity	-0.235***		-0.147**		-0.206**	0.218***	0.188**	0.260***	
	(0.089)		(0.063)		(0.089)	(0.076)	(0.087)	(0.086)	
Urban		0.204***				0.146**		0.285***	0.081**
		(0.062)				(0.064)		(0.080)	(0.036)
Water			0.132**						
			(0.056)						
Flushed Toilet							0.251***		0.087**
							(0.091)		(0.043)
Cooking	0.413***	0.366***	0.408***	0.462***	0.567***			0.189*	0.290***
	(0.104)	(0.072)	(0.079)	(0.082)	(0.083)			(0.104)	(0.046)
Married	-0.441***	-0.128*	-0.324***			-0.266***	-0.345***		-0.211***
	(0.109)	(0.070)	(0.081)			(0.097)	(0.109)		(0.040)
Widowed	-0.204**	-0.179**							-0.097**
	(0.101)	(0.083)							(0.038)
Age		0.015*				-0.003*			
		(0.009)				(0.002)			
Age^2		-0.000**		-0.000**	-0.000***				-0.000***
		(0.000)		(0.000)	(0.000)				(0.000)
Male	0.380***		0.288***	0.121**		0.310***	0.258**		0.166***
	(0.109)		(0.079)	(0.057)		(0.098)	(0.113)		(0.037)
Primary							0.144**		
							(0.056)		
Secondary			0.224***		0.187**		0.168**		
			(0.083)		(0.085)		(0.083)		
Primary Max					0.181*				
					(0.104)				
Secondary Max	0.143**	0.127**						0.162**	0.077***
	(0.069)	(0.064)						(0.065)	(0.023)
Iron	0.198***		0.123**	0.199***	0.112**				0.084***
	(0.071)		(0.057)	(0.055)	(0.054)				(0.022)
Refreigirator	0.295**	0.403***		0.286**			0.247*	0.235*	0.151**
	(0.129)	(0.113)		(0.118)			(0.149)	(0.138)	(0.060)
Land Rented	-0.130**			-0.089*		-0.083*			
	(0.058)			(0.053)		(0.046)			
Mobile	0.220***	0.226***	0.262***	0.298***	0.251***	0.201***	0.235***	0.145*	0.205***
	(0.071)	(0.052)	(0.073)	(0.053)	(0.063)	(0.063)	(0.060)	(0.081)	(0.022)
Cattele	0.135**	0.135**		-0.142**		0.192***	0.210***	0.235***	0.114***
	(0.063)	(0.056)		(0.071)		(0.066)	(0.065)	(0.061)	(0.031)
Radio	0.192***	0.090*	0.185***	0.154***	0.166***		0.124**	0.149***	0.135***
	(0.055)	(0.053)	(0.048)	(0.048)	(0.049)		(0.052)	(0.050)	(0.018)
Improved Stove	0.200**	0.142*	0.144*	0.155**	0.225**				0.140***
	(0.086)	(0.073)	(0.074)	(0.075)	(0.091)				(0.031)
TV	0.226*				0.326***	0.205*		0.197**	0.136***
	(0.123)				(0.104)	(0.111)		(0.095)	(0.041)
Mattress		0.173**							0.076***
		(0.075)							(0.026)
Car		0.269*			0.493***		0.213*		0.175**
		(0.158)			(0.177)		(0.128)		(0.072)
Watch		0.151***			0.116**	0.119**		0.128***	0.069***
		(0.051)			(0.049)	(0.049)		(0.048)	(0.017)
Goat				0.158**				0.242***	0.070***
				(0.061)				(0.055)	(0.024)
Sheep				-0.173**	0.196**				
				(0.086)	(0.084)				
Land Ownership				0.122*					
				(0.065)					
Bicycle						0.085*			
						(0.047)			
Sewing Machine						0.266***		0.157*	0.064**
						(0.095)		(0.083)	(0.029)
Motorcycle						0.280*			
						(0.143)			
Chicken								0.108*	
								(0.057)	
Wheelbarrow									0.080*
									(0.044)
Ajusted-R^2	0.639	0.643	0.651	0.654	0.695	0.516	0.494	0.548	0.580
Observations	503	504	504	504	504	502	501	503	4025

Notes: OLS estimator is used for all regressions. Standard errors in parentheses are clustered at the village level. *** p<0.01, ** p<0.05, * p<0.1. See notes to Table 1.1 for other details.

Table A8: PMT Regressions: no stepwise procedure

	(1) Module 1	(2) Module 2	(3) Module 3	(4) Module 4	(5) Module 5	(6) Module 6	(7) Module 7	(8) Module 8	(9) All
Hhsize	-0.247***	-0.251***	-0.194***	-0.231***	-0.230***	-0.129***	-0.162***	-0.209***	-0.200***
	(0.039)	(0.042)	(0.045)	(0.037)	(0.043)	(0.027)	(0.031)	(0.041)	(0.015)
Hhsize^2	0.012***	0.013***	0.011***	0.010***	0.011***	0.006***	0.007***	0.009***	0.009***
	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.001)	(0.001)	(0.002)	(0.001)
Elderly	0.058	0.114*	-0.009	-0.041	-0.034	-0.013	-0.029	-0.110	-0.005
	(0.062)	(0.059)	(0.072)	(0.061)	(0.063)	(0.060)	(0.063)	(0.069)	(0.024)
Young Children	-0.081**	-0.079**	-0.139***	-0.063*	-0.122***	-0.110***	-0.081**	-0.084**	-0.097***
	(0.038)	(0.034)	(0.035)	(0.037)	(0.032)	(0.035)	(0.033)	(0.038)	(0.012)
Children	-0.005	-0.014	-0.051	-0.010	-0.003	-0.025	-0.019	-0.047	-0.024**
	(0.028)	(0.028)	(0.033)	(0.031)	(0.034)	(0.032)	(0.028)	(0.034)	(0.012)
Mud/Dirt Floor	-0.117	-0.093	-0.077	-0.079	-0.058	-0.136**	-0.066	0.004	-0.076**
	(0.077)	(0.065)	(0.063)	(0.065)	(0.065)	(0.057)	(0.066)	(0.067)	(0.031)
Thatch Roof	0.093	0.009	-0.061	-0.165***	-0.026	-0.094*	-0.154***	-0.089	-0.068***
	(0.065)	(0.066)	(0.058)	(0.059)	(0.062)	(0.051)	(0.057)	(0.061)	(0.026)
Mud Walls	-0.285***	-0.085	-0.285***	-0.221***	-0.227***	-0.175**	-0.058	-0.130	-0.193***
	(0.084)	(0.073)	(0.066)	(0.074)	(0.068)	(0.088)	(0.086)	(0.092)	(0.031)
N Rooms	0.021	0.040*	0.007	0.036*	0.028	0.028	0.070***	0.079***	0.039***
	(0.020)	(0.021)	(0.017)	(0.019)	(0.026)	(0.017)	(0.017)	(0.018)	(0.007)
Electricity	-0.092	0.151	-0.156**	0.176*	-0.118	0.247***	0.289***	0.356***	0.073*
	(0.093)	(0.103)	(0.069)	(0.090)	(0.094)	(0.082)	(0.103)	(0.091)	(0.044)
Urban	0.167**	0.262***	0.071	-0.005	0.091	0.106	-0.056	0.242***	0.102**
	(0.079)	(0.074)	(0.077)	(0.075)	(0.091)	(0.075)	(0.081)	(0.089)	(0.041)
Water	-0.058	0.071	0.073	-0.095	-0.091	0.006	0.027	0.007	-0.008
	(0.081)	(0.073)	(0.062)	(0.080)	(0.093)	(0.087)	(0.079)	(0.070)	(0.038)
Flushed Toilet	0.125	0.129	0.115	0.031	0.229**	0.124	0.290***	0.099	0.161***
	(0.099)	(0.107)	(0.096)	(0.089)	(0.112)	(0.110)	(0.101)	(0.108)	(0.053)
Cooking	0.565***	0.422***	0.500***	0.680***	0.812***	0.095	0.068	0.153	0.421***
	(0.122)	(0.090)	(0.086)	(0.100)	(0.108)	(0.107)	(0.118)	(0.103)	(0.051)
Married	-0.250**	-0.171*	-0.215**	-0.077	0.011	-0.272***	-0.370***	0.172	-0.151***
	(0.119)	(0.088)	(0.091)	(0.106)	(0.107)	(0.095)	(0.117)	(0.130)	(0.042)
Widowed	-0.147	-0.184*	0.036	-0.155	-0.016	-0.121	-0.085	0.014	-0.107***
	(0.107)	(0.095)	(0.077)	(0.100)	(0.116)	(0.093)	(0.118)	(0.107)	(0.039)
Age	0.003	0.021**	0.007	0.020**	-0.001	-0.009	0.001	0.001	0.004
	(0.009)	(0.009)	(0.009)	(0.010)	(0.008)	(0.009)	(0.010)	(0.011)	(0.003)
Age^2	-0.000	-0.000***	-0.000	-0.000**	-0.000	0.000	-0.000	-0.000	-0.000*
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Male	0.345***	0.084	0.338***	0.154	0.108	0.306***	0.277**	-0.009	0.184***
	(0.121)	(0.084)	(0.085)	(0.115)	(0.108)	(0.111)	(0.126)	(0.121)	(0.039)
Primary	0.005	0.120*	-0.006	0.064	-0.006	0.118*	0.148**	0.090	0.072***
	(0.076)	(0.062)	(0.061)	(0.065)	(0.075)	(0.065)	(0.060)	(0.069)	(0.027)
Secondary	0.156	0.144	0.234**	0.036	0.339***	0.083	0.292**	0.070	0.153***
	(0.120)	(0.100)	(0.103)	(0.101)	(0.124)	(0.098)	(0.116)	(0.117)	(0.042)
Primary Max	0.176	-0.023	0.093	0.014	0.228**	-0.078	-0.038	-0.064	0.016
	(0.116)	(0.118)	(0.102)	(0.117)	(0.113)	(0.114)	(0.109)	(0.102)	(0.043)
Secondary Max	0.143*	0.155**	0.139*	0.127	0.098	0.035	-0.041	0.146	0.105***
	(0.083)	(0.077)	(0.072)	(0.090)	(0.088)	(0.079)	(0.078)	(0.093)	(0.028)
Ajusted-R^2	0.584	0.596	0.618	0.592	0.642	0.462	0.440	0.472	0.535
Observations	503	504	504	504	504	502	501	503	4025

