Responding to shocks: Cash, transfers and imagery



Ashley Pople

Hertford College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy in Economics

March 2022

Responding to shocks: Cash, transfers and imagery

Ashley Pople

Hertford College

University of Oxford

A thesis submitted for the degree of Doctor of Philosophy in Economics

March 2022

This thesis consists of three standalone chapters that speak to the broad theme of responding to shocks in low- and middle-income countries. Chapter 1 investigates the impact of receiving cash transfers in anticipation of an extreme climate disaster, thereby enabling low-income households to cope with the shock on their own terms. We present quasi-experimental evidence showing the effectiveness of a timelier forecast-based cash transfer delivered to low-income households just before the peak of severe flooding in Bangladesh. We find that the 'anticipatory' cash transfer enabled households to change their behaviour and improve their welfare, both during the flood and three months after the intervention. Chapter 2 explores the role of mental imagery in overcoming trauma – largely induced by exposure to violence – by fostering the ability to imagine the future. We design and test a ten-session entrepreneurship training curriculum that teaches mental imagery in business decision making through a randomised control trial in Colombia. We find that mental imagery can be taught as a skill and those who receive imagery training improve their earnings, relative to those who receive a placebo entrepreneurship training. Chapter 3 explores the interaction between public and private cash transfers, recognising that social protection may interact with existing coping strategies. I estimate the extent to which receiving an old-age pension crowds out remittances in post-Apartheid South Africa, using a regression discontinuity design. I find evidence of substantial crowding out in a middle-income country context: for every 1 rand of pension received, 0.22 rand of remittances are displaced on average.

Supervisors This thesis was supervised by Stefan Dercon and Abigail Adams-Prassl.

Total words This thesis contains approximately 76,250 words.

Statement of authorship Chapter 1 is joint work with Stefan Dercon, Ruth Hill and Ben Brunckhorst. Chapter 2 is joint work with Nava Ashraf, Gharad Bryan, Alexia Delfino, Emily Holmes and Leonardo Iacovone. Chapter 3 is single authored.

Acknowledgement of funding I acknowledge funding from the Rhodes Trust, Department of Economics and Hertford College.

Acknowledgements

I embarked on the PhD in 2016 with the hope to reflect on big questions, engage in thought-provoking discussions, and deepen my passions. The fact that I was able to achieve all of these is a real testament to the support I have received from my community. Drawing on my South African childhood, "ubuntu" comes to mind: *I am because we are*. I could not have reached this milestone without travelling the journey with others, to whom I am deeply grateful.

To my supervisors, Stefan Dercon and Abi Adams-Prassl: I am deeply appreciative for your mentorship and wisdom over the years. I have learnt so much from you both. Thank you for inspiring me and exemplifying the type of economist I wish to be. Every interaction still leaves me brimming with ideas and excitement, while reminding me what it means to be human.

To all my co-authors and research partners: thank you for making research so enriching and fun. You made the PhD a far happier journey and reinforced my belief that research should be fundamentally collaborative. I will always remember our debates, and the nights we took off to dance salsa or share a meal.

To my academic community at the Centre for the Study of African Economies (CSAE): Simon Quinn, Doug Gollin, Pramila Krishnan, Chris Adam, Anandi Mani and Sanjay Jain, as well as Hamish Low, Johannes Abeler and Ian Crawford in the Applied Micro group. Thank you for all that you have taught me about economics and the friendly feedback that developed my ideas. A big thank you to Rose Page, the captain of the CSAE mothership, for providing never-ending research support.

To the Rhodes Trust: thank you for enabling me to realise my dream of coming to Oxford and for offering me a community where deep friendships were made. I also appreciate the funding I received from the Department of Economics and Hertford College, and the various grants that made my research possible.

To the wonderful humans who have walked with me through the DPhil journey and have become the closest of friends: Katherine Stapleton, Alex Copestake, Diala AlMasri and Jamie Walsh. Thank you for the walks, hugs and warmth that carried me through the hard times, and the laughter and fun that characterised the

good. Thank you to the broader DPhil community for bringing sunshine to my DPhil days: Verena, Vanessa, Ben, Hannah, Rocco, Patrick, Sanghamitra, Giulio, Vatsal, Muhammad, Shihang, Binta and Lukas.

To my community of housemates at 24 Marlborough Road: Laura, Rob, Seamus, Chris, Mizu, Emily, Laure, Davo, Joanna, Tanyah, Fritzi, Rebecca, and Katherine. Thank you for creating home for me in Oxford over the past five years. I always looked forward to coming home and sharing meals. You became my family, especially during lockdown. I would not have kept sane through the DPhil and pandemic years without you (and all our adventures on kayaks, in rivers or Hinksey lake, and in the final months, morning coffees at Paperboat Cafe).

To my community of friends: thank you for me making me who I am today. To the London lads, Mayanka, Andrew, Michelle, Jacob, Peter and KP: you reminded me to laugh and have fun throughout my Oxford experience, while always being a rock of support. To Ahmed, Tinashe, Vafa and Carl: thank you for cheering me up with your spontaneous phone calls and long walks. Thank you to my friends back home – Jess, Lily, Flavia, Isabeau, Beata, David and Stephanie – for your steadfast friendships and the roles you played in shaping the path I am on today. To Maria Paula and Katarina: thank you for our adventures in Bogotá and Lilongwe, and giving me a second home during my research. I am grateful to Mariella for teaching me how to do a headstand and grounding me through biweekly yoga during my final year of the DPhil.

To my family: thank you for your love, unwavering encouragement, and support throughout my education. David, Véronique, Georgina, Julian, Michele, Dorrie, Mamoune, Mike – you empowered me to pursue my dreams. To my grandfather, Robert Toletti, thank you for being my role model in life. I wish you could have been here to see me to the end, but I will forever carry forward the lessons you have taught me about how to live a good life and survive out in the mountains.

Last, but not least, thank you to JJ for your being my biggest advocate in life. Thank you for filling my final year of the DPhil with so many flowers, almonds, salsa, weekend adventures and hugs.

While I look ahead with excitement to the next chapter, I also reflect on the last five magical years with fondness and gently draw this chapter to a close.

Ashley Pople
Oxford, October 2021

Contents

Introduction **1**

1 Anticipatory Cash Transfers in Climate Disaster Response **9**

1.1 Introduction 11

1.2 Intervention 16

1.3 Sample and data 20

1.4 Empirical strategy 32

1.5 Results 37

1.6 Heterogeneity analysis 44

1.7 Robustness to model specification and sample choice 52

1.8 Results five months post-intervention 59

1.9 Conclusion 62

Appendices of Chapter 1 **66**

2 Learning to See the World's Opportunities: The Impact of Imagery on Entrepreneurial Success **85**

2.1 Introduction 87

2.2 A primer on imagery, future thinking and trauma 93

2.3 Correlations between imagery and economic outcomes 98

2.4 Intervention 103

2.5 Experimental design 109

2.6 Measurement . . . 115
2.7 Empirical strategy . . . 120
2.8 Results . . . 121
2.9 Robustness checks for attrition . . . 142
2.10 Concluding remarks . . . 146

Appendices of Chapter 2 **149**

3 Cash Transfers and Crowding Out of Remittances: Evidence from South Africa **163**

3.1 Introduction . . . 164
3.2 Context . . . 167
3.3 Data . . . 168
3.4 Empirical strategy . . . 172
3.5 Results . . . 180
3.6 Robustness to model specification and placebo tests . . . 182
3.7 Crowding out over time . . . 191
3.8 Conclusion . . . 193

Appendices of Chapter 3 **197**

Conclusion **205**

References **209**

List of Figures

1.1 Timeline of triggers and cash transfers 19
1.2 Map of the WFP intervention along the Jamuna River 19
1.3 Effect of receiving a cash transfer 38
1.4 Number of households receiving cash relative to the local flood peak date 46
1.5 Non-parametric analysis on the timing of the cash transfer 50
1.6 Standardised treatment effects on sub-samples by wallet-use 54
1.7 Robustness checks to alternative model specifications 56
1.8 Welfare effects over time within panel unions 60
1.9 Child and adult food consumption 10 and 20 weeks after the intervention 62
1.10 Food security during the flood 84

2.1 Baseline imagery, assets and life satisfaction 99
2.2 Imagery and earnings pre-Covid-19 101
2.3 Subjective probability of imagined scenarios and earnings in the months before Covid-19 102
2.4 Quality of positive and negative imagery by baseline trauma 104
2.5 Timeline of implementation and data collection 111
2.6 Effects of receiving the imagery training on downstream outcomes . 125
2.7 Effect of receiving the imagery training relative to the no-intervention group 129

2.8 Effect of receiving the placebo training relative to the no-intervention group . . . 130
2.9 Linear prediction of imagery training on economic outcomes, by baseline trauma . . . 138

3.1 Probability of pension receipt by age of oldest household member . . 176
3.2 Density of age of oldest household member . . . 177
3.3 Continuity test of predetermined household characteristics at the age eligibility threshold . . . 179
3.4 Mean predicted value of remittances, by age of the oldest household member . . . 182
3.5 RD effect of the pension on remittances for different bandwidth choices . . . 184
3.6 Probability of receiving a pension, with placebo thresholds . . . 189
3.7 RD effect of the pension on remittances for placebo thresholds . . . 190
3.8 Real household remittances, by age of the oldest household member 194
3.9 Employment and migration patterns around the age eligibility threshold for Wave 3 . . . 198
3.10 Continuity test of the race variable for Wave 1 . . . 199
3.11 Donut hole test . . . 203

List of Tables

1.1 Sample size, by transfer date and treatment status 23
1.2 Balance and summary statistics 25
1.3 Treatment effect of receiving anticipatory cash transfer 39
1.4 Treatment effect of receiving anticipatory cash transfer one day earlier 47
1.5 Timing effects using 14 July treatment group only 48
1.6 Treatment effect of receiving anticipatory cash transfer by land type . 51
1.7 Robustness check: Intent-to-treat estimates using 14 July only 57
1.8 Lee bounds to correct for differential non-response 59
1.9 Standardised mean treatment effect on welfare outcomes within panel unions (second round only) 61
1.10 Balance across treatment groups by transfer date 67
1.11 Balance in sample from second round of surveys 68
1.12 Variable construction for welfare measures 69
1.13 Variable construction for other outcome measures 70
1.14 False Discovery Rate q-values and wild bootstrap inference 72
1.15 Treatment effect on secondary outcomes 73
1.16 Treatment effect on main outcomes for 14,15,16 July transfers vs 30 July transfers 74
1.17 Marginal effect of an earlier cash transfer on secondary outcomes . . 74
1.18 Marginal effect of an earlier cash transfer - IV 75
1.19 Marginal effect of an earlier cash transfer - excluding 30 July 75

1.20 Summary statistics by land type 76
1.21 Treatment effect on secondary outcomes by land type 77
1.22 Standardised treatment effects on sub-samples by wallet-use 78
1.23 Standardised treatment effects on sub-samples by wallet-use 79
1.24 Robustness to alternate model specifications 80
1.25 Standardised mean treatment effects - welfare 81
1.26 Standardised mean treatment effects - livelihood, finance and actions 81
1.27 Robustness to controlling for satellite measures of flood intensity at mauza level 82
1.28 Welfare effects over time - panel unions 83
1.29 Test for substitution of adult and child food consumption over time in panel control households 84

2.1 Baseline Balance Table 113
2.2 Endline Attrition Table 115
2.3 Imagery vs Placebo Training: Effect on Imagery 123
2.4 Imagery vs Placebo Training: Effect on Downstream Outcomes . . . 126
2.5 Imagery vs Placebo Training: Effect on Secondary Outcomes 128
2.6 Lee bounds to correct for differential non-response 132
2.7 Treatment effects on imagery, by trauma (business scenarios) 137
2.8 Treatment effects on imagery, by gender (business scenarios) 140
2.9 Effect on downstream outcomes, by gender 141
2.10 Baseline characteristics by gender 143
2.11 Imagery vs Placebo: Effect on Imagery for "Territorial" Sub-Stratum . 144
2.12 Imagery vs Placebo: Effect on Economic Outcomes for "Territorial" Sub-Stratum 145
2.13 Breakdown of sample by subdivision and treatment status 150
2.14 Measurement: Imagery 151
2.15 Measurement: Psychological resilience 152
2.16 Measurement: Economic Outcomes Pre-Covid 152

2.17 Measurement: Economic Outcomes During Covid 153
2.18 Effect on non-business imagery for business sample 154
2.19 Effect on non-business imagery for entire sample 155
2.20 Imagery vs No-intervention: Effect on Downstream Outcomes 156
2.21 Placebo Training vs No-intervention: Effect on Downstream Outcomes 156
2.22 Imagery vs No int: Effect on Downstream Outcomes for "Territorial" 157
2.23 Placebo vs No int: Effect on Downstream Outcomes for "Territorial" . 157
2.24 Effect on downstream outcomes, by trauma 158
2.25 Attrition, conditional on acceptance 159
2.26 Attrition in the first follow-up survey 160
2.27 Treatment effects on imagery, by trauma (non-business scenarios) . . 161
2.28 Entire sample: Treatment effects on imagery (business scenarios) . . 161
2.29 Entire Sample: Treatment Effects on Downstream Outcomes 162

3.1 Summary statistics 170
3.2 Descriptive statistics 171
3.3 RD effect of the state pension on household remittances 181
3.4 Robustness to alternative model specifications 186
3.5 Robustness to controls from the previous wave 187
3.6 Donut hole test 188
3.7 RD effect of the state pension on household remittances, over time . 192
3.8 Descriptive statistics for households that receive or send remittances in Wave 3 vs Wave 5 195
3.9 Relationship of the sender of remittances for control households . . . 198
3.10 RD effect on predetermined covariates at the age eligibility threshold 199
3.11 RD effect of the pension on remittances, by gender of oldest household member 200
3.12 RD effect of the pension on remittances, by household wealth 201
3.13 RD effect of the pension on household remittances, by bandwidth size 202

3.14 RD effect of the pension on household remittances, using placebo thresholds 204

Introduction

Low-income households are disproportionately exposed to shocks and variations in income, largely due to factors beyond their control. The severity and frequency of these shocks are likely to worsen over time, as conflict and climate disasters become more prevalent. There is potential for large welfare gains from smoothing consumption, given that negative coping strategies can have long-lasting effects. What are the best levers for keeping household consumption smooth in the face of these shocks? This thesis explores the question through three different angles. Chapter 1 investigates the impact of receiving cash transfers in anticipation of a climate disaster in a low-income setting, thereby enabling households to cope with the shock on their own terms. Recognising that cash transfers alone may be insufficient if the shock induces behavioural constraints ex post, Chapter 2 explores the role of imagery in overcoming trauma by fostering the ability to imagine the future. However, these policy levers do not operate in a vacuum. Traditionally, transfers shared both within and across households have served as one of the primary means to insure against shocks. Chapter 3 examines the interaction between public cash transfers and these private support networks. I find evidence that the effectiveness of public cash transfers may be neutralised by an endogenous response in remittance behaviour.

With global surface temperatures continuing to increase, large covariate shocks such as drought and floods have been more frequent and more intense (IPCC, 2021), while violent conflict is on the rise (United Nations and World Bank, 2018). As exemplified by the Covid-19 pandemic, these crises tend to affect the poor disproportionately and often surpass their ability to cope. Households resort to negative coping strategies, such as foregoing consumption, costly borrowing,

selling assets, taking children out of school or migrating (Dercon, 2002). These coping strategies – often temporarily adopted – can have long-lasting impacts on welfare and development (Maccini and Yang, 2009; Dercon and Porter, 2014). In light of these long-term effects of shocks, there are large welfare gains that can be obtained from smoothing consumption.

In response, there is a large literature highlighting the effectiveness of social protection, specifically in the form of cash transfers, in cushioning the negative impacts of shocks and supporting households to smooth their consumption (de Janvry et al., 2006; Del Carpio and Macours, 2010; Jensen et al., 2017; Asfaw et al., 2017; Knippenberg and Hoddinott, 2017; Adhvaryu et al., 2018; Premand and Stoeffler, 2020). However, despite a rapid expansion in cash transfer programmes during the Covid-19 pandemic (Gentilini et al., 2021), the coverage of affected households who are targeted remains limited. The response also tends to be ex-post and too late, after households have felt the worst impacts of the shock and adopt negative coping strategies. At the same time, these crises are becoming increasingly more predictable, with advances in computing power and availability of rich, non-traditional data.[1] These technological innovations present a new opportunity to manage risks in advance, rather than retroactively addressing the impact of the shocks. As a result, forecast-based 'anticipatory' transfers targeting at-risk households are becoming more prevalent in both the humanitarian and development sectors.

Despite the increasing attention placed on 'shock-responsive' social protection, there is limited evidence on the impact of timelier cash transfers on household welfare. Chapter 1 (joint with Stefan Dercon, Ruth Hill, and Ben Brunckhorst) presents quasi-experimental evidence on the welfare impact of a one-off anticipatory cash transfer delivered in advance of a climate shock. In July 2020, the World Food Programme sent approximately USD $53 (equivalent to two weeks of household food expenditure) using mobile money to over 23,000 extreme poor households that were forecast to experience severe flooding in Bangladesh. The floods that followed were some of the most severe and protracted in decades. To

[1] Analysis of UN-coordinated appeals suggests that at least 55 percent of crises are somewhat predictable (Weingärtner and Spencer, 2019).

assess impact, we compare households who received cash to comparable households within the same local area who did not receive cash due to having access to the incorrect mobile wallet or inactive bKash account at the time of the verification call. We surveyed more than 9,000 households using phone surveys three months after the intervention.

We find that the small transfer not only provided immediate relief by reducing the number of days households went without eating during the flood, but also helped households change their behaviour in ways that could have long-term benefits. In particular, the transfer enabled households to take action and prevent the loss of productive assets. Treated households were less likely to borrow after the flood and borrowed less at lower rates. After three months, we find that these households were better off: they had higher food consumption, wellbeing and earning potential. Our most striking result is the protection in child food consumption, given that short-term under-nutrition is known to have lasting impacts on education and labour market outcomes (Chen and Zhou, 2007; Maccini and Yang, 2009; Dercon and Porter, 2014; Dinkelman, 2017; Victora et al., 2021). Children in treated households were three percentage points more likely to consume three or more meals on the day prior to the survey. This result is robust to a variety of alternative model specifications. We also find that the speed of delivery matters: cash transfers received one day earlier before the flood peak improves consumption. We conclude that the small cash transfer set households on a different trajectory by alleviating liquidity constraints at a critical juncture of the shock.

Although highly impactful, cash alone may be insufficient in mitigating the long-lasting effects of a shock if behavioural constraints are induced ex post. For example, exposure to shocks can be highly traumatic and the resulting trauma can in turn affect the ability to imagine the future and thus make non-biased forward-looking decisions.[2] Studies in neuroscience and psychology have shown that trauma increases the weight placed on negative imagined future scenarios relative to positive scenarios (McNally et al., 1995; Tulving, 2002; Moore and Zoellner, 2007; Kleim and Ehlers, 2008; Kleim et al., 2014).

[2]Over 70% of people are estimated to have experienced at least one traumatic event in their lifetime drawing on multi-country mental health surveys, highlighting the broader relevance of this phenomenon (Benjet et al., 2016; Kessler et al., 2017).

Chapter 2 (joint with Nava Ashraf, Gharad Bryan, Alexia Delfino, Emily Holmes and Leonardo Iacovone) explores whether teaching the ability to imagine the future improves entrepreneurship outcomes for low-income micro-entrepreneurs in Colombia. We design a ten-session entrepreneurship training curriculum that teaches mental imagery in business decision making in a balanced way. Mental imagery is defined as the process of mental experiencing and "seeing with the mind's eye", by accessing information from memory. We test the impact of teaching imagery on entrepreneurship outcomes through a randomised control trial with roughly 2,000 would-be entrepreneurs in Bogotá, whereby a large share of the sample report high symptoms of trauma. We randomly assign entrepreneurs to three treatment arms: (1) the imagery training; (2) a placebo entrepreneurship training; and (3) a no-intervention group. The imagery and placebo training arms have an identical structure and delivery method, and cover the same entrepreneurship content. The sole difference can be found in the imagery exercises in the imagery training and the use of written or verbal group-based activities in their place in the placebo training. During the Covid-19 pandemic, we conduct two follow-up phone surveys 8 and 14 months after the end of the intervention.

We measure the frequency of use and quality of imagery employed using standard psychological scales and find that imagery can be trained. Consistent with the notion that imagery training could rebalance the way in which the future is imagined, we also find that imagery participants with higher levels of baseline trauma report significantly more vivid positive future imagery relative to low trauma participants. Moreover, those who receive the imagery training have significantly higher earnings, both before and after the start of the Covid-19 pandemic.

However, when we compare the imagery training to a group who receives no training, we find small to no results across all outcomes of interest. When we turn to the placebo training arm, we find that placebo participants report significantly lower earnings relative to the no-intervention group. There is some evidence for differential attrition in survey response during the pandemic, but we are unable to rule out the hypothesis that the placebo training may have caused entrepreneurs to perform worse. In this case, imagery could have served to neutralise the harmful effects of the business training. Owing to the Covid-19 pandemic, we were limited

to collecting data via phone surveys, which in turn restricts our ability to unpack mechanisms linking imagery to economic outcomes. Nevertheless, our findings provide a strong proof of concept that investing in the effort to imagine the future is a worthwhile endeavour and can be taught through public policies that are scalable.

The first two chapters of this thesis explore policy levers that could be used to respond to the welfare effects of shocks. However, public policies do not operate in a vacuum: they may interact with existing coping strategies. In the absence of public support, households respond to shocks by drawing on a wide range of private coping strategies, such as relying on risk-sharing networks, remittances, and savings. From a partial equilibrium perspective, the effectiveness of public policies in reducing poverty and vulnerability of direct recipients could be diminished by perverse endogenous responses in household behaviour. The way in which public cash transfers might interact with private transfers is theoretically ambiguous (Becker, 1974; Barro, 1974; Cox, 1987). Thus, identifying whether crowding out of private support becomes an empirical question. In high-income countries where social protection systems are well-established and private transfers are small, several empirical studies conclude that the threat of displacement may be overstated (Altonji et al., 1992, 1997; Künemund and Rein, 1999). In contrast, more recent empirical work suggests that crowding out may be a relevant concern for low- and middle-income countries, where credit markets do not work well, risk-sharing networks are widespread and social protection may be expanding in size or coverage (Cox and Jakubson, 1995; Cox et al., 1998; Albarran and Attanasio, 2003; Jensen, 2003; Cox et al., 2004; Juarez, 2009; Gerardi and Tsai, 2014; Amuedo-Dorantes and Juarez, 2015). This literature largely draws on panel data and instrumental variable estimation methods to address the self-selection into social protection programmes.

Chapter 3 builds on this literature by evaluating the extent to which receiving a public cash transfer in the form of an old-age pension crowds out remittances in post-Apartheid South Africa. This paper is the first to employ a (fuzzy) regression discontinuity design to estimate the crowding out effect, exploiting the discontinuity in the probability of receiving the pension created by age eligibility rules.

Using data from the nationally representative National Income Dynamics Study, I examine the crowding out effect in the context of modern-day South Africa where social spending had doubled within a decade (National Treasury, 2012). I find a crowding out effect of 0.22 rand for every 1 rand of pension income received on average. The result is robust to a wide range of alternative model specifications. Age-eligible households located further away from the age eligibility threshold continue to experience a reduction in remittances within the cross-sectional data of interest.

This paper contributes to a growing literature that find substantial crowding out effects in low- and middle-income countries. In most of these contexts, as in my own, social protection had rapidly expanded in size or coverage in the years prior to the analysis. Yet I show that the crowding effect is only observed in the year 2012 and not in later years. In fact, remittances grow in both size and coverage over the years across the board. Although I cannot pin down the precise reason for this increase, I find consistent descriptive evidence that sending households are becoming wealthier over time and receiving households appear qualitatively poorer over time. The growing social protection system does not appear to be unravelling private support networks, but further research is required to verify this hypothesis.

These results have important implications for the efficacy of public cash transfers, as the pension is in effect shared across many household members. The welfare effect of the pension for direct recipients is likely to be partly neutralised by the endogenous response in remittances. However, if younger populations are indeed the primary source of remittances, the crowding out thereof in response to social protection may relieve the burden for working-age populations to support the elderly and instead free up capital for more productive uses. The broader welfare effects of the crowding out remains an open question for further investigation.

Contributions

Chapter 1 on anticipatory cash transfers provides three key contributions to the literature on cash transfers and climate disaster response. The first and main contribution is to evaluate the welfare impact of a one-off cash transfer received in

advance of a sudden climate disaster. The evidence base to date largely focuses on the role of a regular flow of cash transfers in cushioning the negative income effects of shocks (de Janvry et al., 2006; Del Carpio and Macours, 2010; Jensen et al., 2017; Asfaw et al., 2017; Knippenberg and Hoddinott, 2017; Adhvaryu et al., 2018; Premand and Stoeffler, 2020) or the impact of cash transfers delivered in the recovery phase after a disaster (De Mel et al., 2012; Mansur et al., 2017; Ivaschenko et al., 2020). Second, the paper is the first large-scale evaluation that rigorously measures the welfare impact of humanitarian cash transfers (Doocy and Tappis, 2017). More broadly, there is a dearth of evidence in the humanitarian sector, reflecting challenges with identifying a valid counterfactual group and conducting baseline surveys before a disaster strikes (Puri et al., 2017). Finally, little is known about how the modality of cash transfers affect their impact, such as timing or targeting. To this end, we present evidence that a timelier cash transfer is welfare improving.

Chapter 2 on mental imagery contributes to the entrepreneurship literature and existing work examining behavioural constraints in decision making. First, our results show the importance of mental imagery for entrepreneurship outcomes, suggesting that imagining the future is an important micro-foundation for inter-temporal decision making (Becker and Mulligan, 1997). Second, our paper affirms existing studies that find long-lasting effects of trauma on decision making (Voors et al., 2012; Callen et al., 2014) and shows that mental imagery could be a policy response to address these biases. Finally, we show that the promotion of 'soft' skills (here, in the form of mental imagery) enhances the effectiveness of entrepreneurship training, adding to a growing evidence base on this topic (McKenzie, 2021; McKenzie et al., 2021).

In Chapter 3, I find a relatively large crowding out effect of public cash transfers on private income transfers in post-Apartheid South Africa. The effect is similar in size to other studies in middle-income countries (Jensen, 2003; Cox et al., 2004; Lai and Orsuwan, 2009; Gerardi and Tsai, 2014; Amuedo-Dorantes and Juarez, 2015), yet I use an entirely different empirical approach in the form of a regression discontinuity design – the first and main contribution of this paper. This chapter further contributes to the literature on crowding out by examining whether the crowding out effect persists over time. Finally, this chapter

speaks more broadly to a literature on the effects of cash transfers on household behaviour, where the South African old-age pension has been subject to extensive study (Case and Deaton, 1998; Duflo, 2003; Bertrand et al., 2003; Edmonds et al., 2004; Edmonds, 2006; Ranchhod, 2006; Posel and Casale, 2006; Ardington et al., 2009). These papers show that the pension can positively affect household composition, labour supply and migration. I build on these existing studies by presenting evidence that the pension can also interact negatively with remittance behaviour, thereby diminishing income available to direct recipients.

CHAPTER 1

Anticipatory Cash Transfers in Climate Disaster Response

with **Ruth Hill, Stefan Dercon and Ben Brunckhorst**

Abstract In the face of increasing climate volatility and stretched aid budgets, more effective ways to support households in times of crisis are needed. This paper examines the welfare impact of an anticipatory cash transfer provided to ultra-poor households forecast to experience extreme floods in Bangladesh. Evidence on the impact of one-off cash transfers in disaster settings is limited despite their widespread use, and the welfare cost of delaying assistance is rarely considered. To assess impact, we compare households who received cash to comparable households who did not receive cash due to having access to the incorrect mobile wallet or inactive bKash account at the time of the verification call. We find that anticipatory cash was mostly spent on food and water, and beneficiary households were 36% less likely to go a day without eating during the flood. Three months after the flood, beneficiary households had significantly higher child and adult food consumption and subjective well-being. We find that they experienced lower asset loss, engaged in less costly borrowing after the flood, and reported higher earning potential. These benefits accrue in the months before a conventional humanitarian response would arrive, highlighting the welfare gains from acting early. We show further evidence that small changes in timing matter: receiving cash a day earlier resulted in a small increase in food consumption months later.

Acknowledgements *We are grateful to Daniel Pfister, Dirk-Jan Omtzigt, Leonardo Milano, Hannah Ker, Ahmadul Hassan and our World Food Programme partners for their ongoing support and provision of data; without them, this project would not have been possible. We thank Abi Adams-Prassl, Christopher Woodruff, Michelle Rao, Paris Kazis, Niger Dilnahar, Johannes Abeler, Hamish Low, Simon Quinn, Ferdinand Rauch and seminar participants at Oxford University, Centre for the Study of African Economies (CSAE), Innovation for Poverty Action and the Global Poverty Action Lab's Annual Researcher Gathering, Office for the Coordination of Humanitarian Affairs (OCHA), World Bank and the Bangladesh anticipatory action learning group for their helpful comments. Lastly, we acknowledge Data Analysis and Technical Assistance (DATA) for their support in conducting our large surveys. This randomised control trial was registered in the American Economic Association RCT Registry under trial number 6576 (https://doi.org/10.1257/rct.6576-1.1). It received ethics approval by the Oxford Central University Research Ethics Committee (Protocol No. ECONCIA20-21-24).*

1.1 Introduction

When disasters strike, governments or humanitarian organisations often step in to provide time-limited support to those worst affected. Increasingly, support is provided in the form of one-off cash transfers. This trend was emphasised by the unprecedented scale of temporary cash transfers introduced by governments in response to Covid-19. Gentilini et al. (2020) record 1,179 social protection measures in 212 countries during the first six months of the pandemic, with half taking the form of cash transfers. The impact evaluation literature on regular transfers indicates that they help households cope with shocks (de Janvry et al., 2006; Jensen et al., 2017; Asfaw et al., 2017; Knippenberg and Hoddinott, 2017; Adhvaryu et al., 2018; Premand and Stoeffler, 2020). Despite their popularity, evidence on the impact of one-off transfers in a crisis is limited, reflecting more broadly a dearth of evaluations in the humanitarian sector (Puri et al., 2017). As a result, little is known about whether one-off transfers have impacts beyond providing immediate support to consumption, and how their timing, size, and targeting affects welfare. Recently in the humanitarian sector, forecast-based 'anticipatory' transfers are becoming more prevalent. These transfers target vulnerable households in advance of a shock to help them mitigate impacts, protect livelihoods and cope in the aftermath, but little is known about the impacts of these either. In a world of increasing climate volatility and stretched aid budgets, a better understanding of how to support households effectively in times of crisis is needed.

Our paper addresses this gap by examining the welfare impact of a one-off anticipatory cash transfer provided to households facing extreme floods. In July 2020, the World Food Programme (WFP) sent BDT 4,500 (approximately US$53, equivalent to two weeks of household food expenditure) using mobile money to 23,434 ultra-poor households along the Jamuna River that were forecast to experience severe flooding. The floods that followed were some of the most severe and protracted in decades. This novel 'anticipatory action' approach to humanitarian relief uses impact-based forecasts to provide support before the peak of the disaster. It contrasts with a more traditional response that reacts ex post to a materialised need. In 2020, the anticipatory cash transfer reached affected households

100 days earlier than previous WFP interventions in the same context. As far as we are aware, this is the first large-scale evaluation that rigorously measures the welfare impact of humanitarian cash transfers provided either in anticipation or ex post response to a sudden onset climate disaster (Doocy and Tappis, 2017). Gros et al. (2019) and Gros et al. (2020), the papers most similar to our own, show that anticipatory cash transfers in Bangladesh and Mongolia, respectively, are promising, but these studies face dual limitations of small sample sizes and imbalanced comparison groups.