Notes: OLS estimator is used for all regressions. Standard errors in parentheses are clustered at the village level. *** p<0.01, ** p<0.05, * p<0.1. See notes to Table 1.1 for other details.

Table A9: Predictive Performances: food consumption only

	(1)	(2)
	$\hat{y}_{ik}$	$\hat{\mu}_{ik}$
Formula 1	-0.197***	0.105***
	(0.015)	(0.017)
Formula 2	-0.045***	0.039***
	(0.012)	(0.011)
Formula 3	-0.102***	0.070***
	(0.014)	(0.013)
Formula 4	-0.303***	0.122***
	(0.012)	(0.021)
Formula 5	-0.277***	0.168***
	(0.020)	(0.026)
Formula 6	-0.210***	0.084***
	(0.013)	(0.016)
Formula 7	-0.145***	0.061***
	(0.014)	(0.015)
F-statistics	178.50***	6.89***
Observations	4024	4024
Number of Households	503	503
Mean in *Formula* 8	12.489	0.283

Notes: OLS estimator is used for both regressions. Robust standard errors clustered at the village level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. See notes to Table 1.2 for other details.

Table A10: Predictive Performances: consumption per adult equivalent

	(1)	(2)
	$\hat{y}_{ik}$	$\hat{\mu}_{ik}$
Formula 1	-0.194***	0.112***
	(0.016)	(0.019)
Formula 2	-0.061***	0.040***
	(0.012)	(0.012)
Formula 3	-0.093***	0.079***
	(0.015)	(0.016)
Formula 4	-0.269***	0.112***
	(0.013)	(0.020)
Formula 5	-0.238***	0.149***
	(0.020)	(0.024)
Formula 6	-0.216***	0.094***
	(0.013)	(0.018)
Formula 7	-0.164***	0.085***
	(0.017)	(0.018)
F-statistics	149.26***	7.78***
Observations	4024	4024
Number of Households	503	503
Mean in *Formula* 8	12.402	0.303

Notes: OLS estimator is used for both regressions. Robust standard errors clustered at the village level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. See notes to Table 1.2 for other details.

Table A11: Predictive Performances: extended list of covariates in PMT

	(1)	(2)
	$\hat{y}_{ik}$	$\hat{\mu}_{ik}$
Formula 1	-0.200***	0.134***
	(0.018)	(0.017)
Formula 2	-0.074***	0.059***
	(0.014)	(0.012)
Formula 3	-0.085***	0.103***
	(0.018)	(0.015)
Formula 4	-0.273***	0.157***
	(0.018)	(0.021)
Formula 5	-0.231***	0.157***
	(0.021)	(0.022)
Formula 6	-0.221***	0.092***
	(0.012)	(0.015)
Formula 7	-0.170***	0.087***
	(0.015)	(0.015)
F-statistics	192.30***	10.75***
Observations	4024	4024
Number of Households	503	503
Mean in *Formula* 8	12.621	0.249

Notes: OLS estimator is used for both regressions. Robust standard errors clustered at the village level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. See notes to Table 1.2 for other details.

Table A12: Predictive Performances: no stepwise procedure in PMT

	(1)	(2)
	$\hat{y}_{ik}$	$\hat{\mu}_{ik}$
Formula 1	-0.190***	0.103***
	(0.014)	(0.017)
Formula 2	-0.055***	0.042***
	(0.011)	(0.011)
Formula 3	-0.090***	0.069***
	(0.014)	(0.014)
Formula 4	-0.275***	0.119***
	(0.013)	(0.020)
Formula 5	-0.236***	0.142***
	(0.018)	(0.021)
Formula 6	-0.214***	0.075***
	(0.009)	(0.015)
Formula 7	-0.155***	0.076***
	(0.015)	(0.016)
F-statistics	225.52***	8.53***
Observations	4024	4024
Number of Households	503	503
Mean in *Formula* 8	12.621	0.287

Notes: OLS estimator is used for both regressions. Robust standard errors clustered at the village level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. See notes to Table 1.2 for other details.

Table A13: Targeting Performances ($1.25 Poverty Line): food consumption only

	(1) TE_{ik}	(2) IE_{ik}	(3) EE_{ik}	(4) TE^2_{ik}	(5) IE^2_{ik}	(6) EE^2_{ik}
Formula 1	0.064***	0.217***	-0.190***	0.053***	0.092***	-0.014
	(0.023)	(0.029)	(0.035)	(0.017)	(0.023)	(0.022)
Formula 2	0.048**	0.140***	-0.106***	0.030***	0.052***	-0.005
	(0.021)	(0.024)	(0.034)	(0.011)	(0.014)	(0.019)
Formula 3	0.048**	0.150***	-0.122***	0.028***	0.051***	-0.011
	(0.021)	(0.026)	(0.033)	(0.010)	(0.014)	(0.013)
Formula 4	0.074***	0.303***	-0.307***	0.039***	0.105***	-0.071***
	(0.026)	(0.031)	(0.033)	(0.014)	(0.019)	(0.015)
Formula 5	0.095***	0.315***	-0.270***	0.079***	0.153***	-0.042*
	(0.027)	(0.032)	(0.039)	(0.021)	(0.029)	(0.022)
Formula 6	0.022	0.172***	-0.228***	0.011	0.048***	-0.052***
	(0.022)	(0.025)	(0.031)	(0.010)	(0.013)	(0.013)
Formula 7	0.036*	0.140***	-0.138***	0.013	0.042***	-0.036***
	(0.019)	(0.024)	(0.028)	(0.008)	(0.011)	(0.013)
F-statistics	3.22	16.81	16.41	3.78	6.39	4.44
Observations	4024	2512	1512	4024	2512	1512
Number of Households	503	314	190	503	314	190
Mean in *Formula* 8	0.278	0.185	0.434	0.054	0.037	0.082

Notes: LPM is used for regressions 1–3. OLS is used for regressions 4–6. Standard errors in parentheses are clustered at the village level. *** p<0.01, ** p<0.05, * p<0.1. See notes to Table 1.3 for other details.