To assess impact, we compare households that received the cash transfer to otherwise comparable households that did not. Beneficiary households were selected if they were on the pre-existing lists of vulnerable households and had access to an active bKash mobile money account at the time of the verification call. Our control group is comprised of households that also met the first criteria but were deemed not to have access to the correct mobile wallet account that was also active during the beneficiary verification. As a result, only some households were selected to receive the cash transfer, despite meeting the same vulnerability criteria and having access to a mobile wallet account. There is no reason to believe that the selected households are any different from those who did not have an active bKash account at the time of the verification call. 10 to 12 weeks after the cash intervention, we randomly sample 9,130 households, using phone surveys, of which 8,954 households met our criteria for inclusion.

We find that the cash transfer significantly improves child and adult food consumption, and subjective wellbeing, even when measured three months after the intervention. Children in treated households were three percentage points more likely to consume three or more meals on the day prior to the survey (relative to 80% of control households). This result is salient in light of a large body of literature highlighting the long-term consequences of temporary child undernutrition on educational, income and health outcomes (Chen and Zhou, 2007; Maccini and Yang, 2009; Dercon and Porter, 2014; Dinkelman, 2017; Victora et al., 2021).

Our analysis points to a number of behavioural changes that could have contributed to these welfare effects three months after receiving the modest cash transfer. Cash was mostly spent on food and water; treated households were 10 per-

centage points less likely to go a day without eating during the flood shock itself (relative to 28% of control households). Treated households were more likely to evacuate household members and livestock relative to control households. The cash transfer significantly decreased asset loss (especially for productive assets) and costly borrowing. We also find that treated households reported higher earning potential three months later in terms of wage hours and crop cultivation. Using a second survey conducted five months after the transfers, we observe a persistent positive effect on food consumption for adults, but not for children who catch up among control households. Although the cash transfers likely conveyed information about the severity of the flood, we believe that the cash transfer played a primary role in alleviating liquidity constraints at a critical juncture in the flood. Our results are robust to alternative model specifications, different criteria for the comparison groups, and bounding treatment effects to account for differential non-response.

Although we are unable to directly compare anticipatory action to a traditional humanitarian response, these welfare benefits were all recorded before the traditional response arrived. Without cash, comparable ultra-poor households experienced lower food consumption for both adults and children, lower wellbeing, higher asset loss, higher borrowing, and less employment in the period before the arrival of a traditional response. Some of these deficits, even when temporary, are known to have permanent scarring effects - for example, child undernutrition (Victora et al., 2021) and costly borrowing - suggesting significant gains from acting early, even if later support helps recovery (Dercon, 2004; Alderman et al., 2006; Maccini and Yang, 2009). By arriving earlier, the small cash transfer may have set households on a different trajectory by easing liquidity constraints at a critical juncture.

There are three theoretical reasons why receiving the cash earlier before the shock may be more beneficial than receiving it later during recovery. Firstly, an earlier cash transfer could widen the choice set for coping strategies, such as evacuating, protecting productive assets or stocking up on food supplies. Secondly, there are typically market disruptions for the duration of a climate shock, resulting in a limited ability to trade and the inflation of prices, such as increased boat

prices during floods. Lastly, sudden onset disasters are traumatic, limiting mental bandwidth, and decisions made at a critical juncture could have long run consequences. Supporting food security and easing liquidity constraints at the onset of a disaster could improve decision making so to mitigate impacts and cope in the aftermath.

To investigate timing further, we explore whether small changes in when transfers were received relative to flood dynamics matter. We use exogenous variation in the number of days between receiving a transfer and peak local flooding in 630 mauzas (approximately village level) based on high resolution satellite data. We then investigate the relationship between the timing of the cash transfer and welfare effects using linear regression and non-parametric methods to estimate the marginal effect of receiving cash one day earlier. We find evidence to suggest speed matters: an earlier anticipatory transfer is welfare improving. Cash received earlier relative to the flood event had stronger effects on food consumption three months later. Welfare effects tend to dissipate for each day transfers are delayed.

Our paper contributes to a large literature on the role of cash transfers in alleviating poverty. There is strong evidence that a regular flow of cash transfers can cushion the negative income effects of shocks and thus partially limit the use of costly coping mechanisms (de Janvry et al., 2006; Del Carpio and Macours, 2010; Hou, 2010; Macours et al., 2012; Aker et al., 2016; Jensen et al., 2017; Asfaw et al., 2017; Knippenberg and Hoddinott, 2017; Adhvaryu et al., 2018; Premand and Stoeffler, 2020). However, rigorous evidence on the effectiveness of a one-off cash transfer remains scarce, especially in response to shocks or disasters, whereby existing evidence focuses on recovery. For instance, Haushofer and Shapiro (2016) show that receiving a GiveDirectly cash transfer in Kenya in a lump-sum or in monthly payments result in similar welfare improvements, albeit different expenditure patterns, but these cash transfers were not shock-responsive. Closer to our work, two papers consider the impact of a top-up payment delivered to households from regular cash transfer programmes after a sudden onset climate disaster (tropical cyclone Winston in Fiji). Using a regression discontinuity design, both papers found that households that received cash transfers in the months after the shock recovered more quickly than those who did not (Mansur et al.,

2017; Ivaschenko et al., 2020). In the aftermath of a slow-onset climate disaster (a drought), Del Carpio and Macours (2010) and Macours et al. (2012) use a cluster randomised control trial to show that a bi-monthly conditional cash transfer implemented by the Nicaraguan government had positive persistent impacts on child health, development, and labour. In a different domain, a randomised control trial implemented in Sri Lanka several months after the 2004 tsunami found that a one-off cash grant to firms facilitated business recovery and investment (De Mel et al., 2012). Our paper evaluates the impact of a one-off ex ante cash transfer on ultra-poor households' ability to cope with a sudden onset climate disaster.

Our paper also makes a significant contribution to the literature on humanitarian evaluations. A review of 900 studies of humanitarian programmes found that only 31 could be classified as impact evaluations, of which only eight studies focused on the emergency relief phase (Puri et al., 2017). Weingärtner et al. (2020) specifically review the evidence for anticipatory humanitarian action and find that little rigorous evidence of impact is available. Evaluations of humanitarian interventions have often focused on the impact of regular programming in protracted crises, for example the impact of regular cash transfer programmes (for example: Lehmann and Masterson (2014); Battistin (2016); Aker (2013); Schwab (2019)).[1] The reasons are varied. For instance, it is often difficult to justify the ethics of a randomised control trial in life-or-death situations. The need for speed and the lack of transparency in implementation obfuscates the identification of a valid counterfactual. It is challenging to conduct a baseline when it is unknown *a priori* where a disaster will strike, and disasters ex-post disrupt the supply of basic services and infrastructure, including those needed for data collection. We address these challenges that are present in our own context by collecting a large post-intervention survey and exploiting the plausibly exogenous variation created by the incomplete targeting of households due to administrative constraints. Our identification strategy enables causal impact analysis of a humanitarian intervention in a fast-onset emergency, with broader policy relevance for development programmes.

[1] As another example, there are several papers that examine the impact of nutrition interventions (for example: Aguayo et al. (2015); Tranchant et al. (2019)), but reviews highlight that there are few papers that have rigorously measured the impact of nutrition interventions in emergencies (Webb et al., 2014; Webb, 2015).

An extensive body of literature compares the modality of cash versus in-kind interventions (Leroy et al., 2010; Aker, 2013; Skoufias et al., 2013; Cunha, 2014; Hidrobo et al., 2014; Cunha et al., 2019), but timing as a modality remains underexplored (Aker et al., 2016; Asfaw et al., 2017; Clavijo, 2020). Our paper showcases that a faster transfer offers the opportunity to limit the welfare cost of the shock, with policy implications that go beyond the humanitarian sector. With recent advances in data and forecasting technologies, the onset of many disasters is more predictable. As a result, there has been growing interest in the role of 'shock-responsive' safety nets that can temporarily expand both in transfer size and coverage in response to shocks, without excessively burdening budgets (Andrews et al., 2018; World Bank, 2018). The expansion in social protection during the Covid-19 pandemic provides many opportunities for assessing the impact of one-off transfers, and a research base on this is starting to emerge (Abay et al., 2020; Arndt et al., 2020; Bottan et al., 2021). In addition to providing evidence on the impact of humanitarian interventions in sudden climate disasters, our analysis speaks to the potential welfare benefits from quickly and pro-actively adapting safety nets in response to shocks.

Our paper is structured as follows. The next two sections describe the intervention and sample. Section 3.4 outlines our empirical strategy and measurement. In Section 3.5, we present our results on the effects of cash relative to no cash, and the channels through which these welfare improvements might occur. We explore how the impacts varied with the timing of the transfer in Section 1.6.1 and by land type in Section 1.6.2. Section 3.6 demonstrates the robustness of our results to mobile money use, alternative model specifications, and bounding exercises. Section 1.8 uses a second round of cross sectional data to present effects of the cash transfer five months later, and Section 2.10 offers conclusions.

1.2 Intervention

The 2020 monsoon floods in Bangladesh were the second highest since 1989 and the second longest since 1998. More than one million households were inundated and 5.5 million people were directly affected by flooding at the beginning

of August 2020 (United Nations Resident Coordinator Office, 2020). Flood waters halted agricultural production, damaged infrastructure, and disrupted food markets, schools and health services. The Ministry of Agriculture estimates that 110,000 hectares of crop land was damaged, while 257 people lost their lives due to the floods (United Nations Resident Coordinator Office, 2020).

In July 2020, the United Nations piloted a novel approach to humanitarian action by employing a data-driven forecast to predict the impact of excess flooding along the Jamuna River[2], a particularly flood-prone area of northern Bangladesh. This forecast was used to trigger the release of anticipatory cash transfers prior to peak flooding. With support from the International Federation of Red Cross and Red Crescent Societies and the Bangladesh Red Crescent Society (BDRCS), the WFP sent BDT 4,500 (approximately $53) using mobile money accounts to 23,434 vulnerable households in 131 unions (the smallest administrative unit in Bangladesh). The cash transfer is equivalent to approximately two weeks' food expenditure (World Bank, 2019). The cash was intended to mitigate the worst effects of the flood shock on household food consumption and mortality during the flood itself, as opposed to post-flood recovery. Water-tight storage containers and animal feed were also provided to 7,000 and 11,760 families, respectively, by the Food and Agriculture Organization of the United Nations (FAO). The United Nations Population Fund (UNFPA) also delivered hygiene, dignity, and health kits to 15,000 women, girls, and transgender people. There was little overlap between the cash and non-cash interventions, as the UN agencies mostly operated in different districts. This analysis focuses on the cash transfers made by WFP only; we account for any overlap in Section 3.4.

The anticipatory action approach is novel for the humanitarian sector. The approach uses predictive analytics, pre-arranged finance and pre-agreed actions to respond to foreseeable humanitarian needs. The earlier response contrasts to a traditional humanitarian response, which reacts to a materialised need after a severe shock occurs. For instance, during the last severe floods in 2019, beneficiaries received cash only about 100 days after the flood peak, despite a very rapid approval

[2]Jamuna River is the local name for the Brahmaputra River.

for funding made by the UN Central Emergency Response Fund. In this anticipatory action pilot, most beneficiaries received cash several days prior to peak flooding.[3] Not only is an anticipatory approach faster, but it can also be operationally cheaper to plan for and deliver an intervention before a climate disaster.

The timing of the anticipatory cash transfers was determined by a pre-defined set of triggers, based on forecasts of water flow from upstream at a centrally located gauge station, as illustrated by Figure 1.1. Given the accuracy with which forecasted water flow predicts actual water flow in Bangladesh, these forecasts were used in the design of two triggers indicating the onset of an extreme flood event. The first 'readiness' trigger was activated on 4 July 2020 and set in motion preparatory activities, including beneficiary verification calls.[4] The second 'activation' trigger - activated on 11 July - initiated the cash transfers.[5] The forecasts predicted the flood peak at the Bahadurabad gauging station for 17 July 2020. WFP sent cash to households over a period of five days, on 14, 15, 16, 18 and 30 July, both just prior to and immediately after the flood peak. In contrast to this novel anticipatory approach, the traditional humanitarian response was initiated only after a needs assessment had been conducted on 4 August – a month after the first trigger. Figure 1.2 illustrates the geographic spread along the Jamuna River.

[3]Funding for the cash intervention was approved by the Central Emergency Response Fund 35 days after the flood peak in 2019, whereas advance plans and protocols enable the rapid disbursement of funds 14 days before the flood peak at Bahadurabad gauging station, the basis for forecasts, in 2020.

[4]The first trigger was activated once water flows forecasted by the Global Flood Awareness System (GloFAS) and/or the Bangladesh Flood Forecast and Warning Centre (FFWC) 15-day probabilistic warning models was predicted to be more than 50% likely to cross the one-in-five year return period threshold (100,000 m^3/s) over a period of three consecutive days at the Bahadurabad gauging station, with a 10-day lead time. This threshold was set to be 0.85 metres above the government danger level of 19.5 metres. The GloFAS is a global hydrological forecast and monitoring system that couples weather forecast with a hydrological model calibrated for the Jamuna River in Bangladesh. The FFWC is a government agency responsible for flood forecasts.

[5]The second activation trigger was reached once the water level forecasted by the FFWC's five-day lead time model crossed the government-defined danger level by an additional 0.85 metres at the Bahadurabad gauging station.

Figure 1.1: Timeline of triggers and cash transfers

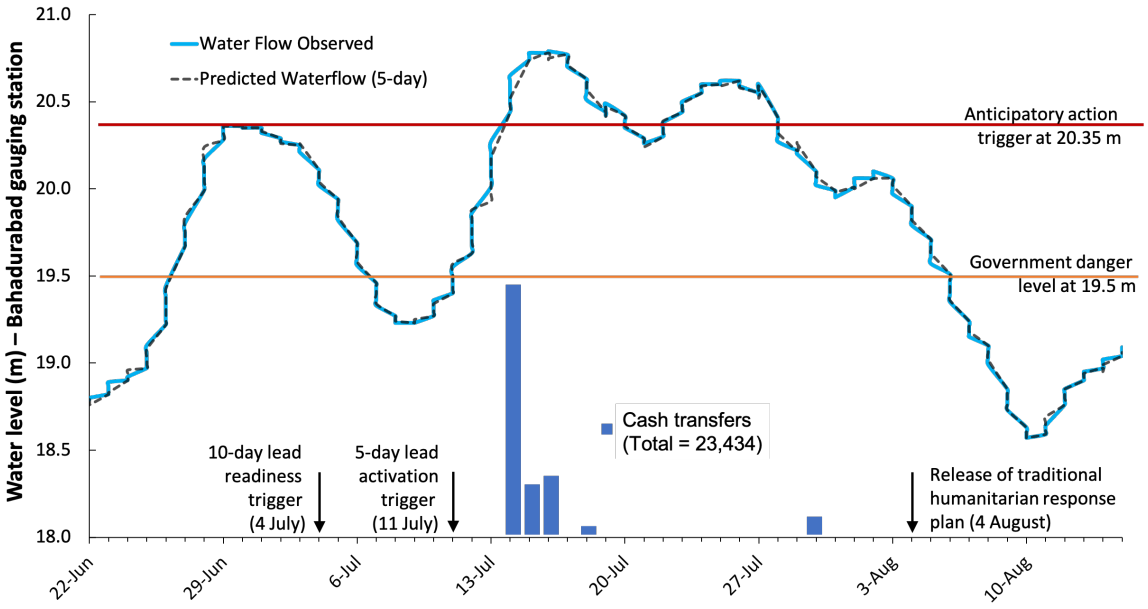
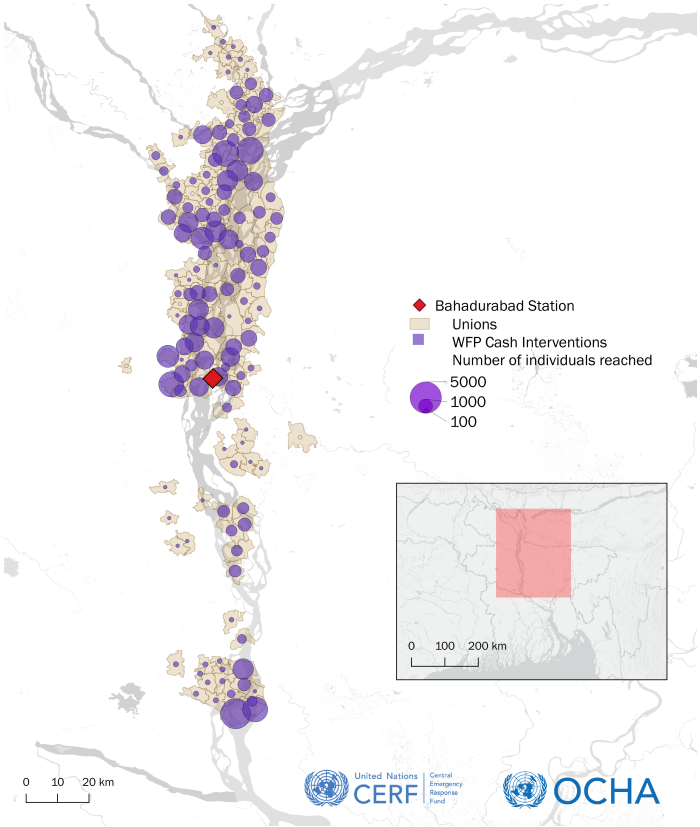


Figure 1.2: Map of the WFP intervention along the Jamuna River



It is important to note that cash transfers not only provide liquidity to some of the most vulnerable and poorest households, but also convey information about the severity of the oncoming flood. WFP did not inform households that the cash transfer was sent in anticipation of the incoming flood. Nevertheless, households may have interpreted the receipt of a cash transfer by WFP during rising waters as indicating the severity of the oncoming shock or as a call to action. Just under two thirds of households across the treatment and control groups in our sample reported receiving early warning from formal or informal sources. In the absence of a credible early warning system, the information component of the transfer could be particularly powerful. In earlier studies, early warnings were critical features in ensuring that affected populations adopted preventative actions before floods (Carsell et al., 2004; Kreibich et al., 2005; Thieken et al., 2007).

1.3 Sample and data

1.3.1 Sampling strategy

We leverage exogenous variation in treatment induced by the programme roll-out as a natural experiment to assess whether a one-off anticipatory cash transfer reduces the effects of a flood shock on household welfare. We randomly sampled beneficiary households and control households; the latter were otherwise comparable but did not receive the cash transfer, because they not have access to the correct active mobile wallet.

The process of household selection proceeded as follows. After activation of the readiness trigger in early July, the WFP and the International Federation of Red Cross and Red Crescent Societies Climate Centre first selected districts and unions in which to target beneficiaries based on a joint assessment of their flood risk and socio-economic vulnerability. Due to constraints imposed by the Covid-19 pandemic, pre-existing lists of vulnerable households were used to identify households in each union. The households on these lists had received assistance through UN interventions or government safety nets in previous years.

After the readiness trigger and union selection, WFP contacted as many households as they could reach on these pre-existing lists via phone over a pe-

riod of five days. These households were drawn at random across the prioritised unions and conditional on meeting the vulnerability criteria, all households had an equal probability of being selected. Nearly all households in our sample have access to one or more mobile wallet account, which was verified in our survey. There is very little perceived difference across service providers, both in terms of cost and function, except that bKash has a larger network coverage. Households frequently own more than one account, often to take advantage of sales promotions or gain access to government services, which were historically rolled out through different providers. During the call, the household's identity and location were verified and they were asked whether they had access to an active bKash account. Time constraints meant that WFP had only been able to establish a contract with one mobile money provider (bKash) for cash distribution, despite the prevalence of several mobile money services in the area.[6] Households with the incorrect mobile wallet account or inactive bKash account were excluded from receiving the cash. Those with an inactive bKash account (regardless of whether they have access to other accounts) could reactivate dormant accounts within a few days in order to receive the cash. Note that a bKash account becomes inactive if the account is not used to make or receive a financial transaction within six months, but reactivating the account is a simple and costless matter of contacting a bKash agent in the village.[7]

In sum, beneficiary households were selected if they were on the pre-existing lists of vulnerable households and had access to an active bKash mobile money account at the time of the verification call. Our control group is comprised of households that also met the first criteria but were deemed not to have access to the correct mobile wallet account that was also active during the beneficiary verification.[8] As a result, only some households were selected to receive the cash transfer, despite meeting the same vulnerability criteria and having access to a mobile wallet account. There is no reason to believe that the selected households are

[6]bKash currently holds approximately 75% of the market share, with Nagad and Rocket the main competitors holding; together, Nagad and Rocket hold approximately 20% of the market (International Finance Corporation, 2016).

[7]At the time of the survey, nearly all control households (97.4%) reported having a bKash account, highlighting how easy it is to set up this mobile wallet account.

[8]During the beneficiary verification, these households were recorded having access to a mobile money account with a different provider (e.g., Nagad or Rocket), a frozen bKash account, or no bKash account at all.

any different from those who did not have an active bKash account at the time of the verification call. Moreover, selection is independent of their experience of the flooding.

More than half of the cash transfers were sent to treated households on 14 January 2020 based on the above selection criteria used during the beneficiary verification call. If households did reactivate dormant bKash accounts in time, they received the cash transfer on 15, 16, 18 and 30 July 2020. WFP stated that they did not inform people during the verification call that they would be receiving a cash transfer. If this were true, households who reactivated their accounts and received the cash transfer during the later dates simply responded to the question about access to an active bKash account during the verification call. Therefore, we include these households in the treated group in the main analysis. Nevertheless, in order to account for the possibility that these households may have responded to the treatment, we will present a robustness check where we treat households that received cash after 14 July as a problem of compliance. In other words, they received the treatment when they should not have according to the selection criteria above and thus will be included in the control group. We also show that results are robust to controlling for wallet usage.

We randomly sampled households that met these criteria within targeted unions where there were at least 10 beneficiaries and relevant control households. For treated households, we randomly sampled all or 60 beneficiaries (whichever was smaller) from each union for those who received a cash transfer from 14 to 16 July, and all households receiving cash on 30 July.[9] For the control households, we similarly sampled all or a random sub-sample of households that were recorded as excluded by WFP for one of the aforementioned reasons in equal proportions across unions.

[9] A small number of households received cash on 18 July and we used this group to pretest the questionnaire. They do not appear in our sample.

1.3.2 Description of the sample

Sample size

Phone surveys were conducted in Bangla by trained phone survey enumerators with 9,130 households, 10 to 12 weeks after the intervention between 21 September and 8 October 2020. We targeted the household member whose names were on the WFP list of potential beneficiaries and met the baseline vulnerability criteria. Respondents were asked a series of questions, including on demographics, behavioural response to the flooding, food consumption, household assets, life satisfaction, work, and use of the cash transfer (if applicable). Respondents received BDT 100 (approximately $1.18) in phone credit for completing the survey.

Of the 9,130 households surveyed, 8,954 households across 111 unions met our criteria for treated and control households based on their survey responses; 6,566 treated households reported receiving a cash transfer via their bKash mobile money from WFP in July, whereas 2,388 control households reported no transfer, but had access to a mobile money account at the time of the survey. Table 1.1 summarises the number of households targeted by WFP and surveyed, by treatment status and transfer date.

Table 1.1: Sample size, by transfer date and treatment status

	No. of households receiving cash (by transfer date)	No. of households surveyed (by transfer date)
Treatment	**23,434**	**6,566**
14 July	14,345	3,312
15 July	2,903	1,218
16 July	3,384	1,085
30 July	1,036	670
Date not confirmed	0	281
Control	0	**2,388**

Notes: The second column shows the number of treated households targeted by WFP across the four transfer dates, according to their administrative records. The third column shows the number of households that were successfully surveyed. Treated households reported receiving the cash transfer and having access to a bKash mobile money account by the transfer date, whereas control households had access only to a mobile money account at the time of the survey. A small sample of households (281 households) were unable to confirm the date on which they received the cash transfer.

Sample characteristics and flood experience

Table 1.2 describes our sample and compares the mean value of time-invariant individual and household characteristics across treatment and control groups. The target population is vulnerable: 97% of intended beneficiaries were female, of whom only a third have completed primary school. Just over a quarter live in the most fragile of housing structures, which could easily be damaged in the flooding, and there is a high number of dependents relative to adults within households. We also collected variables measuring the use of technology, since the intervention and the definition of our control group was based on access to mobile money. Around half our sample had used mobile money in the last six months (excluding the transfer from WFP); 80% own a mobile phone and almost all respondents had used a mobile in the last week at the time of the survey.

Table 1.2: Balance and summary statistics

	Control mean	Treatment mean	Δ	Norm. Diff.	*p*-value
Individual characteristics					
Age	37.47	38.79	1.32	0.10	0.108
Female respondent	0.97	0.97	0.00	0.00	0.426
Household head	0.18	0.22	0.03	0.08	0.705
Completed primary school	0.34	0.30	-0.04	-0.08	0.576
Household characteristics					
Household size	4.68	4.73	0.05	0.03	0.872
Dependency ratio	0.75	0.76	0.01	0.02	0.282
Raw material house	0.26	0.27	0.01	0.02	0.183
Distance to large water body (m)	1332.73	1249.13	-83.60	-0.06	0.394
Protected mainland	0.43	0.33	-0.09	-0.19	0.470
Unprotected mainland	0.27	0.27	0.00	0.00	0.100
Char land	0.30	0.40	0.09	0.20	0.018
Anticipatory action					
Received WFP cash transfer	0.00	1.00	1.00	.	.
Received dignity kit from UNFPA	0.06	0.14	0.09	0.29	0.000
Received feed or storage from FAO	0.04	0.07	0.03	0.15	0.011
Technology					
Used digital wallet in last six months	0.50	0.47	-0.03	-0.05	0.417
Own mobile	0.82	0.80	-0.02	-0.06	0.633
Uses someone else's mobile	0.16	0.18	0.03	0.07	0.482
Uses mobile at least once a week	0.97	0.96	-0.01	-0.03	0.609
Observations	2388	6566			

Notes: Δ reports the treatment mean minus the control mean. Norm. Diff. reports the normalised difference between the treatment and control group means, following Imbens and Rubin (2015). The last column reports the *p*-value from ordinary least squares regressions of each variable on the treatment dummy to test equivalence of means, controlling for union fixed effects and clustering standard errors at union level as in our main specification.

Our sample was also heavily affected by the floods, as demonstrated in survey responses. 85% of control households reported flooding at or above floor level, with 5% above roof level. Based on satellite imagery, 30% of the land in local areas where our sample is located was covered in water at the peak of flooding on average.[10] Moreover, the flooding persisted at above half the local maximum flood extent for an average of 45 days, an indication that this was one of the most protracted flood events in decades. Although 17 July was the predicted flood peak

[10]Satellite data was used to measure flood intensity at mauza level, the administrative unit below the union or approximately village level.

at the centrally located gauging station used in the triggers described in Section 2.4, the most common date of peak flooding was 22 July for households in our sample. The date of peak flood extent varied across mauzas from mid-July until the beginning of August, according to estimates derived from satellite imagery.

Only 61% of the control group reported receiving early warning and 53% took any action to prepare for the flood. Over half migrated to live elsewhere between the flood and the time of the survey. In terms of asset loss, control households lost the same number of small livestock (goats, sheep, and pigs) as they owned at the time of the survey (one animal for every two households); 60% of control households reported loss of poultry, another important productive asset; 46% lost cultivated crops and of these, only half have been able to replant since the flood. Coping strategies were exhausted. Almost a third of control households reported going at least one day without eating any food during the flood, which highlights the extreme poverty and vulnerability of households in our sample.

Balance across treatment and control groups

Households appear well balanced across treatment and control groups.[11] The normalised difference column shows that the largest discrepancy in means is no more than 0.1 times the standard deviation in almost all cases.[12] On average, socio-economic variables and mobile phone use are very similar across groups. Exceptions include the land type beneficiaries live on, and whether they were a recipient of other anticipatory interventions from UN organisations.

We report the p-values from ordinary least squares regressions of time-invariant and technology use variables on the treatment dummy to test equivalence of means.[13] We fail to reject at the 10% level in all cases except for the

[11] Appendix Table 1.10 shows that households receiving cash across the four transfer dates are similar in terms of time-invariant variables.

[12] The normalised difference is strictly defined as the difference in means, divided by the square root of half the sum of treatment and control group variances, following Imbens and Rubin (2015). They use propensity score matching to show that normalised differences of 0.25 or less are well balanced in terms of experimental results but covariate adjustment is unreliable for values of 1 or more.

[13] We control for union fixed effects and cluster for standard errors at union level, as in our main specification.

proportion of households on unprotected char land, and those receiving other interventions from the UN. 40% of treatment households are located on unprotected char land, relative to 30% of the control group, and are therefore potentially more exposed to flooding. There was a limited amount of geographic overlap between the cash transfers and other UN interventions received by a small subset of our sample due to the use of common pre-existing lists within those areas. A higher share of treated households received these additional interventions relative to control households: 14% of beneficiaries in our sample received a dignity kit from UNFPA, while 7% received animal feed or food storage from FAO, relative to 6% and 4% of the control group, respectively.

Although treatment and control households are balanced on observable characteristics and their use of mobile money and phones, it is possible that there may be important unobservable differences that limit our ability to attribute the estimated impacts to the cash transfer provided. For example, a higher proportion of treated households on char lands, which are typically fertile, may be better able to recover after the floods, thus resulting in an overestimation of the treatment effect. At the same time, these households are extremely exposed to the flooding and thus may be disproportionately affected by the flood shock, resulting in an underestimation of the effect on flood mitigation, even if recovery is faster. Similarly, treated households receiving other UN interventions may reflect other baseline vulnerabilities. In order to address these concerns, we test the robustness of our results against a variety of alternative model specifications, as discussed in Section 1.4.3 below. Specifically, we show that our results are robust to excluding households that received other anticipatory action interventions and conducting subgroup analysis within each land type.

1.3.3 Other data used in the analysis

Second round of phone surveys

Between 22 and 28 December 2020, additional phone surveys were conducted as part of a separate evaluation of WFP's post-flood cash transfers. Some of the surveys were conducted in unions that had been excluded from the post-flood cash

transfer, because these were deemed to have experienced less flooding. In these unions, 1,291 surveys were conducted with households consistent with our definitions of treatment and control groups. These households are located in only 64 out of the 111 unions covered in the first round of data collection. Given that only a small number of treated households were surveyed in both rounds of data collection, we treat these two rounds as cross-sectional data.[14] This survey provides additional, although limited, data for our analysis.

Appendix Table 1.11 compares household characteristics across the treatment and control groups in the smaller second round of data collection. Households appear well balanced and we fail to reject equivalent means at the 10% level in all cases except digital wallet use, which is relatively more frequent in the control group.[15]. Moreover, the follow-up sample is broadly representative of households contacted during the first survey. Of 1,291 households in this sample, more than half were included in the initial evaluation.[16] We also confirm that other sources of cash assistance were very limited. In the second survey, only 5% of treated households and 6% of the control group reported receiving cash from a source other than WFP since the flooding began.[17]

Satellite data

We complement survey data with satellite imagery to estimate objective measures of flood timing and severity, given that self-reported flood exposure is a weak proxy for true flood exposure (Guiteras et al., 2015). In collaboration with the UN Office for the Coordination of Humanitarian Affairs's Centre for Humanitarian Data and MapAction, we employ the European Space Agency's Sentinel-1 SAR imagery, providing images with 10-metre spatial resolution. In light of its ability

[14]Of the 1,291 relevant households in the second round of data collection, 93% of 659 control households were contacted in the first survey, but only 32% of the 632 treated households.