Table A14: Targeting Performances ($1.25 Poverty Line): consumption per adult equivalent

	(1) TE_{ik}	(2) IE_{ik}	(3) EE_{ik}	(4) TE_{ik}^2	(5) IE_{ik}^2	(6) EE_{ik}^2
Formula 1	0.044**	0.188***	-0.133***	0.027	0.084***	-0.042***
	(0.019)	(0.025)	(0.028)	(0.017)	(0.027)	(0.016)
Formula 2	0.020	0.112***	-0.093***	-0.002	0.028***	-0.039***
	(0.014)	(0.020)	(0.021)	(0.008)	(0.009)	(0.012)
Formula 3	0.032*	0.097***	-0.049*	0.009	0.031***	-0.018
	(0.017)	(0.022)	(0.026)	(0.010)	(0.011)	(0.016)
Formula 4	0.048**	0.242***	-0.190***	0.011	0.079***	-0.072***
	(0.020)	(0.028)	(0.026)	(0.013)	(0.019)	(0.016)
Formula 5	0.054**	0.245***	-0.181***	0.028*	0.099***	-0.059***
	(0.021)	(0.029)	(0.027)	(0.015)	(0.022)	(0.016)
Formula 6	0.042**	0.188***	-0.137***	0.003	0.053***	-0.058***
	(0.019)	(0.025)	(0.026)	(0.011)	(0.014)	(0.015)
Formula 7	0.028	0.126***	-0.093***	0.002	0.038***	-0.044***
	(0.017)	(0.023)	(0.026)	(0.011)	(0.014)	(0.015)
F-statistics	1.10	14.13	10.96	1.71	3.91	4.05
Observations	4024	2216	1808	4024	2216	1808
Number of Households	503	277	226	503	277	226
Mean in *Formula* 8	0.247	0.139	0.481	0.059	0.029	0.124

Notes: LPM is used for regressions 1–3. OLS is used for regressions 4–6. Standard errors in parentheses are clustered at the village level. *** p<0.01, ** p<0.05, * p<0.1. See notes to Table 1.3 for other details.

Table A15: Targeting Performances ($1.25 Poverty Line): extended list of covariates in PMT

	(1) TE_{ik}	(2) IE_{ik}	(3) EE_{ik}	(4) TE^2_{ik}	(5) IE^2_{ik}	(6) EE^2_{ik}
Formula 1	0.064***	0.203***	-0.241***	0.033***	0.065***	-0.038*
	(0.024)	(0.028)	(0.037)	(0.012)	(0.015)	(0.020)
Formula 2	0.032	0.113***	-0.146***	0.028***	0.039***	0.005
	(0.021)	(0.022)	(0.044)	(0.008)	(0.011)	(0.013)
Formula 3	0.044**	0.119***	-0.120***	0.021**	0.028***	0.005
	(0.021)	(0.024)	(0.041)	(0.008)	(0.010)	(0.015)
Formula 4	0.066**	0.229***	-0.291***	0.046***	0.088***	-0.047**
	(0.026)	(0.029)	(0.040)	(0.014)	(0.018)	(0.018)
Formula 5	0.070***	0.232***	-0.285***	0.040***	0.083***	-0.054***
	(0.024)	(0.028)	(0.038)	(0.013)	(0.016)	(0.019)
Formula 6	0.024	0.133***	-0.215***	0.010	0.036***	-0.046**
	(0.018)	(0.022)	(0.035)	(0.010)	(0.011)	(0.019)
Formula 7	0.018	0.087***	-0.133***	0.011	0.031***	-0.034*
	(0.019)	(0.022)	(0.033)	(0.009)	(0.009)	(0.019)
F-statistics	1.53	12.18	14.68	3.84	6.05	1.87
Observations	4024	2760	1264	4024	2760	1264
Number of Households	503	345	158	503	345	158
Mean in *Formula* 8	0.221	0.122	0.437	0.040	0.021	0.081

Notes: LPM is used for regressions 1–3. OLS is used for regressions 4–6. Standard errors in parentheses are clustered at the village level. *** p<0.01, ** p<0.05, * p<0.1. See notes to Table 1.3 for other details.

Table A16: Targeting Performances ($1.25 Poverty Line): no stepwise procedure in PMT

	(1) TE_{ik}	(2) IE_{ik}	(3) EE_{ik}	(4) TE^2_{ik}	(5) IE^2_{ik}	(6) EE^2_{ik}
Formula 1	0.058**	0.186***	-0.222***	0.031**	0.061***	-0.035*
	(0.022)	(0.026)	(0.040)	(0.012)	(0.015)	(0.018)
Formula 2	0.042**	0.125***	-0.139***	0.025**	0.050***	-0.029**
	(0.017)	(0.020)	(0.033)	(0.010)	(0.013)	(0.012)
Formula 3	0.042**	0.125***	-0.139***	0.026**	0.046***	-0.017
	(0.019)	(0.023)	(0.034)	(0.011)	(0.015)	(0.017)
Formula 4	0.056**	0.235***	-0.335***	0.041***	0.090***	-0.066***
	(0.024)	(0.026)	(0.037)	(0.015)	(0.019)	(0.016)
Formula 5	0.062**	0.232***	-0.310***	0.039***	0.085***	-0.063***
	(0.024)	(0.026)	(0.038)	(0.014)	(0.018)	(0.016)
Formula 6	0.024	0.145***	-0.241***	0.005	0.042***	-0.075***
	(0.019)	(0.022)	(0.034)	(0.010)	(0.011)	(0.022)
Formula 7	0.032*	0.099***	-0.114***	0.008	0.035***	-0.051**
	(0.018)	(0.021)	(0.035)	(0.010)	(0.009)	(0.022)
F-statistics	1.67	13.48	16.08	1.96	5.19	4.15
Observations	4024	2760	1264	4024	2760	1264
Number of Households	503	345	158	503	345	158
Mean in *Formula* 8	0.256	0.148	0.494	0.060	0.032	0.120

Notes: LPM is used for regressions 1–3. OLS is used for regressions 4–6. Standard errors in parentheses are clustered at the village level. *** p<0.01, ** p<0.05, * p<0.1. See notes to Table 1.3 for other details.

Table A17: Targeting Performances (30% Poverty Threshold): food consumption only

	(1) TE_{ik}	(2) IE_{ik}	(3) EE_{ik}	(4) TE_{ik}^2	(5) IE_{ik}^2	(6) EE_{ik}^2
Formula 1	0.036**	0.026	0.060*	0.019**	0.010	0.041
	(0.016)	(0.020)	(0.033)	(0.009)	(0.008)	(0.025)
Formula 2	-0.008	-0.006	-0.013	0.008	0.008	0.006
	(0.014)	(0.014)	(0.033)	(0.009)	(0.007)	(0.026)
Formula 3	0.028*	0.020	0.046	0.004	-0.004	0.022
	(0.017)	(0.021)	(0.037)	(0.009)	(0.007)	(0.025)
Formula 4	0.020	0.014	0.033	0.011	0.002	0.032
	(0.015)	(0.018)	(0.030)	(0.009)	(0.005)	(0.026)
Formula 5	0.036**	0.026	0.060	0.015	0.005	0.036
	(0.017)	(0.021)	(0.038)	(0.009)	(0.008)	(0.025)
Formula 6	0.032*	0.023	0.053	0.008	0.004	0.015
	(0.018)	(0.021)	(0.040)	(0.008)	(0.008)	(0.018)
Formula 7	0.030*	0.026	0.040	0.005	0.011	-0.008
	(0.017)	(0.020)	(0.043)	(0.008)	(0.007)	(0.022)
F-statistics	1.47	0.54	1.00	1.34	1.39	1.14
Observations	4024	2816	1208	4024	2816	1208
Number of Households	503	352	151	503	352	151
Mean in *Formula* 8	0.262	0.187	0.437	0.057	0.028	0.125

Notes: LPM is used for regressions 1–3. OLS is used for regressions 4–6. Standard errors in parentheses are clustered at the village level. *** p<0.01, ** p<0.05, * p<0.1. See notes to Table 1.4 for other details.