[15]61% of control group respondents have used mobile money in the last six months, compared to 49% of the treatment group (p=0.030)

[16]As in the first survey, the average respondent is female, 37–38 years old, and lives in a household of five with two or three dependents. Just over one third completed primary school, one quarter live in the most vulnerable housing structure, and almost all have used a mobile phone in the last week.

[17]Almost all these other cash transfers were almost all received from NGOs and government organisations; 80% of the other cash transfers reported were received within 12 weeks of the flood, i.e., the time of the first survey.

to capture information even in the presence of cloud cover, Sentinel-1 SAR data has been frequently applied to flood mapping, including in Bangladesh (Uddin et al., 2019; Singha et al., 2020). Water bodies are identified from their dark appearance. We use a change detection and thresholding approach to identify flooded areas, following a methodology developed by the UN-SPIDER Knowledge Portal. This methodology entails comparing pixel intensity and the degree of change between before-flooding and after-flooding images for the area of interest, with the baseline constructed by using an average of all images from the previous four years.

Using this approach, we estimate the flooding extent by area covered at the mauza level. Furthermore, as the temporal frequency of the Sentinel-1 imagery can be up to 12 days between images, we estimate flooding extent at daily intervals by fitting the Sentinel-1 data points to a Gaussian function, thereby enabling us to identify the estimated date of flood peak (or maximum extent) for each mauza.[18] We validate these estimates against four different external data sources: (1) river discharge measurements from GloFAS and water level measurements from the Flood Forecast and Warning Centre at four points along the Jamuna River; (2) perceived flooding extent from key informants on the ground in 20 unions; (3) optical satellite imagery from Sentinel-2; and (4) satellite flood extent data from UNOSAT for 21 July. Compared to gauging station data, satellite estimates obscure multiple flood peaks occurring in a short period, but overall trends in flooding are very similar. In unions with key informants, perceived flood trends show general agreement, and in many cases flood magnitudes are also similar. Comparison to optical satellite imagery also shows a high degree of overlap with visible surface water.

1.3.4 Outcome measures

In constructing our outcome measures, we follow a detailed pre-analysis plan, registered in the American Economic Association Registry.[19] We collect all data via phone surveys and therefore our measures are limited by this format. Appendix

[18] When using satellite data, we exclude nine mauzas where the Gaussian model fit of flood dynamics was poor; this accounts for only 0.5% of households.

[19] https://www.socialscienceregistry.org/trials/6576

Tables 1.12 and 1.13 provide more detail on the precise construction of each variable.

We consider three measures of household welfare: (1) child food consumption; (2) adult food consumption; and (3) wellbeing. We focus primarily on food consumption, given that the cash was intended to mitigate the effects of the flood shock on food insecurity. Sacrificing food consumption is one of the most frequent coping mechanisms in response to a negative income shock and one that cash is most likely to affect (Aker et al., 2016; Haushofer and Shapiro, 2016; Jensen et al., 2017; Asfaw et al., 2017). Child food consumption is captured by a dummy variable indicating whether children in the household have consumed at least three meals in the day prior to the survey. The adult food consumption index is a measure of quality, rather than quantity, of nutritional intake. The first component captures the number of days during which expensive protein (meat, fish, or eggs) were consumed by any household member in the seven days prior to the phone survey. The second component is the food consumption score (FCS), a measure of general nutrient intake calculated using the frequency of consumption of different food groups in the week prior to the survey, on a scale of 0–112. We exclude rice from the calculation of this measure, as more than 95% of households reported eating rice every day in the past week.

As the last welfare outcome, we measure wellbeing using a 10-scale Cantril's

ladder of life satisfaction.[20] Flood shocks are likely to be extremely distressing events and the existing evidence on cash transfers has been shown to have large increases on psychological wellbeing in the short- and long-term (Haushofer and Shapiro, 2016). The Cantril's Ladder is widely used in phone surveys through the Gallup World Poll and correlates strongly with other welfare measures, such as income (Deaton, 2008).

To uncover the channels through which welfare improvements may have occurred, we pre-specified five additional variables: (1) pre-emptive actions; (2) asset loss; (3) costly borrowing; (4) remittances; and (5) earning potential. Our focus on measuring pre-emptive actions is motivated by the evidence showing that actions taken before floods can limit asset loss and damage by up to 50%, largely due to effective warning systems (Carsell et al., 2004; Kreibich et al., 2005; Thieken et al., 2007). Like early warning systems, cash was delivered in anticipation of the flood shock to a large proportion of our sample. In the absence of a standard pre-emptive actions measures, we employ a relatively crude index that measure the number of actions taken to prepare for flooding in mid-July before the flood peak, including purchasing food, evacuating, or reinforcing walls, among others.

We construct our asset loss index by combining the number of livestock that died in the two months following the flood peak, the number of asset categories that were lost or damaged due to the flood (out of a list of 15), and the area of

[20] We pre-specified that the wellbeing index would combine with Cantril's ladder of life satisfaction and self-reported hours slept the previous night, as a shorthand proxy for psychological distress given the phone survey medium for data collection. We re-evaluated this after a review of the relevant economics and psychology literature revealed that sleep outcomes are extremely challenging to impact through one-off interventions and may be a poor measure of psychological wellbeing. While lab experiments in high-income countries have shown benefits from temporarily increased sleep duration on mental, psychological, and physical outcomes (Lim and Dinges, 2010; Banks and Dinges, 2007), sleep interventions in field settings have produced limited results at best, especially in low-income contexts. For instance, a targeted three-week sleep intervention providing information, encouragement, and tools to enhance the sleep environment increased nightly sleep by only 27 minutes in urban India, as measured using state-of-the-art technology (Bessone et al., 2020). Moreover, none of the expected effects on economic outcomes materialised as a result of increased sleep duration, including cognition, productivity, decision making, and even wellbeing. Bessone et al. (2020) also highlight the importance of measuring sleep using objective measures, as self-reported measures of sleep often fail to capture true sleep efficiency (Castro et al., 2013; Schokman et al., 2018). Hence, it is highly unlikely that our cash treatment – a small indirect intervention – will impact self-reported hours of sleep, so we exclude this measure from the index. Nonetheless, our results remain robust to correcting for multiple hypothesis testing when we include an additional outcome for self-reported hours of sleep as another independent hypothesis.

cultivated crops lost in decimals (1 decimal $\approx$ 40 sq metres) due to flooding. The costly borrowing index is conditional on borrowing and combines both the amount borrowed and the highest monthly interest rate incurred in the two months following the flood peak. We use a dummy variable for whether a household received remittances to construct the standardised remittance outcome, noting that only 9% of households reported receiving any remittances when we reviewed preparatory data blind to treatment status and amended the pre-analysis plan accordingly. Lastly, the earning potential index is constructed by combining a dummy variable for whether a household did not lose crops from flooding or was able to replant, and the number of paid hours of work per adult in the week prior to the survey.

We standardise all variables for comparability, following Kling et al. (2007) in constructing indices. In particular, we first ensure that all variables are consistently signed within a particular index. We then sum individual response items within each scale, before standardising the summed scale by subtracting the control group mean and dividing by the control standard deviation. Additionally, if there are multiple sub-scales within an index, we also sum the standardised subscales, before re-standardising the final index using the control mean and standard deviation.

Variables for land type (char land, unprotected mainland or protected mainland) are defined at mauza level. We use spatial data to categorise each of the 639 mauzas in our sample by the predominant land type based on their location relative to the braided shape of the Jamuna River and about 800 km of existing flood embankments.

1.4 Empirical strategy

1.4.1 Three empirical specifications

The average treatment effect of cash

We first estimate the average treatment effect of the cash on a variety of outcomes by using the following empirical model:

$$Y_i = \beta_0 + \beta \cdot T_i + \gamma \cdot X_i + \delta \cdot \eta_i + \varepsilon_i \tag{1.1}$$

where Y_i is the outcome of interest for household i and T_i is a dummy variable indicating whether a household received an anticipatory cash transfer from WFP, pooling across all transfer dates. X_i is a vector of pre-specified controls to increase precision in our estimates. These include gender, education level, household size, dependency ratio, house structure, UNFPA/FAO recipient status and land type (flood exposure). We control for the union, as our geographic fixed effect (η_i). ε_i is a mean zero error term. We estimate robust standard errors to correct for heteroskedasticity and cluster at the union level, our sampling frame.

Heterogeneity analysis by timing

In addition to exploring the impact of an anticipatory cash transfer relative to the control group, we examine whether the impact varies by timing and land type. Our second specification assesses whether an earlier timing of the cash transfer relative to the flood shock matter for household outcomes. Within mauzas, households received cash on different dates in July 2020. This variation was owing to the large number of transfers that needed to be conducted in a short period of time and the fact that some households were contacted later and/or given the opportunity to activate bKash accounts. Approximately 50% of our sample received cash on 14 July, with the rest receiving cash on 15, 16 and 30 July. We compare the date of cash transfers to the the local flood peak estimated at the mauza level using satellite data. We estimate the average treatment effect of receiving cash centred on the flood peak date and the marginal welfare effect of receiving the cash one day earlier, as follows:

$$Y_i = \beta_0 + \beta_1 \cdot T_i + \beta_2 \cdot T_i \cdot D_i + \gamma \cdot X_i + \delta \cdot \eta_i + \varepsilon_i \tag{1.2}$$

where T_i is still a dummy variable indicating whether a household i received a cash transfer from WFP and β_1 captures the average treatment effect of the cash transfer centred on the local flood peak date in comparison to the control group. D_i indicates the number of days between the cash transfer date and the local flood peak date, with negative values indicating when households received cash after the flood peak. β_2 captures the marginal effect of receiving cash one day earlier. For every outcome, we test the null hypothesis that an earlier cash transfer has

no impact. We control for mauza fixed effects in this specification to reduce any confounds between flood timing and other spatial characteristics. We also control for the survey date, given that control households were interviewed a few days later than treated households on average. We check that our results are robust to instrumenting the number of days between cash transfer and local flood peak with transfer dates and excluding households which received cash on 30 July. We also use the variation in timing induced by local flood dynamics, while controlling for transfer dates, by analysing results for 14 July recipients only and controlling for union geographic fixed effects.

Heterogeneity analysis by land type

Lastly, we explore whether the impacts of the cash transfers vary for households living on different types of land.[21] Land type captures at least three important dimensions of initial vulnerability to the flood shock: (1) exposure to flooding; (2) socio-economic status; and (3) market isolation. Moreover, land type is easily observed and hence a useful tool for targeting the intervention.

We split the sample into three land types due to their unique features: (1) unprotected char land, which includes low-lying islands along the course of the Jamuna River; (2) unprotected mainland located outside existing flood embankments; and (3) protected mainland. Char lands are formed by silt deposits and the flat, fertile soil makes them attractive for farming. However, they are also extremely vulnerable to flooding and erosion, and they are more isolated from markets and services. More vulnerable transient households typically settle on char lands ('char households'). Unprotected mainlands are located outside flood embankments and are also highly exposed to flooding as a result ('unprotected households'). In contrast, households located on protected mainland – the third land type – are within flood embankments and are less exposed to floods ('protected households'). Roughly a third of the sample is located on each land type, but as noted in Section 1.3.2, there is a greater proportion of treated households relative to control households on char land.

[21] We specified land type as the relevant dimension of heterogeneity in our pre-analysis plan.

To conduct the heterogeneity analysis, we estimate a fully interacted model, using unprotected households as the base category. Building on specification (1), we also include dummies for char households (L_{1i}) and protected households (L_{2i}), and their interaction with treatment status ($T_i * L_{1i}$ and $T_i * L_{2i}$):

$$Y_i = \beta_0 + \beta_1 \cdot T_i + \beta_2 \cdot L_{1i} + \beta_3 \cdot T_i * L_{1i} + \beta_4 \cdot L_{2i} + \beta_5 \cdot T_i * L_{2i} + \gamma \cdot X_i + \delta \cdot \eta_i + \varepsilon_i \quad (1.3)$$

where β_3 captures the additional average treatment effect for char households, and where β_5 captures the additional average treatment effect for protected households, relative to unprotected households.

1.4.2 Approach to inference

In addition to presenting *p*-values from the Wald test, we correct for multiple hypothesis testing across our main outcomes of interest. In particular, we present sharpened *q*-values after correcting for the false discovery rate, following the two-stage procedure developed by Benjamini et al. (2006) and implemented with code by Anderson (2008).

In the Appendix, we also present *p*-values from wild bootstrap inference with clustering at the union level to account for differences in cluster size, which can bias standard cluster inference even with many clusters (Mackinnon and Webb, 2017). There is an uneven distribution of households across the 111 unions, because the WFP was restricted in its targeting efforts by the limited availability of households on the pre-existing beneficiary lists.

1.4.3 Robustness checks

We check that our results are robust to alternative definitions of the control group and model specifications in Section 3.6. First, we show that our results are robust to mobile money use by splitting the sample into households that have previously used a mobile money account and households with a new account (38% of the sample). Treated households were required to have access to an active bKash mobile money account to receive the intervention. All control households reported

having access to an active mobile money account at the time of survey.[22] Although no differences in average mobile money usage are observed between control and treatment households, it is important to ensure that our results are robust to any unobserved or nonlinear differences in mobile money usage that may be present.

For similar reasons, we also check that our results are robust to the exclusion of households that received the cash transfer on 30 July. These households were targeted at a later date, in part because they were given the opportunity to reactivate or set up new bKash accounts. We observe that households that received cash on 30 July were significantly less likely to have used a mobile money account in the six months prior to the survey (38% vs 48% of households receiving cash on 14–16 July; p=0.000).

Similarly, we account for the possibility that households that received cash on 15, 16, 18 and 30 July may have responded to the treatment by reactivating dormant bKash accounts. We present a robustness check where we treat households that received cash after 14 July as a problem of compliance. In other words, they received the treatment when they should not have according to the selection criteria above and thus will be included in the control group.

In addition, we check for robustness to the use of more conservative geographic fixed effects and the exclusion of outliers and households that also received assistance from other UN bodies (UNFPA and FAO). Furthermore, we exclude covariates and control for various measures of flood intensity (estimated at the mauza level from satellite imagery).

Lastly, given the severity of the flooding and the associated asset loss and displacement, we expected low response rates. We sampled a large number of treated and control households as a result, with replacement. However, given that control households were less likely to be contacted by phone, we expected a higher replacement rate for the control group relative to treated households. In total, we contacted 7,819 treated households and reached 6,670, whereas we contacted 4,459 control households and reached 2,460. In Section 1.7.3, we provide several reasons

[22] In the pre-analysis plan, we proposed testing the robustness of our results by excluding control households that did not have access to an active bKash account. However, this turned out to be only 63 households at the time of the survey.

to believe that the patterns of non-response suggest that our treatment effects are likely to be lower bounds, and we show that our main results are robust to the use of Lee bounds (Lee, 2009).