Table A18: Targeting Performances (30% Poverty Threshold): consumption per adult equivalent

	(1) TE_{ik}	(2) IE_{ik}	(3) EE_{ik}	(4) TE^2_{ik}	(5) IE^2_{ik}	(6) EE^2_{ik}
Formula 1	0.040***	0.028*	0.066**	0.021**	0.005	0.060*
	(0.013)	(0.016)	(0.032)	(0.010)	(0.005)	(0.031)
Formula 2	0.024**	0.017	0.040	0.021*	0.007	0.053
	(0.012)	(0.015)	(0.025)	(0.011)	(0.005)	(0.034)
Formula 3	0.036***	0.026	0.060*	0.025**	0.003	0.074**
	(0.014)	(0.016)	(0.034)	(0.011)	(0.005)	(0.034)
Formula 4	0.036**	0.026	0.060*	0.022**	0.005	0.062*
	(0.015)	(0.018)	(0.033)	(0.010)	(0.005)	(0.033)
Formula 5	0.036**	0.026	0.060*	0.025**	0.009	0.062*
	(0.015)	(0.018)	(0.033)	(0.011)	(0.007)	(0.033)
Formula 6	0.034**	0.026	0.053	0.016**	0.008	0.035*
	(0.016)	(0.019)	(0.035)	(0.007)	(0.005)	(0.020)
Formula 7	0.024	0.017	0.040	0.014	0.006	0.031
	(0.015)	(0.018)	(0.038)	(0.010)	(0.005)	(0.032)
F-statistics	1.49	0.49	0.80	1.44	0.85	1.17
Observations	4024	2816	1208	4024	2816	1208
Number of Households	503	352	151	503	352	151
Mean in *Formula* 8	0.227	0.162	0.377	0.061	0.011	0.178

Notes: LPM is used for regressions 1–3. OLS is used for regressions 4–6. Standard errors in parentheses are clustered at the village level. *** p<0.01, ** p<0.05, * p<0.1. See notes to Table 1.4 for other details.

Table A19: Targeting Performances (30% Poverty Threshold): extended list of covariates in PMT

	(1) TE_{ik}	(2) IE_{ik}	(3) EE_{ik}	(4) TE^2_{ik}	(5) IE^2_{ik}	(6) EE^2_{ik}
Formula 1	0.020	0.017	0.026	0.031***	0.009	0.083**
	(0.016)	(0.020)	(0.038)	(0.012)	(0.009)	(0.033)
Formula 2	0.016	0.011	0.026	0.022**	0.012	0.046**
	(0.016)	(0.018)	(0.038)	(0.009)	(0.010)	(0.021)
Formula 3	0.036*	0.026	0.060	0.021**	0.007	0.054**
	(0.019)	(0.023)	(0.041)	(0.009)	(0.009)	(0.023)
Formula 4	0.020	0.014	0.033	0.029**	0.011	0.071**
	(0.017)	(0.020)	(0.040)	(0.012)	(0.010)	(0.032)
Formula 5	0.028	0.020	0.046	0.016*	0.002	0.050**
	(0.018)	(0.020)	(0.047)	(0.009)	(0.009)	(0.024)
Formula 6	0.024	0.017	0.040	0.015*	0.012	0.022
	(0.017)	(0.019)	(0.040)	(0.008)	(0.010)	(0.014)
Formula 7	0.020	0.014	0.033	0.009	0.009	0.009
	(0.018)	(0.022)	(0.041)	(0.007)	(0.009)	(0.010)
F-statistics	0.63	0.22	0.36	1.32	0.84	1.26
Observations	4024	2816	1208	4024	2816	1208
Number of Households	503	352	151	503	352	151
Mean in *Formula* 8	0.219	0.156	0.364	0.036	0.026	0.061

Notes: LPM is used for regressions 1–3. OLS is used for regressions 4–6. Standard errors in parentheses are clustered at the village level. *** p<0.01, ** p<0.05, * p<0.1. See notes to Table 1.4 for other details.

Table A20: Targeting Performances (30% Poverty Threshold): no stepwise procedure in PMT

	(1)	(2)	(3)	(4)	(5)	(6)
	TE_{ik}	IE_{ik}	EE_{ik}	TE_{ik}^2	IE_{ik}^2	EE_{ik}^2
Formula 1	0.028**	0.020	0.046	0.010	0.009	0.014
	(0.014)	(0.016)	(0.033)	(0.008)	(0.009)	(0.016)
Formula 2	-0.000	-0.000	-0.000	0.004	0.004	0.005
	(0.012)	(0.014)	(0.028)	(0.006)	(0.008)	(0.013)
Formula 3	0.024	0.017	0.040	0.002	-0.003	0.014
	(0.015)	(0.018)	(0.034)	(0.008)	(0.010)	(0.015)
Formula 4	0.012	0.009	0.020	-0.003	-0.008	0.009
	(0.014)	(0.016)	(0.033)	(0.004)	(0.005)	(0.009)
Formula 5	0.016	0.011	0.026	0.005	0.002	0.013
	(0.014)	(0.016)	(0.033)	(0.007)	(0.008)	(0.015)
Formula 6	0.020	0.014	0.033	0.003	-0.000	0.011
	(0.013)	(0.014)	(0.030)	(0.005)	(0.004)	(0.013)
Formula 7	0.024	0.017	0.040	-0.001	0.006	-0.018
	(0.016)	(0.018)	(0.039)	(0.008)	(0.007)	(0.022)
F-statistics	1.07	0.40	1.02	0.51	1.25	0.45
Observations	4024	2816	1208	4024	2816	1208
Number of Households	503	352	151	503	352	151
Mean in *Formula* 8	0.254	0.182	0.424	0.064	0.044	0.111

Notes: LPM is used for regressions 1–3. OLS is used for regressions 4–6. Standard errors in parentheses are clustered at the village level. *** p<0.01, ** p<0.05, * p<0.1. See notes to Table 1.4 for other details.

Appendix B

Appendix to Chapter 2

Table B1: IV estimates

	Migration (excl. returns)			Migration (incl. returns)		
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	0.051**	0.048**	0.048**	0.063**	0.058**	0.058**
	(0.022)	(0.021)	(0.021)	(0.026)	(0.025)	(0.025)
Extended controls		✓	✓	No	✓	✓
Island FE			✓			✓
Control mean	0.078	0.078	0.078	0.128	0.128	0.128
Observations	2181	2181	2181	2181	2181	2181

Notes: This table reports LATE estimates of the program. Random assignment is used as an IV for actually treated households (according to survey data). See notes to Table 2.4 for other details. Standard errors in parentheses are clustered at the level of the treatment (household). *** p<0.01, ** p<0.05, * p<0.1.

Table B2: Differential attrition test

	Control		Treatment			
	Mean	SD	Mean	SD	Diff	p-value
Attrition rate	0.044	0.206	0.037	0.189	0.007	0.39
Observations	900	900	1372	1372	2272	2272

Notes: This table displays the difference in mean attrition between treatment and control groups.

Table B3: Attrition reasons

	Control		Treatment			
	Mean	SD	Mean	SD	Diff	p-value
Attrition reason						
Duplicate household	0.002	0.047	0.002	0.047	0.000	0.99
Refusal	0.007	0.081	0.004	0.066	0.002	0.46
Absent	0.009	0.094	0.009	0.093	0.000	0.97
Dissolved household	0.020	0.140	0.019	0.136	0.001	0.86
Too sick	0.001	0.033	0.001	0.027	0.000	0.76
Other	0.006	0.074	0.002	0.047	0.003	0.19
Observations	900	900	1372	1372	2272	2272

Notes: This table displays difference in mean attrition rates between treatment and control groups by attrition reasons.