1.5 Results

1.5.1 The anticipatory cash transfer improves welfare

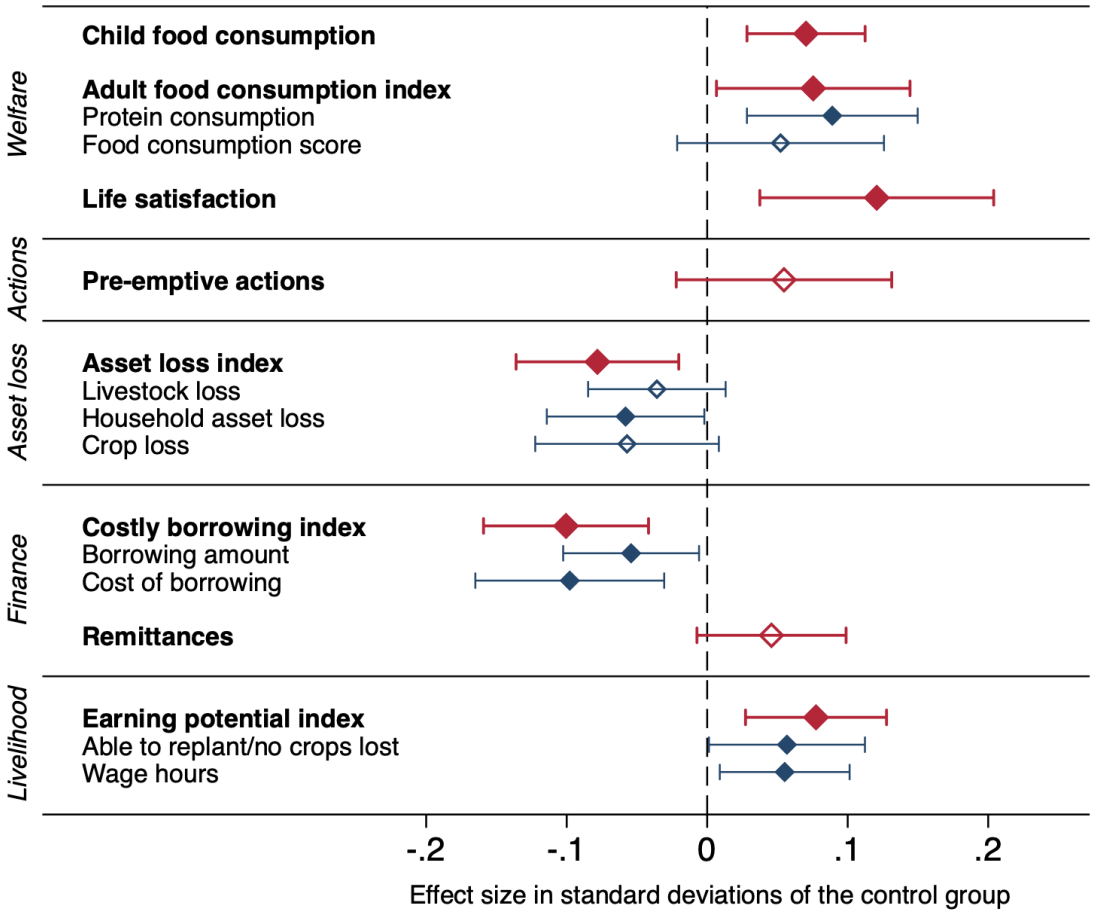
The first question of interest is whether a small, one-off cash transfer succeeds in improving welfare. We find that the anticipatory cash transfer significantly improves welfare: child and adult food consumption and well-being are significantly higher for treated households, even when measured three months after the intervention. Figure 1.3 shows the average effect of the cash transfer on pre-specified, standardised outcomes with 90% confidence intervals. Table 1.3 reports the p-values and sharpened q-values for the standardised treatment effects and the percentage change relative to the control mean.[23] All results reported below remain statistically significant at conventional levels after correcting for the false discovery rate and wild bootstrap inference (see Appendix Table 1.14).

Children in treated households are 0.07 standard deviations or three percentage points more likely (p=0.006) to have consumed three or more meals in the day prior to the survey, relative to a control mean of 80%. The effect remains highly significant across multiple robustness checks (see Section 3.6) and to defining the outcome as the number of meals consumed. The importance of improved child food consumption months after the flood shock is salient, in light of a large body of literature highlighting the negative effects of short-term disruptions in nutritional intake for children during droughts on long-term educational, earnings and health outcomes – even decades later (Chen and Zhou, 2007; Maccini and Yang, 2009; Dercon and Porter, 2014; Dinkelman, 2017; Victora et al., 2021).

Similarly, adult food consumption is 0.075 standard deviations higher (p=0.072) for treated households relative to control households. The treatment

[23]The sharpened q-values are sometimes smaller in size than the unadjusted p-values. This is due to the fact that when there are low p-values for the majority of outcomes implying many true rejections, several false rejections can be tolerated while still maintaining a low false discovery rate.

Figure 1.3: Effect of receiving a cash transfer



Notes: Markers indicate the standardised mean treatment effect of receiving the cash transfer on pre-specified outcomes (red) and sub-indices (blue), with 90% confidence intervals shown. Solid markers indicate significance at the 10% level. Covariates include age, gender, education level, household size, dependency ratio, house structure, UNFPA/FAO recipient status and land type. Union fixed effects are included. Standard errors are clustered at union level.

Table 1.3: Treatment effect of receiving anticipatory cash transfer

	Std. treatment effect	Control mean	Δ	% Δ	*p*-value	*q*-value	N
Children consumed three meals (0/1)	0.070***	0.80	0.03	+3.8%	0.006	0.027	7563
Adult food consumption index	0.075*				0.072	0.046	8951
Days protein products consumed (0–7)	0.089**	2.64	0.18	+6.8%	0.017		8951
Food consumption score (0–112)	0.052	39.21	0.82	+2.1%	0.242		8951
Life satisfaction (0–10)	0.121**	2.08	0.26	+12.5%	0.018	0.028	8941
Number of pre-emptive actions (0–6)	0.055	0.95	0.06	+6.3%	0.237	0.099	8947
Asset loss index	-0.078**				0.027	0.033	8950
Number of livestock died in last two months	-0.036	0.59	-0.04	-6.8%	0.228		8950
Asset categories lost/damaged (0–15)	-0.058*	1.30	-0.07	-5.4%	0.090		8950
Area of cultivated crops lost (decimals)	-0.057	15.58	-1.87	-12.0%	0.150		8950
Costly borrowing index	-0.100***				0.005	0.027	6061
Amount borrowed in last two months (BDT)	-0.054*	8927.52	-564.79	-6.3%	0.066		6061
Highest interest rate (%/month)	-0.098**	4.83	-0.56	-11.6%	0.017		6061
Received remittances (0/1)	0.046	0.08	0.01	+12.5%	0.153	0.072	8950
Earning potential index	0.077**				0.012	0.027	8945
No crops lost/able to replant (0/1)	0.057*	0.65	0.03	+4.6%	0.092		8945
Paid hours of work/adult last week	0.055*	9.08	0.60	+6.6%	0.050		8945

Notes: The first column shows the standardised mean treatment effect for pre-specified outcomes (bold) and sub-indices. The second column shows the control mean, followed by the non-standardised treatment effect (Δ) and percentage change relative to the control mean (% Δ). *p*-values are reported on all outcomes and sub-indices, with standard errors clustered at union level. False discovery rate *q*-values for eight hypotheses are calculated over the main outcomes following the sharpened two-stage procedure of Benjamini et al. (2006). Covariates include age, gender, education level, household size, dependency ratio, house structure, UNFPA/FAO recipient status and land type. Includes union fixed effects. * p<0.10, ** p<0.05, *** p<0.01

effect on the adult food consumption index is largely driven by a 6.8% increase (p=0.017) in the number of days during which more expensive forms of protein (meat, fish, or eggs) were consumed in the week prior to the survey, relative to a control mean of 2.6 days. The persistent effects in food consumption three months after the intervention is consistent with other studies that show longer-term improvements in food security from (albeit much larger) cash transfers (Haushofer and Shapiro, 2016, 2018).

These findings are consistent with the recalled number of days without eating during the flood shock, in data collected for a subset of unions during the second round of phone surveys. Results show that treated households were 10 percentage points less likely to go a day without eating during the flood (p=0.027), relative to 28% of control households. On average, they endured half the number of days without eating (p=0.011), compared to a control mean of 0.62 days per household

(see Section 1.8 and Figure 1.10). These results are consistent with the reported use of the cash transfer: treated households mostly reported spending the cash on food or water (91% of households), followed by medicine or health services (33%), agricultural inputs (26%), loan repayments (12%), clothing (11%), and repairs on home or assets (11%).[24] However, since the modest transfer is approximately equivalent to only 2 weeks of food expenditure, finding effects after 10-12 weeks points to other possible mechanisms discussed in Section 1.5.2. Results strongly suggest that the treatment effects that we estimate on food consumption three months after the intervention are a lower bound compared to the size of the treatment effects during the shock itself.

Control households report an extraordinarily low level of wellbeing at the time of the survey, with a mean of 2 on a 10-item Cantril's ladder of life satisfaction, evidencing the distressing nature of flood shocks. We find that the cash transfer helps mitigate the flood's impact on general wellbeing: at the time of the survey, average life satisfaction for the treated households was 0.12 standard deviations or 12.5% higher (p=0.018) than in the control group. However, we note that even with this increase, this is an extremely low level of well-being.

1.5.2 Mechanisms through which welfare is improved

The improvements to welfare suggest that the cash transfer served to smooth a negative income shock created by disruptions to income generation activities and the destruction to private assets. In this section, we show that the cash transfer also significantly mitigated asset loss and costly borrowing, even three months after the flood shock. These effects could arise due to actions taken either before, during or after the flood shock. While collecting data via a phone survey two to three months after the intervention limits the granularity with which we can disentangle the timing of behaviours, our results suggest that anticipatory cash bolstered existing coping strategies, rather than expanding the choice set of strategies.

[24]Consistent with these results, the WFP Monitoring and Evaluation report found that treated households spent approximately 67% of the cash (approximately BDT 2000) on meeting food needs, followed by recovery (46%), preparedness actions (28%) and health (23%)(WFP, 2020).

Just over half (53%) of control households took action prior to the flood to reduce its impacts, with a control mean of one action per household. Treated households did not take significantly more actions to prepare for flooding (see Figure 1.3), nor did they pursue different strategies, with the exception of evacuation. We acknowledge that our pre-specified measure of preventative actions is blunt. Exploring further, we observe that treated households were three percentage points more likely to evacuate both household members (p=0.032; control mean of 29%) and livestock (p=0.026; control mean of 17%), after controlling for covariates and union fixed effects. At the margin, we find that treated households were four percentage points more likely (p=0.051) to take some action in preparation for the flood peak.[25]

Although little difference in preventative actions was recorded, we find that the cash transfer was effective in mitigating asset loss during the shock and in the two months after. The cash transfer decreased asset loss and damage by 0.078 standard deviations (p=0.027) relative to the control group. The sub-indices of livestock, household items and crop loss are themselves only marginally significance at conventional levels, but in combination, they produce an asset loss index that is statistically significant in the treatment group at the 5% level. We find that treated households were 2.3 percentage points less likely (p=0.097) to lose small livestock after flooding and 3.2 percentage points less likely (p=0.082) to lose any poultry during flooding, controlling for covariates and union fixed effects.[26] Lower mortality rates of small livestock and poultry (important productive assets) may explain why households benefited from higher food consumption three months after the intervention, especially on the protein-specific sub-index. Our findings are consistent with other literature showing that evacuation in advance of flooding can substantially reduce damages elsewhere by enabling households to remove pos-

[25]The treatment effects for evacuation and taking any action are reported in Appendix Table 1.15.

[26]The results for livestock mortality are reported in Appendix Table 1.15. Treated households own 17.2% more (p=0.001) small livestock (goats, sheep and pigs) at the time of the survey relative to a control mean of 0.54 animals per household, controlling for covariates and union fixed effects. We do not identify this difference in livestock ownership as a result of receiving cash, since we cannot control for baseline livestock assets. However, higher livestock ownership in the treatment group at baseline would only reinforce the significance of lower livestock mortality that we find among recipients of the cash transfer. Unsurprisingly, we find no differences for larger livestock (cows, calves, and buffalo), which are owned by far fewer households.

sessions, evacuate poultry and farm animals, move agricultural equipment, seeds and partially harvest crops (Carsell et al., 2004; Subbiah et al., 2008).

Households have high liquidity needs during and in the months following the flood shock. Almost three in four control households borrowed money between the flood and survey, borrowing an average amount of BDT 8,928 (approximately $105). Notably, almost a third of the sample borrowed more than BDT 10,000 (approximately $118), double the amount of the cash transfer. However, conditional on borrowing, those receiving the cash transfer borrowed less and more cheaply than control households. The pre-specified costly borrowing index was 0.1 standard deviations lower (p=0.005) on average for treated households relative to the control group. This result remains highly significant correcting for multiple hypothesis testing and across robustness checks (see Section 3.6). On the extensive margin, treated households were also 2.4 percentage points less likely (p=0.082) to borrow relative to a control mean of 70% (Appendix Table 1.15).[27] Conditional on borrowing, treated households borrowed lower amounts at cheaper interest rates. We find they borrowed BDT 565 (approximately $7) less on average in the two months after the onset of the flood if they borrowed at all. Treated borrowers accessed credit at an interest rate that was 0.57 percentage points lower than the control mean of 4.83% per month, and were 10% more likely (p=0.008) to borrow at a zero interest rate (from informal sources), reducing the burden of future indebtedness. At the same time, treated households also reported borrowing food or receiving help from others on 14% fewer days ($p<0.001$) in the week before the survey relative to control households.

There is no evidence that the cash transfer crowds out remittances on average. Treated households were 12.5% more likely to report receiving remittances on average (a one percentage point increase relative to a control mean of 8%), although this result is not statistically significant with a p-value of 0.155. The absence of crowding-out effects assuages fears that the cash transfer could replace informal

[27]The WFP Monitoring and Evaluation report found that loans were largely obtained to meet food needs during the crisis period (40% of households), providing a direct link between the role of the cash transfer, reduction in food insecurity and costly borrowing (WFP, 2020).

support networks (Cox, 1987; Cox and Jakubson, 1995; Jensen, 2003; Juarez, 2009).[28]

Thus far, the results point to modest increases in food consumption and well-being, and lower levels of asset loss and costly borrowing. We also explore whether households receiving the cash transfer appear to be recovering faster. We find that treated households reported a 0.077 standard deviation increase (p=0.012) in the pre-specified earning potential index, relative to the control households. Treated households are 3.9 percentage points more likely (p=0.001) to have a household member working for remuneration (see Appendix Table 1.15), and the average hours worked for remuneration per adult were 6.6% higher (p=0.050), equivalent to an additional 36 minutes per week. The cash transfer appears to be playing a small but not insignificant role in post-flood recovery, helping to restore livelihoods and the capacity to cope with future shocks.

We cannot determine precisely what proportion of the impact of the transfer arises from the injection of liquidity at a critical juncture compared to the early warning (information) aspect of the transfers. However, evidence strongly points to the primary importance of liquidity. First, households receiving cash were no more likely to report receiving early warning. Second, control households that did evacuate (they had warning but not extra liquidity from cash transfer) were not unambiguously better off than those that did not. Third, control households that bought food to prepare (having liquidity to do so) had significantly higher welfare than those that did not. Lastly, the fact that food is the most common expenditure (91% of households) and the type of pre-emptive actions were mostly similar across treated and control households points to liquidity as the dominant factor in households' ability to cope.

Overall, the treatment effects of the cash transfer are statistically significant, but small in magnitude. The size of these effects is unsurprising, considering that the amount of the one-off cash transfer was very modest, approximately two weeks' worth of food expenditure during floods lasting more than one month. The

[28]For ultra-poor households in our sample, zero-interest borrowing from informal networks seems to be more important than remittances to cope with the shock, whereas fewer than one in 10 households reported receiving remittances. We found that treated households had greater access to these informal loans.

level of borrowing sheds light on the liquidity needs of households during this period. Instead, what should be striking is that we do measure significant effects across such a wide range of outcomes three months after the intervention, highlighting the effectiveness of a timely transfer.

1.6 Heterogeneity analysis

1.6.1 Timing of cash transfers

Although we are unable to directly compare anticipatory action to a traditional response, the welfare benefits we report were all recorded before the arrival of a traditional response. Without cash, comparable ultra-poor households experienced lower food consumption (both adults and children), lower wellbeing, higher asset loss, higher borrowing, and less employment in the period before a traditional response. Some of these deficits, even when temporary, are known to have permanent scarring effects – for example, child undernutrition (Victora et al., 2021) – suggesting that there are significant advantages from acting early. The small cash transfer may set households on a different trajectory by easing liquidity constraints and permitting better decisions at a critical juncture.