Table B4: Indirect treatment effects

	Migration (excl. returns)			Migration (incl. returns)		
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	0.033	0.033	0.033*	0.033	0.033	0.032
	(0.022)	(0.021)	(0.017)	(0.028)	(0.027)	(0.021)
40% villages (β_2)	-0.001	-0.002	-0.003	-0.006	-0.006	-0.008
	(0.023)	(0.019)	(0.018)	(0.028)	(0.025)	(0.023)
Treatment x 40% villages (β_3)	-0.006	-0.009	-0.009	0.008	0.003	0.003
	(0.026)	(0.025)	(0.024)	(0.036)	(0.035)	(0.029)
$\beta_2 + \beta_3$	-0.007	-0.011	-0.012	0.002	-0.003	-0.005
	(0.024)	(0.018)	(0.016)	(0.031)	(0.021)	(0.020)
Extended controls		✓	✓		✓	✓
Island FE			✓			✓
Control mean (in 20% villages)	0.079	0.079	0.079	0.131	0.131	0.131
Observations	2181	2181	2181	2181	2181	2181

Notes: This table reports LPM estimates of indirect treatment effects using equation (2.18). See notes to Table 2.4 for other details. Standard errors in parentheses are clustered at the level of the treatment (village). *** p<0.01, ** p<0.05, * p<0.1.

Table B5: Liquidity channel

	Migration (excl. returns)					
	(1)	(2)	(3)	(4)	(5)	(6)
Savings (dummy):						
Treatment (β_1)	0.062***	0.063***	0.063***			
	(0.020)	(0.020)	(0.020)			
High savings	0.025	0.027	0.032			
	(0.019)	(0.019)	(0.171)			
Treatment × High savings (β_3)	-0.051**	-0.053**	-0.055**			
	(0.025)	(0.025)	(0.025)			
$\beta_1 + \beta_3$	0.011	0.010	0.008			
	(0.015)	(0.015)	(0.015)			
Control mean (low savings)	0.064	0.064	0.064			
Savings (continuous):						
Treatment (β_1)				0.038***	0.039***	0.038***
				(0.013)	(0.013)	(0.013)
Savings				0.001	0.001	0.001
				(0.001)	(0.001)	(0.007)
Treatment × Savings (β_3)				-0.002*	-0.002*	-0.002*
				(0.001)	(0.001)	(0.001)
Control mean (savings = 0)				0.073	0.073	0.073
Extended controls	✓	✓	✓	✓	✓	✓
Island FE		✓	✓		✓	✓
Savings × Controls			✓			✓
Observations	2181	2181	2181	2181	2181	2181

Notes: This table reports LPM estimates of conditional effects using equation (2.19). See notes to Table 2.4 for other details. Standard errors in parentheses are clustered at the level of the treatment (household). *** p<0.01, ** p<0.05, * p<0.1.

Table B6: Risk aversion channel

	Migration (excluding returns)					
	(1)	(2)	(3)	(4)	(5)	(6)
Risk aversion (dummy):						
Treatment (β_1)	-0.003	-0.002	-0.009			
	(0.042)	(0.042)	(0.042)			
High risk aversion	-0.105**	-0.104**	-0.426			
	(0.046)	(0.047)	(0.388)			
Treatment $\times$ High risk aversion (β_3)	0.114*	0.113*	0.127**			
	(0.062)	(0.063)	(0.062)			
$\beta_1 + \beta_3$	0.111**	0.111**	0.118**			
	(0.044)	(0.045)	(0.046)			
Control mean (low risk aversion)	0.140	0.140	0.140			
Risk aversion (continuous):						
Treatment (β_1)				0.048	0.048	0.046
				(0.030)	(0.030)	(0.033)
Risk aversion				-0.043*	-0.043	-0.406*
				(0.026)	(0.026)	(0.238)
Treatment $\times$ Risk aversion (β_3)				0.041	0.041	0.039
				(0.035)	(0.036)	(0.035)
Control mean (risk aversion $=0$)				0.107	0.107	0.107
Extended controls	✓	✓	✓	✓	✓	✓
Island FE		✓	✓		✓	✓
Risk aversion $\times$ Controls			✓			✓
Observations	476	476	476	476	476	476

Notes: This table reports LPM estimates of conditional effects using equation (2.19). See notes to Table 2.4 for other details. Standard errors in parentheses are clustered at the level of the treatment (household). *** p<0.01, ** p<0.05, * p<0.1.

B1 Mathematical Appendix

Proof of Proposition 1

The lifetime utilities of financing migration in t_1 (Case 1), financing migration in t_2 (Case 2), or not migrating at all (Case 3) are:

$$\begin{aligned}
U^{Case1} &= u(s_0 - c + 2w_d) \\
U^{Case2} &= u(s_0 - c + w_o + w_d) \\
U^{Case3} &= u(s_0 + 2w_o)
\end{aligned}$$

Case 1: migration in t_1 Financing migration in t_1 is only feasible if the initial level of savings s_0 is large enough to finance the upfront cost of migration ($s_0 \geq c$). If feasible, a household member migrates in t_1 if:

$$\begin{cases}
U^{Case1} > U^{Case2} \Leftrightarrow w_d - w_o > 0 \\
U^{Case1} > U^{Case3} \Leftrightarrow 2(w_d - w_o) > c
\end{cases}$$

Case 2: migration in t_2 Financing migration in t_2 is only feasible if the household can save enough in t_1 to pay the upfront cost of migration in t_2 ($s_0 + w_o > c$). If $s_0 \geq c$, migrating in t_1 is always preferable to migrating in t_2. If $c - w_o \leq s_0 < c$, a household member migrates in t_2 if:

$$U^{Case2} > U^{Case3} \Leftrightarrow w_d - w_o > c$$

Proof of Proposition 2

The lifetime utilities of financing migration in t_1 (Case 1), financing migration in t_2 (Case 2), or not migrating at all (Case 3) are:

$$\begin{aligned}
U^{Case1} &= u(s_0 - c + 2w_d + \tau) \\
U^{Case2} &= u(s_0 - c + w_o + w_d + \tau) \\
U^{Case3} &= u(s_0 + 2w_o + \tau)
\end{aligned}$$

Case 1: migration in t_1 Financing migration in t_1 is only feasible if the initial level of savings s_0 is large enough to finance the upfront cost of migration ($s_0 \geq c$). This budget constraint is not affected by the unconditional cash transfer, as it is received after the decision to migrate in t_1. If feasible, a household member migrates in t_1 if:

$$\begin{cases} U^{Case1} > U^{Case2} \Leftrightarrow w_d - w_o > 0 \\ U^{Case1} > U^{Case3} \Leftrightarrow 2(w_d - w_o) > c \end{cases}$$

Case 2: migration in t_2 Financing migration in t_2 is only feasible if the household can save enough in t_1 to pay the upfront cost of migration in t_2 ($s_0 + w_o + \tau > c$). If $s_0 \geq c$, migrating in t_1 is always preferable to migrating in t_2. If $c - w_o - \tau \leq s_0 < c$, a household member migrates in t_2 if:

$$U^{Case2} > U^{Case3} \Leftrightarrow w_d - w_o > c$$

Proof of Proposition 3

The lifetime utilities of financing migration in t_1 (Case 1), financing migration in t_2 (Case 2), or not migrating at all (Case 3) are:

$$\begin{aligned} U^{Case1} &= u(s_0 - c + 2w_d) \\ U^{Case2} &= u(s_0 - c + w_o + w_d + \tau) \\ U^{Case3} &= u(s_0 + 2w_o + \tau) \end{aligned}$$

Case 1: migration in t_1 Financing migration in t_1 is only feasible if the initial level of savings s_0 is large enough to finance the upfront cost of migration ($s_0 \geq c$). This budget constraint is not affected by the conditional cash transfer. If feasible, a household member migrates in t_1 if:

$$\begin{cases} U^{Case1} > U^{Case2} \Leftrightarrow w_d - w_o > \tau \\ U^{Case1} > U^{Case3} \Leftrightarrow 2(w_d - w_o) > c + \tau \end{cases}$$

If $\tau > c$, then the first condition is more stringent than the second one ($U^{Case2} > U^{Case3}$). If $\tau < c$, then the second condition is more stringent than the first one ($U^{Case2} < U^{Case3}$)

Case 2: migration in t_2 Financing migration in t_2 is only feasible if the household can save enough in t_1 to pay the upfront cost of migration in t_2 ($s_0 + w_o + \tau > c$). If $c - w_o \leq s_0 < c$, migration cannot be financed in t_1. In this case, a household member migrates in t_2 if:

$$U^{Case2} > U^{Case3} \Leftrightarrow w_d - w_o > c$$

If $s_0 \geq c$, migration can be financed in both t_1 and t_2. In this case, a household member migrates in t_2 if:

$$\begin{cases} U^{Case2} \geq U^{Case1} \Leftrightarrow w_d - w_o \leq \tau \\ U^{Case2} > U^{Case3} \Leftrightarrow w_d - w_o > c \end{cases}$$

Proof of Proposition 4

If $s_0 \geq c$, the household does not need to borrow to finance migration in t_1. Therefore, borrowing only occurs if borrowing is necessary and sufficient to finance migration in t_1, which occurs if $c - B \leq s_0 < c$. If the household borrow, it will borrow the amount $c - s_0$, which is the minimum loan that allows financing migration in t_1. The household will not borrow more as borrowing is costly ($r \geq 0$ and as consumption smoothing is irrelevant following the assumption that households are maximizing lifetime wealth.

The lifetime utilities of financing migration in t_1 with savings (Case 1A), financing migration in t_1 with a loan (Case 1B), financing migration in t_2 (Case 2), or not migrating at all (Case 3) are:

$$\begin{aligned} U^{Case1A} &= u(s_0 - c + 2w_d) \\ U^{Case1B} &= u(s_0 - c + 2w_d - (c - s_0)r) = \\ U^{Case2} &= u(s_0 - c + w_o + w_d) \\ U^{Case3} &= u(s_0 + 2w_o) \end{aligned}$$

If borrowing is necessary and sufficient to finance migration in t_1 ($c - B \leq s_0 < c$), borrowing is optimal if:

$$\begin{cases} U^{Case1B} > U^{Case2} \Leftrightarrow w_d - w_o > r(c - s_0) \\ U^{Case1B} > U^{Case3} \Leftrightarrow 2(w_d - w_o) > c + r(c - s_0) \end{cases}$$

Proof of Proposition 5

Expected utilities The lifetime expected utility of a household attempting to migrate in t_1 is:

$$
\begin{aligned}
U^{Case1} &= p[u(s_0 - c + 2w_d)] + (1-p)[u(s_0 - c + 2w_o)] \\
&= u(s_o - c + 2pw_d + 2(1-p)w_o - \pi_1)
\end{aligned}
$$

where π_1 is the risk premium associated with migrating in t_1. The lifetime expected utility of a household attempting to migrate in t_2 is:

$$
\begin{aligned}
U^{Case2} &= p[u(s_0 - c + w_o + w_d)] + (1-p)[u(s_0 - c + 2w_o)] \\
&= u(s_o - c + w_o + pw_d + (1-p)w_o - \pi_2)
\end{aligned}
$$

where π_2 is the risk premium associated with migrating in t_2. The lifetime utility of a household who does not finance migration is:

$$U^{Case3} = u(s_0 + 2w_o)$$

Case 1: migration in t_1 Financing migration in t_1 is only feasible if the initial level of savings s_0 is large enough to finance the upfront cost of migration ($s_0 \geq c$). It is straightforward that $U^{Case1} > U^{Case2}$: the probability of success and the bad outcome are the same for these two lotteries, while the good outcome is better in Case 1 (given the assumption $u' > 0$). Therefore, if feasible ($s_0 \geq c$), a household member migrates in t_1 if:

$$U^{Case1} > U^{Case3} \Leftrightarrow 2p(wd - wo) > c + \pi_1$$

Case 2: migration in t_2 Financing migration in t_2 is only feasible if the initial level of savings s_0 and the wage at origin are large enough to finance the upfront cost of migration in t_2 ($s_0 + w_o \geq c$). If $s_0 \geq c$, migration in t_2 is never optimal as $U^{Case1} > U^{Case2}$. If $c - w_o \leq s_0 < c$, a household member migrates in t_2 if:

$$U^{Case2} > U^{Case3} \Leftrightarrow p(wd - wo) > c + \pi_2$$

Proof of Proposition 6

Expected utilities The lifetime expected utility of a household attempting to migrate at the beginning of t_1 and benefiting from an unconditional cash transfer at the end of t_1 is:

$$\begin{aligned} U^{Case1} &= p[u(s_0 - c + 2w_d + \tau)] + (1-p)[u(s_0 - c + 2w_o + \tau)] \\ &= u(s_o - c + 2pw_d + 2(1-p)w_o + \tau - \pi_1') \end{aligned}$$

where π_1' is the risk premium associated with migrating in t_1. The lifetime expected utility of a household benefiting from an unconditional cash transfer at the end of t_1 and attempting to migrate at the beginning of t_2 is:

$$\begin{aligned} U^{Case2} &= p[u(s_0 - c + w_o + w_d + \tau)] + (1-p)[u(s_0 - c + 2w_o + \tau)] \\ &= u(s_o - c + w_o + pw_d + (1-p)w_o + \tau - \pi_2') \end{aligned}$$

where π_2' is the risk premium associated with migrating in t_2. The lifetime utility of a household benefiting from an unconditional cash transfer at the end of t_1 and not attempting to migrate is:

$$U^{Case3} = u(s_0 + 2w_o + \tau)$$

Case 1: migration in t_1 Financing migration in t_1 is only feasible if the initial level of savings s_0 is large enough to finance the upfront cost of migration ($s_0 \geq c$). It is straightforward that $U^{Case1} > U^{Case2}$: the probability of success and the bad outcome are the same for these two lotteries, while the good outcome is better in Case 1 (given the assumption $u' > 0$). Therefore, if feasible ($s_0 \geq c$), a household member migrates in t_1 if:

$$U^{Case1} > U^{Case3} \Leftrightarrow 2p(wd - wo) > c + \pi_1'$$

Case 2: migration in t_2 Financing migration in t_2 is only feasible if the sum of the initial level of savings s_0, the wage at origin w_o and the cash tranfer τ is large enough to finance the upfront cost of migration in t_2 ($s_0 + w_o + \tau \geq c$). If $s_0 \geq c$, migration in t_2 is never optimal because $U^{Case1} > U^{Case2}$. If $c - w_o - \tau \leq s_0 < c$, a household member migrates in t_2 if:

$$U^{Case2} > U^{Case3} \Leftrightarrow p(wd - wo) > c + \pi_2'$$

The budget constraint in t_2 is eased by the cash transfer (liquidity channel). Furthermore, if the utility function is characterized by decreasing absolute risk aversion (DARA), $\pi_1' < \pi_1$ and $\pi_2' < \pi_2$, implying that households are less risk averse thanks to the transfer and more willing to accept the risk associated with migration. By contrast, if the utility function is characterized by increasing absolute risk aversion (IARA), $\pi_1' > \pi_1$ and $\pi_2' > \pi_2$, implying that households become more risk averse with the cash transfer and less willing to accept the risk of migrating. If the household is risk neutral, $\pi_1' = \pi_1$ and the only effect of the cash transfer is through the liquidity channel.

B2 Sub-analysis outlined in the PAP

Remittances For each household member who was reported as having migrated to Mayotte between the baseline and follow-up survey, and still in Mayotte at follow-up, we collected data on remittances sent to the household of origin. In Table B7, we present the impact of the program on two main variables: (i) a dummy indicating whether the migrant sent remittances to his or her household of origin (using a 12 months recall period); (ii) the total amount of remittances sent. While the program seem to have had a positive effect on remittances, coefficients are small in absolute terms and non-significant. The latter could be explained by the fact that a minority of migrants started to remit (migration usually takes time to become profitable. Alternatively, it may also be due to a crowding out-effect of the program on remittances, though we should observe a negative coefficient if this mechanism was widespread.