To further investigate the importance of speed in implementation, we explore whether small changes in the timing of the cash transfer relative to flooding matter for household welfare. Qualitative evidence suggests that the cash may have been received too late on average for households to take an alternative set of actions in preparation of the flood, but instead strengthened the use of existing strategies (WFP, 2020; Adams et al., 2021).[29] To examine this, we calculate the number of days between receiving a transfer and peak local flooding in 630 mauzas (approximately village level and surrounding land) based on high resolution satellite data. We then investigate the relationship between cash transfer timing and welfare effects using linear regression and non-parametric methods to estimate the marginal

[29] In particular, households reported that an earlier cash transfer would have been used to stock up on long-life food, drinking water, fuel and medicine; reinforce the roof or walls; take household items to higher ground; and purchase livestock (Adams et al., 2021).

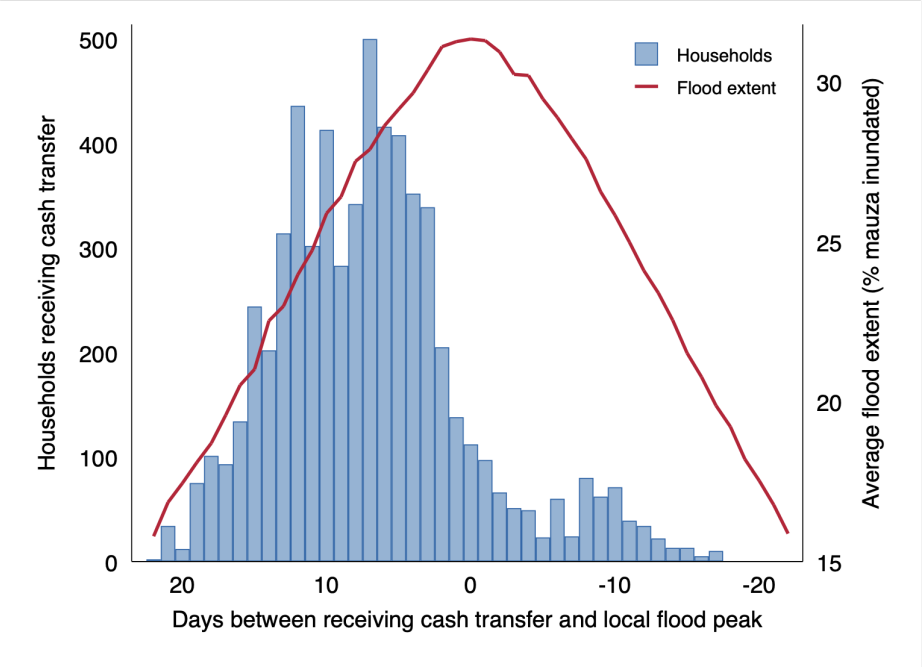
effect of receiving cash one day earlier.[30] Figure 1.4 shows the number of households receiving cash by the number of days between transfer and local flood peak, and the average flood extent corresponding to this timing (negative days indicate receiving cash after the local flood peak). Most households received the cash before the local flood peak. However, dangerous flood levels persisted for several weeks (see also Figure 1.1), so the flood peak date should not be considered a discrete event marking flood shock impacts. Rather, the flood peak provides a useful reference for measuring the timing of the cash transfer relative to flood dynamics over a large geographic region.

We find evidence to suggest speed is welfare improving, even when considering a cash transfer that is already much earlier than the typical response. Table 1.4 reports the average effect of receiving cash centred on the flood peak date and the marginal effect of receiving the cash one day earlier, controlling for mauza fixed effects and the date of the survey. First, we observe that the average treatment effects centred on the flood peak date are broadly consistent with our main results in Section 3.5. When we remove the effect of small changes in timing, the estimated coefficients on the asset loss index and the earning potential index lose significance, but we now estimate a marginally more significant effect on remittances relative to the control group.

Second, we observe that an earlier cash transfer has a stronger effect on adult food consumption. Receiving cash one day earlier increases the adult food consumption index by 0.005 standard deviation ($p<0.05$) when measured 10 to 12 weeks after the intervention. This result does not survive correcting for multiple hypothesis testing using FDR q-values. If we treat this result as only suggestive

[30] In our pre-analysis plan, we specified that we would analyse the effect of receiving a cash transfer one day earlier using only the transfer dates. We now take advantage of the additional granularity offered by the satellite data and control for the date of the survey, as many control households were surveyed a few days after the treated households. We also pre-specified an alternate model comparing households that received the cash on 14, 15 and 16 July to households that received the transfer on 30 July due to administrative delays. The reason for splitting the sample into these two groups, was that the centrally located Bahadurabad gauging station recorded a flood peak on 17 July. While results are comparable, we are concerned that households that received the cash transfer on 30 July may differ in terms of unobservable characteristics, as some of these households were given the opportunity to reactivate or set up new bKash accounts. Results comparing households receiving cash on 15-16 July to those receiving on 30 July are reported in Appendix Table 1.16.

Figure 1.4: Number of households receiving cash relative to the local flood peak date



Notes: The number of households receiving the transfer is shown on the left-hand axis for each day relative to the date of local flood peak (zero). The average extent of flooding in mauzas with treated households is shown on the right-hand axis. Local flood peak dates are derived from a Gaussian model of maximum flood extent at mauza level estimated using satellite imagery. Observations are excluded where errors were identified in Gaussian model fit, and outlying dates are excluded by trimming at the 1% and 99% level, which reduces the size of the treatment group with a defined transfer date from 6,285 to 6,176 households.

evidence, linear extrapolation would indicate that the average effect of cash dissipates entirely if the cash transfer is not received within 12 days of the local flood peak. This 'back-of-the-envelope' calculation implies that potential welfare gains from emergency cash transfers are forfeited by delaying assistance. There are no other significant welfare improvements from receiving cash a day earlier at the time of the survey. Among mechanistic outcomes, households receiving the cash transfer one day earlier also had a 0.009 standard deviation higher costly borrowing index ($p<0.01$).[31] This result does survive correcting for multiple hypothesis testing. This result could partly be explained by the timing of the later cash transfers relative to an important religious fesitval: cash received after peak flooding

[31] Among secondary outcomes, households receiving an earlier transfer are more likely to evacuate, but less likely to replant lost crops (Appendix Table 1.17).

on 30 July may have substituted borrowing during Eid al-Adha 2020 commencing on 31 July, when purchasing gifts and sacrificing livestock are a norm, whereas cash received earlier had already been spent. Overall, these results suggest that an earlier cash transfer may improve welfare, but a greater variation in timing is required in order to disentangle these effects.

Table 1.4: Treatment effect of receiving anticipatory cash transfer one day earlier

	(1) Child food consumption	(2) Adult food consumption index	(3) Life satis-faction	(4) Pre-emptive actions	(5) Asset loss index	(6) Costly borrowing index	(7) Remit-tances	(8) Earning potential index
Transfer	0.084**	0.098***	0.099**	0.052	-0.055	-0.128***	0.061	0.041
	(0.037)	(0.037)	(0.043)	(0.041)	(0.037)	(0.039)	(0.042)	(0.035)
Transfer × days before flood peak	0.003	0.005**	-0.001	0.003	0.002	0.009***	-0.002	-0.001
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
p-value: Transfer × days	0.264	0.048	0.827	0.280	0.408	0.000	0.622	0.596
q-value: Transfer × days	0.726	0.200	1.000	0.726	0.958	0.004	1.000	1.000
Controls	✓	✓	✓	✓	✓	✓	✓	✓
Mauza fixed effects	✓	✓	✓	✓	✓	✓	✓	✓
N	7263	8631	8620	8627	8630	5796	8630	8625
R^2	0.10	0.15	0.17	0.18	0.20	0.17	0.10	0.18

Notes: The standardised mean treatment effect, centred on the date of the flood peak, is shown in the first row. The second row shows the marginal treatment effect of receiving the transfer a day earlier relative to the flood peak. False discovery rate *q*-values are calculated over the marginal treatment effects for the eight outcomes following the sharpened two-stage procedure of Benjamini et al. (2006). In addition to the covariates used in the main analysis, we also control for interview date. Mauza fixed effects are included. Standard errors are clustered at mauza level in parentheses. * p<0.10, ** p<0.05, *** p<0.01

To address the concern that variation in the timing of floods across space might be confounded with other spatial differences, we show equivalent results when instrumenting on whether households received cash on the exogenously-determined transfer dates (14, 15, or 16 July) using union fixed effects (see Appendix Table 1.18). This analysis thus focuses on variation driven by the date of the payment. However, these results are not robust to dropping households receiving cash on 30 July, which are driving most of the variation observed above (see Appendix Table 1.19). It is unclear whether results can be attributed fully to timing or may be due to the fact that households receiving cash on 30 July were slightly different, given their less frequent use of mobile money accounts.

We also exploit variation in timing arising largely due to local flood dynamics and hold constant the transfer dates by presenting results for 14 July recipients only. 14 July recipients make up more than half of the treatment group. Table 1.5

shows these timing results using union fixed effects, given that flood data is reported at the mauza level - the smaller administrative unit. The use of union fixed effects may capture other spatial dynamics, so we would treat these results as suggestive evidence that speed matters for the impact of anticipatory cash transfers. As before, we control for the covariates used in the main analysis, including land type, and the interview date. We find that the positive child food consumption result is almost entirely determined by earlier cash transfers. Receiving cash one day earlier increases the child food consumption index by 0.012 standard deviation ($p<0.05$). However, this result does not survive correcting for testing multiple hypotheses using FDR q-values.

Table 1.5: Timing effects using 14 July treatment group only

	(1) Child food consumption	(2) Adult food consumption index	(3) Life satis-faction	(4) Pre-emptive actions	(5) Asset loss index	(6) Costly borrowing index	(7) Remit-tances	(8) Earning potential index
Transfer	0.014	0.080	-0.003	0.051	-0.055	-0.046	0.065	0.004
	(0.065)	(0.075)	(0.080)	(0.089)	(0.064)	(0.079)	(0.077)	(0.055)
Transfer × days	0.012**	0.007	0.009	0.000	-0.003	0.001	0.001	0.003
before flood peak	(0.006)	(0.007)	(0.006)	(0.008)	(0.006)	(0.007)	(0.007)	(0.005)
p-value: Transfer × days	0.037	0.352	0.157	0.964	0.664	0.836	0.909	0.543
q-value: Transfer × days	0.424	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Controls	✓	✓	✓	✓	✓	✓	✓	✓
Union fixed effects	✓	✓	✓	✓	✓	✓	✓	✓
N	4691	5537	5532	5537	5536	3755	5536	5535
R^2	0.05	0.11	0.11	0.11	0.15	0.11	0.05	0.12

Notes: We restrict the treatment group to 14 July recipients only. The standardised mean treatment effect, centred on the date of the flood peak, is shown in the first row. The second row shows the marginal treatment effect of receiving the transfer a day earlier relative to the flood peak. False discovery rate q-values are calculated over the marginal treatment effects for the eight outcomes following the sharpened two-stage procedure of Benjamini et al. (2006). In addition to the covariates used in the main analysis, we also control for interview date. Union fixed effects are included. Standard errors are clustered at union level in parentheses. * p<0.10, ** p<0.05, *** p<0.01

We further investigate the relationship between the timing of the cash transfer and welfare effects using non-parametric methods. As illustrated by Figure 1.5, we estimate the treatment effect for households receiving cash within a 3-day window and use a standard non-parametric method to trace the treatment effect across the flood timeline, where day 0 indicates the flood peak. We restrict our analysis to a 21-day window (15 days prior to the flood peak and 5 days after the flood peak) and show the number of households receiving cash on each day for reference. In addition to controlling for fixed household characteristics, mobile wallet use and

union effects as in the main specification, we also control for interview date, land type, flood magnitude and flood duration.

We find further evidence that a faster response relative to the evolution of the flood is welfare improving. The non-parametric analysis suggests that the anticipatory cash transfer is most effective in boosting child and adult food consumption when received at least seven days prior to the local flood peak. Welfare effects tend to dissipate for each day that transfers are delayed in the period before peak flooding. Results are robust to varying the bandwidth and a range of alternative methods.[32]

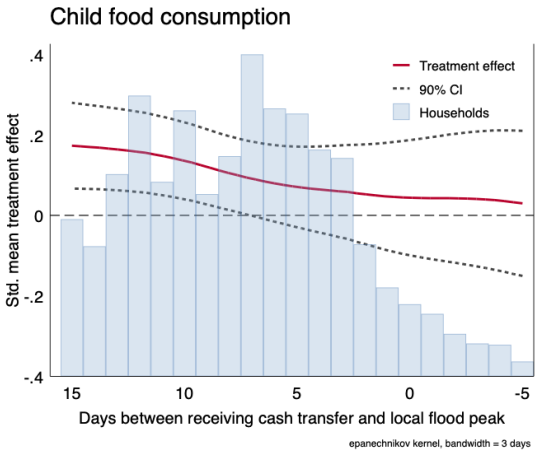
1.6.2 Exposure to flooding

We explore whether the impacts of the cash transfers are different for households living on different types of land. As noted above, land type captures at least three important dimensions of vulnerability in flood affected areas: (1) exposure to flooding; (2) socio-economic status; and (3) market isolation. Overall, we find the strongest treatment effects for households living on land that is more exposed to the flood and more vulnerable, both in terms of socio-economic status and isolation. In contrast, the effects of cash for more protected households are smaller on almost all outcomes of interest.

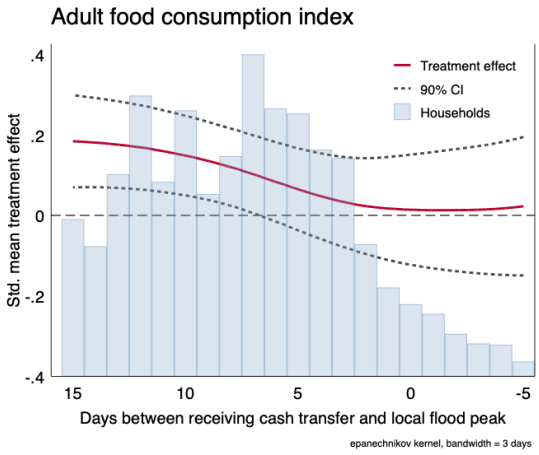
Flood intensity data estimated from satellite imagery confirms that flood waters covered a greater number of mauzas categorised as unprotected or char and for a longer duration than was the case for protected mauzas. Survey data that shows households on protected mainland have a lower dependency ratio, are smaller in size, and are less likely to live in vulnerable housing structures made from raw materials. The distinction between char households and those on unprotected mainland is not as clear on these observable characteristics. Table 1.20 compares the mean value of time invariant characteristics, flood intensity measures and treatment status across land types.

[32]Results are robust to restricting the sample to: (i) similar transfer dates (14-16 July) by excluding 30 July transfers; (ii) more comparable flood experiences by excluding households on protected land, i.e. behind an embankment; and (iii) more similar mobile money users by excluding new subscribers.

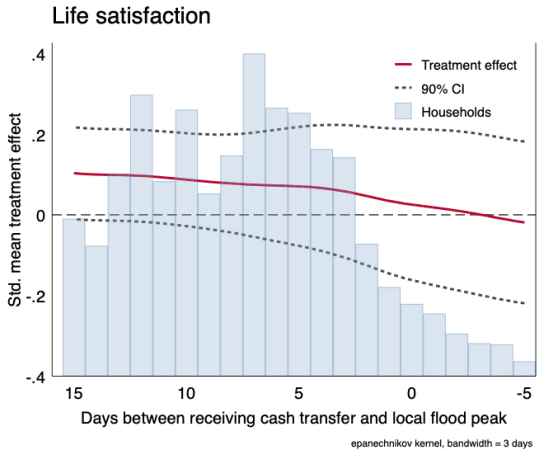
Figure 1.5: Non-parametric analysis on the timing of the cash transfer



(a) Child food consumption



(b) Adult food consumption



(c) Life satisfaction

Notes: The figure illustrates the non-parametric analysis across our three primary outcomes of interest: child food consumption, adult food consumption and life satisfaction. We employ an Epanechnikov kernel and a bandwidth size of three days. We control for fixed household characteristics, mobile wallet use and union effects as in the main specification. We also control for interview date, and mauza-level land type, flood magnitude and flood duration. Standard errors are clustered at union level. 90% confidence intervals are shown.