Migration reasons When respondents reported a migrant, we further inquired about the reason for migrating. The impact of the program by migration reason is presented in Table B8. Respondents declared three main reasons for migrating: economic reasons, health reasons, and family reasons. The overall positive effect on migration we observe seems to be especially driven by individuals migrating for health reasons, followed by family migration, and economic migration. However, Table B9 shows that economic migrants are not the only one to send remittances to their household of origin. People migrating for health

Table B7: Treatment effects on remittances

	Remittances (dummy)			Remittances (amount sent)		
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	0.010	0.008	0.008	0.120*	0.106	0.107
	(0.006)	(0.006)	(0.006)	(0.069)	(0.069)	(0.069)
Extended controls		✓	✓		✓	✓
Island FE			✓			✓
Control mean	0.016	0.016	0.016	0.175	0.175	0.175
Observations	2163	2163	2163	2163	2163	2163

Notes: The dependent variable in columns 1 to 3 is a dummy equal to one if the migrant sent remittances to his or her household of origin. The dependent variable in columns 4 to 6 equals the amount of the remittances. An inverse hyperbolic sine (IHS) transformation has been applied to the amount of the remittances. We do not have information on remittances sent by return migrants during their time in Mayotte. All estimates control for unbalanced covariates. Standard errors in parentheses are clustered at the level of the treatment (household). *** p<0.01, ** p<0.05, * p<0.1.

and family reasons also remit. This suggests that the different migration reasons are not mutually exclusive, even though our survey instruments inquired respondents to select only one type of migration. In addition, people migrating for economic opportunities might state health or family motives because they believe these motives could be seen as more legitimate.

Table B8: Treatment effects by migration reasons

	Migration (excl. returns)					
	(1) Economic	(2) Health	(3) Family	(4) Studies	(5) Tourism	(6) Other
Treatment	0.007	0.020***	0.011	-0.001	-0.004	-0.001
	(0.007)	(0.007)	(0.007)	(0.004)	(0.003)	(0.002)
Extended controls	✓	✓	✓	✓	✓	✓
Island FE	✓	✓	✓	✓	✓	✓
Control mean	0.023	0.025	0.026	0.007	0.005	0.002
Observations	2163	2163	2163	2163	2163	2163
	Migration (incl. returns)					
	(1) Economic	(2) Health	(3) Family	(4) Studies	(5) Tourism	(6) Other
Treatment	0.002	0.024**	0.022**	-0.002	0.001	-0.001
	(0.009)	(0.010)	(0.008)	(0.004)	(0.003)	(0.002)
Extended controls	✓	✓	✓	✓	✓	✓
Island FE	✓	✓	✓	✓	v	✓
Control mean	0.040	0.047	0.030	0.008	0.005	0.002
Observations	2163	2163	2163	2163	2163	2163

Notes: All estimates control for unbalanced covariates. Standard errors in parentheses are clustered at the level of the treatment (household). *** p<0.01, ** p<0.05, * p<0.1.

Table B9: Summary statistics on remittances sent by migration reason

	Remittances			N
	Dummy	Amount sent		
		(All)	(if D=1)	
Economic	0.44	5.00	11.49	62
Health	0.13	1.24	9.83	79
Family	0.24	2.79	11.49	70
Studies	0.14	0.87	6.10	14
Tourism	0.17	1.92	11.51	6
Other	0.00			4
Total	0.23	2.58	11.16	208

Notes: An inverse hyperbolic sine (IHS) transformation has been applied to all remittances amount. The sample is restricted to Mayotte migrants. We do not have information on remittances sent by return migrants during their time in Mayotte.

Heterogeneous effects Finally, we examine heterogeneity in the effect by baseline characteristics. In Table B10, we analyze whether the effect varies with (i) the willingness to migrate, (ii) the number of rounds of CFW received, (iii) the number of working-age adults in the household, (iv) the total consumption per adult equivalent, and (v) the schooling of the household head. Because of the financial constraints highlighted above, we expect the effect to increase with household willingness to migrate and the number of CFW received, and decrease with consumption. The mediating effect of the number of working age adults is more ambiguous. The more working-age adults in the household, the less binding the labor requirement of CFW opportunities. However, the marginal effect of cash received may be smaller in larger households.

Table B10: Heteregeneous Effects

	Migration (excluding returns)				
	(1)	(2)	(3)	(4)	(5)
Treatment	0.022*	-0.035	0.030	0.040	0.028*
	(0.013)	(0.042)	(0.024)	(0.091)	(0.015)
Treatment x Willing to migrate	0.025				
	(0.033)				
Treatment x CFW rounds (N)		0.013			
		(0.009)			
Treatment x Working age adults (N)			-0.001		
			(0.008)		
Treatment x Consumption				-0.002	
				(0.012)	
Treatment x Schooling					0.000
					(0.013)
Extended controls	✓	✓	✓	✓	✓
Island FE	✓	✓	✓	✓	✓
Control mean	0.078	0.078	0.078	0.078	0.078
Observations	2181	2181	2181	2181	2181

Notes: Each column refers to a different LPM estimate using equation (2.19). Estimates control for unbalanced covariates. Standard errors in parentheses are clustered at the level of the treatment (household). *** p<0.01, ** p<0.05, * p<0.1.

The sign of the interaction terms are in line with expectations, but not significant at conventional significance levels. It seems that the effect is stronger for households willing to migrate and receiving more CFW rounds, and lower for more wealthy households. The number of working-age adults does not seem to condition the effect. We explored the presence of potential non-linearities using a quadratic interaction term but the results show no effects.

We investigate heterogeneous effects more comprehensively by implementing the endogenous stratification method, a three-step procedure which allows to assess how different groups are affected by the treatment. First, using control households, we regress the outcome variable (migration to Mayotte) on the baseline characteristics highlighted in Table 3.3. We then use the fitted coefficients to predict migration in the absence of treatment for both the treatment and control groups. Finally, we split the households into different groups on the basis of their predicted migration values and estimate treatment effects across these groups.[1] The results are presented in Table B11. We see that

Table B11: Endogenous stratification

	Household		Individual			
	Migration (excl. returns)	Migration (incl. returns)	Migration (excl. returns)		Migration (incl. returns)	
	(1)	(2)	(3)	(4)	(5)	(6)
Low predicted migration						
Treatment	0.025	-0.009	0.001	0.006	0.002	0.005
SE	(0.015)	(0.021)	(0.003)	(0.003)	(0.003)	(0.004)
Control mean	0.018	0.070	0.011	0.007	0.015	0.011
Medium predicted migration						
Treatment	0.016	0.067		-0.002		0.004
SE	(0.022)	(0.029)		(0.005)		(0.005)
Control mean	0.067	0.090		0.014		0.020
High predicted migration						
Treatment	0.044	0.043	0.011	0.014	0.016	0.018
SE	(0.032)	(0.036)	(0.004)	(0.005)	(0.005)	(0.007)
Control mean	0.150	0.229	0.019	0.023	0.032	0.040
Number of groups	3	3	2	3	2	3
Predictors:						
Extended controls	✓	✓				
Island FE	✓	✓	✓	✓	✓	✓
Individual controls			✓	✓	✓	✓
Observations	2181	2181	14288	14288	14288	14288

Notes: Using the leave-one-out estimation procedure. Standard errors in parentheses are bootstrapped (1,000 repetitions). *** p<0.01, ** p<0.05, * p<0.1.

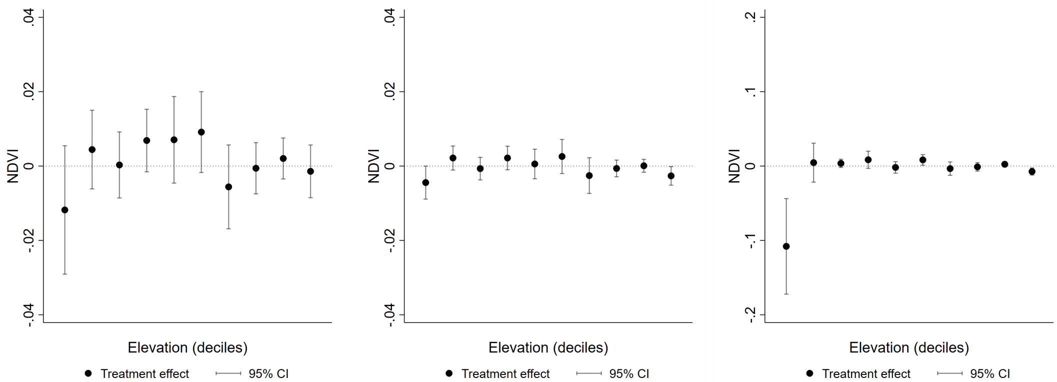
[1] The fitted model is estimated excluding the observation itself to avoid bias (Abadie et al., 2018). We used the *estrat* Stata command with the leave-one-out option which automates the procedure.

Appendix C

Appendix to Chapter 3

Figure C1: Treatment effects by elevation deciles

(a) Treatment (dummy) (b) Treatment (intensity) (c) Treatment (density)



Notes:

Table C1: Satellite and survey-based cultivated area

	(1) Cult. Area (LSMS-ISA)	(2) Cult. Area (LSMS-ISA)	(3) Cult. Area (LSMS-ISA)
Cult. Area (MODIS)	0.108***	-0.018	-0.015
	(0.022)	(0.017)	(0.018)
Woredas FE		✓	✓
Time FE			✓
Observations	479	476	476
R-squared	0.06	0.92	0.92

Notes: In columns (1)-(3), the dependant variable corresponds to the overall cultivated area in woreda w at time t (with $t = 2013 \mid 2015$), derived from the Ethiopian 2013 and 2015 LSMS-ISA surveys. An inverse hyperbolic sine transformation has been applied to all variables. Standard errors in parentheses are clustered at woreda level. *** p<0.01, ** p<0.05, * p<0.1.

Table C2: Determinants of the treatment

	(1) Treatment
Rainfall	0.050***
	(0.015)
$Rainfall^2$	-0.000***
	(0.000)
Temperature	0.681***
	(0.175)
$Temperature^2$	-0.009***
	(0.002)
Elevation	-0.001
	(0.000)
Slope	0.278***
	(0.055)
Start growing season	0.705***
	(0.223)
End growing season	0.774***
	(0.232)
Total area	-0.210
	(0.805)
Cultivated area (% total area)	0.461
	(1.094)
Population density	0.038***
	(0.006)
Night time lights	-1.807***
	(0.262)
NDVI	-10.295***
	(2.441)
Observations	633
R-squared	0.45

Notes: The outcome variable is a dummy equal to one if the woreda received the treatment. A logit estimator is used. Standard errors in parentheses are clustered at the woreda level. *** p<0.01, ** p<0.05, * p<0.1.

Table C3: Treatment and control Woredas by regions

	All woredas			Common support		
	Control	Treatment	Total	Control	Treatment	Total
Addis Ababa	6	0	6	1	0	1
Afar	0	29	29	0	27	27
Amhara	65	63	128	39	57	96
Beneshangul Gumu	17	0	17	6	0	6
Dire Dawa	0	1	1	0	1	1
Gambela	11	0	11	9	0	9
Hareri	0	1	1	0	1	1
Oromia	174	73	247	94	71	165
SNNPR	50	78	128	40	69	109
Somali	24	7	31	24	7	31
Tigray	3	31	34	1	28	29
Total	350	283	633	214	261	475

Table C4: Main results with standardized outcome

	Highlands			Lowlands		
	NDVI	NDVI	NDVI	NDVI	NDVI	NDVI
	(1)	(2)	(3)	(4)	(5)	(6)
Post × Treatment (dummy)	0.034			-0.033		
	(0.030)			(0.032)		
Post × Treatment (intensity)		0.001			-0.013	
		(0.011)			(0.010)	
Post × Treatment (density)			0.010			-0.001
			(0.022)			(0.023)
Woredas FE	✓	✓	✓	✓	✓	✓
Time FE	✓	✓	✓	✓	✓	✓
Time-varying controls	✓	✓	✓	✓	✓	✓
IP-weights	✓	✓	✓	✓	✓	✓
Observations	4606	4606	4606	2044	2044	2044
R-squared	0.91	0.91	0.91	0.95	0.95	0.95

Notes: Standard errors in parentheses are clustered at the level of the treatment (*woredas*). *** p<0.01, ** p<0.05, * p<0.1. See notes to Table 3.4 for other details.

Table C5: Main results with extended controls

	Highlands			Lowlands		
	NDVI	NDVI	NDVI	NDVI	NDVI	NDVI
	(1)	(2)	(3)	(4)	(5)	(6)
Post × Treatment (dummy)	0.003			-0.005		
	(0.002)			(0.004)		
Post × Treatment (intensity)		0.000			-0.002	
		(0.001)			(0.001)	
Post × Treatment (density)			0.001			-0.002
			(0.002)			(0.004)
Woredas FE	✓	✓	✓	✓	✓	✓
Time FE	✓	✓	✓	✓	✓	✓
Time-varying controls	✓	✓	✓	✓	✓	✓
IP-weights	✓	✓	✓	✓	✓	✓
Observations	4606	4606	4606	2044	2044	2044
R-squared	0.91	0.91	0.91	0.95	0.95	0.95

Notes: Time varying controls include climatic variables (i.e. rainfall, temperature, and their respective quadratic terms), night time lights, and population density. Standard errors in parentheses are clustered at the level of the treatment (*woredas*). *** p<0.01, ** p<0.05, * p<0.1. See notes to Table 3.4 for other details.

Table C6: Main results with no restriction to the common support region

	Highlands			Lowlands		
	NDVI	NDVI	NDVI	NDVI	NDVI	NDVI
	(1)	(2)	(3)	(4)	(5)	(6)
Post × Treatment (dummy)	0.001			-0.005		
	(0.002)			(0.003)		
Post × Treatment (intensity)		-0.001			-0.002*	
		(0.001)			(0.001)	
Post × Treatment (density)			0.000			-0.001
			(0.001)			(0.003)
Woredas FE	✓	✓	✓	✓	✓	✓
Time FE	✓	✓	✓	✓	✓	✓
Time-varying controls	✓	✓	✓	✓	✓	✓
IP-weights	✓	✓	✓	✓	✓	✓
Observations	6384	6384	6384	2478	2478	2478
R-squared	0.93	0.93	0.93	0.96	0.96	0.96

Notes: Standard errors in parentheses are clustered at the level of the treatment (*woredas*). *** p<0.01, ** p<0.05, * p<0.1. See notes to Table 3.4 for other details.

Table C7: Impacts on land conservation

	Highlands			Lowlands		
	CA	CA	CA	CA	CA	CA
	(1)	(2)	(3)	(4)	(5)	(6)
Post × Treatment (dummy)	61.811			-5.836		
	(46.214)			(78.157)		
Post × Treatment (intensity)		26.178			-11.407	
		(22.692)			(28.474)	
Post × Treatment (density)			-46.367			107.363
			(35.305)			(74.953)
Woredas FE	✓	✓	✓	✓	✓	✓
Time FE	✓	✓	✓	✓	✓	✓
Time-varying controls	✓	✓	✓	✓	✓	✓
IP-weights	✓	✓	✓	✓	✓	✓
Control mean	548.72	548.72	548.72	143.18	143.18	143.18
Observations	658	658	658	292	292	292
R-squared	0.94	0.94	0.94	0.90	0.90	0.90

Notes: The outcome variable corresponds to the number of pixels cultivated at time t and still cultivated at time $t+1$ (with $t = 2001 \mid 2005$ and $t+1 = 2005 \mid 2013$). Standard errors in parentheses are clustered at the level of the treatment (*woredas*). *** p<0.01, ** p<0.05, * p<0.1. See notes to Table 3.4 for other details.

Table C8: Impacts on migration

	Highlands			Lowlands		
	Mig	Mig	Mig	Mig	Mig	Mig
	(1)	(2)	(3)	(4)	(5)	(6)
Post × Treatment (dummy)	-0.004			-0.198		
	(0.008)			(0.164)		
Post × Treatment (intensity)		0.000			-0.059	
		(0.003)			(0.046)	
Post × Treatment (density)			-0.002			-0.230
			(0.005)			(0.172)
Woredas FE	✓	✓	✓	✓	✓	✓
Time FE	✓	✓	✓	✓	✓	✓
Time-varying controls	✓	✓	✓	✓	✓	✓
IP-weights	✓	✓	✓	✓	✓	✓
Control mean	0.092	0.092	0.092	0.128	0.128	0.128
Observations	1576	1576	1576	600	600	600
R-squared	0.50	0.50	0.50	0.24	0.23	0.23

Notes: Authors' calculations based on IPUMS 2007 data. The outcome variable corresponds to the immigration rate (per 1,000 individuals) for woreda w at time t (with $t = \{2000, 2007\}$). Standard errors in parentheses are clustered at the level of the treatment (*woredas*). *** p<0.01, ** p<0.05, * p<0.1. See notes to Table 3.4 for other details.