Table 1.6 presents regression results; here, we interact treatment status by land type and control for land type dummies. Households on unprotected land form the base category. Across many primary outcomes, the base results are consistent with those presented in Section 1.5.1: unprotected households receiving transfers had higher child and adult food consumption, higher life satisfaction, and lower asset loss, relative to their control group. However, there was no impact for these households on costly borrowing, remittances or earning potential.

Table 1.6: Treatment effect of receiving anticipatory cash transfer by land type

	(1) Child food consumption	(2) Adult food consumption index	(3) Life satis-faction	(4) Pre-emptive actions	(5) Asset loss index	(6) Costly borrowing index	(7) Remit-tances	(8) Earning potential index
Transfer	0.092*	0.154**	0.214***	0.078	-0.152**	-0.014	0.053	0.054
	(0.048)	(0.067)	(0.067)	(0.065)	(0.068)	(0.047)	(0.045)	(0.054)
Char	-0.038	-0.036	0.084	-0.018	-0.039	0.072	-0.020	-0.121
	(0.083)	(0.071)	(0.085)	(0.112)	(0.089)	(0.066)	(0.083)	(0.077)
Transfer × char	-0.018	-0.155**	-0.223**	-0.004	0.088	-0.081	0.032	0.059
	(0.074)	(0.075)	(0.094)	(0.093)	(0.092)	(0.062)	(0.066)	(0.074)
Protected	0.024	-0.031	0.050	0.018	-0.169*	0.189***	0.117	0.016
	(0.085)	(0.083)	(0.084)	(0.112)	(0.087)	(0.063)	(0.096)	(0.077)
Transfer × protected	-0.046	-0.065	-0.034	-0.067	0.129	-0.175**	-0.060	0.006
	(0.063)	(0.094)	(0.108)	(0.110)	(0.079)	(0.078)	(0.070)	(0.080)
Treat effect: Char	0.074	-0.002	-0.009	0.073	-0.064	-0.094**	0.085*	0.113**
Treat effect: Protected	0.045	0.089	0.180**	0.011	-0.023	-0.189***	-0.007	0.060
p-value: Char=Prot.	0.671	0.268	0.018	0.539	0.549	0.197	0.167	0.439
F-test *p*-value:								
Control $\Delta = 0$	0.729	0.864	0.601	0.949	0.107	0.011	0.233	0.234
Treatment effect $\Delta = 0$	0.754	0.107	0.012	0.794	0.264	0.082	0.380	0.634
Controls	✓	✓	✓	✓	✓	✓	✓	✓
Union fixed effects	✓	✓	✓	✓	✓	✓	✓	✓
N	7563	8951	8941	8947	8950	6061	8950	8945
R^2	0.04	0.09	0.10	0.09	0.13	0.09	0.04	0.11

Notes: The baseline is for households on unprotected land. The standardised mean treatment effect of receiving the transfer interacted with land type are shown, with standard errors clustered at union level in parentheses. The second panel shows the treatment effect of receiving cash on for char households and protected households (relative to their corresponding control household), and *p*-values from a test of equivalence. The third panel shows *p*-values from an F-test of equivalent outcomes over the three land types for control households and treatment households. Covariates include age, gender, education level, household size, dependency ratio, house structure, UNFPA/FAO recipient status and land type. Union fixed effects are included. * p<0.10, ** p<0.05, *** p<0.01

We then compare the treatment effects for char versus unprotected households. Relative to unprotected households, char households that received the transfer have significantly lower adult food consumption and life satisfaction, to the extent that there is no treatment effect of the transfer on these outcomes for char households (p=0.041 and p=0.020 respectively). We cannot rule out important effects of cash on adult food security or welfare during the flood, but that

any impacts on these dimensions of welfare did not persist 10 to 12 weeks after flooding. We cannot distinguish between impacts on char households and unprotected households for other primary outcomes. Compared to the control group, the cash transfer significantly decreases the amount and the cost of borrowing for char households, increases the probability of receiving remittances, and improves earning potential (p=0.036, p=0.069 and p=0.020 respectively). Char households are 13% more likely to have a household member working for remuneration ($p<0.001$) if they received the transfer relative to a control household on char land. The treatment effect of cash on earning potential is significantly higher for char households than for protected or unprotected households (p=0.028; p=0.022). In the control group, char households are significantly less likely to work for a wage than control households on other land types, suggestive of lower market integration and barriers to labour force participation on char land that the transfer may help overcome (results reported in Appendix Table 1.21).

The effect of cash on protected households is smaller on almost all outcomes relative to unprotected households, but none of these differences is statistically significant, with the exception of lower costly borrowing. This suggests the cash transfer had fewer welfare impacts on less poor and less exposed households, with implications for targeting limited resources.

1.7 Robustness to model specification and sample choice

Our results are robust to mobile money use and a wide range of alternative alternative model specifications. In Section 1.7.1, we check for robustness to different types of mobile money users, given that this plays a role in how our treatment group is defined. In Section 1.7.2, we check for robustness to alternative use of fixed effects, sub-samples and winsorising. Lastly, in Section 1.7.3, we correct for differential non-response.

1.7.1 Robustness to mobile money use

The cash transfer was delivered via mobile money accounts. Although no differences in average mobile money usage are observed between control and treatment households, it is important to ensure that our results are robust to any unobserved or nonlinear differences in mobile money usage between treatment and control households. Therefore, we check whether our results are robust to mobile money use by splitting the sample into households that have previously used a mobile wallet and households with a new account (38% of the sample).[33] There is a slightly higher proportion of treated households in the category that reported recently opening a mobile money account, which is unsurprising given the incentive for treated households to set up new mobile money accounts. Figure 1.6 shows the standardised treatment effects for each sub-sample; we test for differences in the treatment coefficient across models by using the seemingly unrelated regression test (see Appendix Table 1.23).

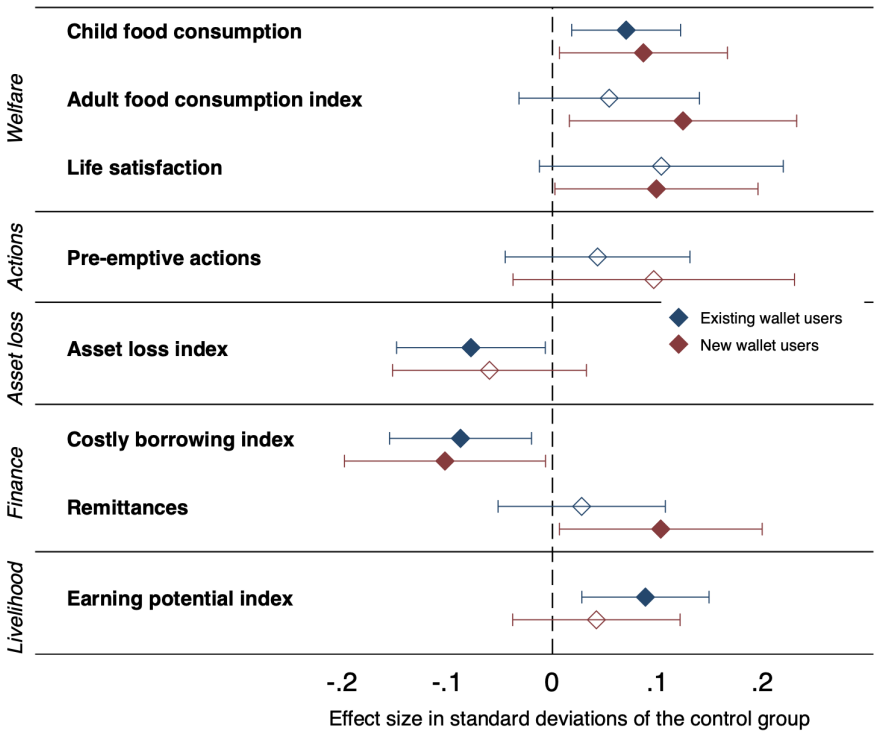
Overall, our results remain robust to splitting the sample by mobile money use. Statistical power is lower when our sample is split, and some of the results are no longer significant. Treatment effects for new wallet users remain significant for four of the six outcome variables that were significant for the full sample. For existing wallet users, four results also remain significant. We find no significant difference in treatment effect between these two groups on any outcome. In addition, we find a positive effect on remittances among new wallet users (p=0.079).[34]

We also test the robustness of results to excluding households that signed up to bKash or reactivated an old account after they were contacted by WFP. Most of these households received their transfers on July 30, so this is done by excluding

[33] We also check robustness by splitting the sample into three groups: (1) frequent wallet users; (2) infrequent wallet users; and (3) new wallet users. Results are reported in Appendix Table 1.22. The existing wallet user sub-sample is split at the median frequency of wallet use, so frequent wallet users are defined as using mobile money two months prior, and infrequent wallet users as using mobile money more than two months prior. Treated households were asked when they had last used their mobile money accounts to conduct a financial transaction prior to receiving the WFP transfer, whereas control households were asked to report when they last used their mobile money accounts prior to the survey.

[34] When we split the sample three ways, the effect on remittances is negative for frequent wallet users, this is suggestive of crowding-out effects among this sub-sample, and significantly different from both infrequent and new wallet user groups. See Appendix Table 1.22.

Figure 1.6: Standardised treatment effects on sub-samples by wallet-use



Notes: Markers indicate the standardised mean treatment effect of receiving the cash transfer on pre-specified outcomes in two independent models for existing wallet users (navy) and new wallet users (maroon) respectively. 90% confidence intervals are shown. Solid markers indicate significance at the 10% level. Covariates include age, gender, education level, household size, dependency ratio, house structure, UNFPA/FAO recipient status and land type. Union fixed effects are included. Standard errors are clustered at union level.

all households that received their transfer on that date. The results are presented in Appendix Table 1.24. Treatment effects are significant at the 10% level for seven of the eight primary outcomes, including all six that were statistically significant for the full sample.

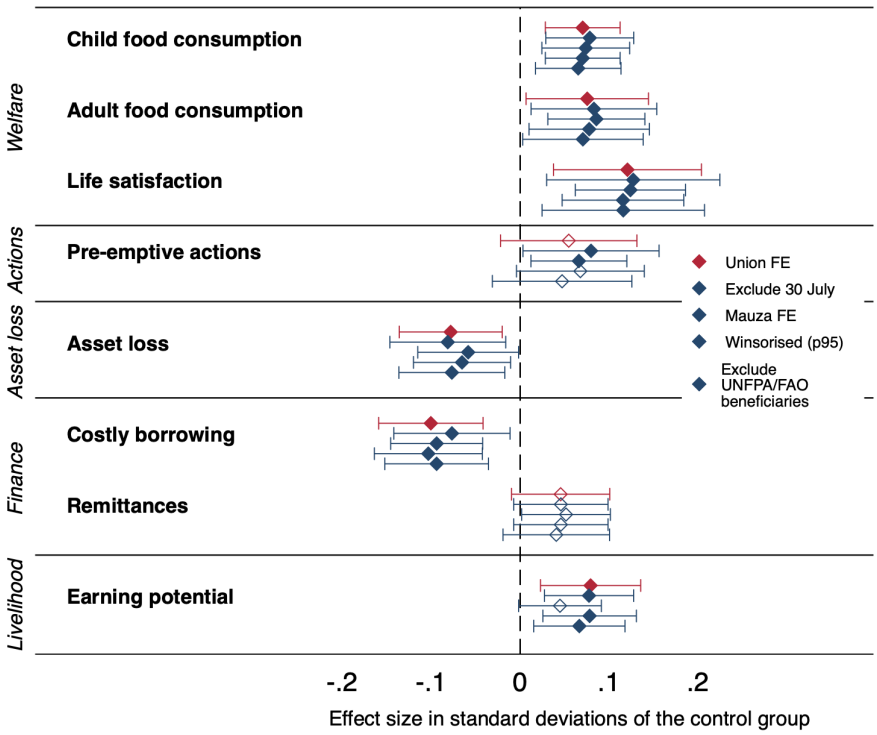
1.7.2 Robustness to alternative model specifications

We check for robustness to a range of alternative model specifications. Figure 1.7 illustrates standardised treatment effects using 90% confidence intervals across five different model specifications, with further details provided in Appendix Table 1.24. Our preferred model specification using union fixed effects and controlling for covariates is shown as the first set of standardised treatment effects for all outcomes. Our results are robust to excluding households receiving a later transfer on July 30 (specification 2), using more conservative fixed effects (specification 3), winsorising outcomes at the 95th percentile (specification 4) and excluding the sub-sample of households that reported receiving assistance from UNFPA and FAO during the same flood period (specification 5). Our results are also robust to excluding covariates, controlling for the survey date, and controlling for various measures of flood intensity estimated at the mauza level from satellite imagery (see Appendix Tables 1.25, 1.26 and 1.27).[35] The latter suggests that the geographic fixed effect in our main specification sufficiently controls for the variation in flood shock across space.

We account for the possibility that households that received cash on 15, 16, 18 and 30 July may have responded to the treatment by reactivating dormant bKash accounts. We present a robustness check where we treat households that received cash after 14 July as a problem of compliance. In other words, they received the treatment when they should not have been treated according to the selection criteria above and thus will be included in the control group. With the caveat that this approach introduces substantial noise given that 14 July recipients only make up half of the treated sample, we present intent-to-treat estimates in Table 1.7. 14 July recipients form the treatment group and the later cash recipients form part of the control group, together with the original control households. We are reassured that

[35]Only the asset loss index loses significance when controlling for survey date.

Figure 1.7: Robustness checks to alternative model specifications



Notes: Markers indicate the standardised mean treatment effect of receiving the cash transfer on pre-specified outcomes for the main specification (red) and alternatives (blue), with 90% confidence intervals shown. Solid markers indicate significance at the 10% level. Covariates include age, gender, education level, household size, dependency ratio, house structure, UNFPA/FAO recipient status and land type. Union fixed effects are included, except when mauza fixed effects are indicated. Standard errors are clustered at union level.

most results follow a similar direction as in the main analysis, with the exception of life satisfaction and costly borrowing. However, we find much smaller effects on all our outcomes. The only effect that is statistically significant is the 4.9% increase in the adult food consumption index relative to the control mean ($p = 0.085$).

Table 1.7: Robustness check: Intent-to-treat estimates using 14 July only

	Std. treatment effect	Control mean	Δ	% Δ	*p*-value	*q*-value	N
Children consumed three meals (0/1)	0.038	0.81	0.01	+1.2%	0.174	0.800	7703
Adult food consumption index	0.052*				0.085	0.800	9121
Days protein products consumed (0–7)	0.064**	2.65	0.13	+4.9%	0.025		9121
Food consumption score (0–112)	0.033	39.38	0.52	+1.3%	0.276		9121
Life satisfaction (0–10)	-0.008	2.30	-0.02	-0.9%	0.809	0.800	9110
Number of pre-emptive actions (0–6)	0.039	0.94	0.04	+4.3%	0.358	0.800	9117
Asset loss index	-0.028				0.389	0.800	9120
Number of livestock died in last two months	-0.030	0.57	-0.03	-5.3%	0.198		9120
Asset categories lost/damaged (0–15)	0.002	1.20	0.00	0.0	0.959		9120
Area of cultivated crops lost (decimals)	-0.026	15.12	-0.85	-5.6%	0.356		9120
Costly borrowing index	0.048				0.221	0.800	6169
Amount borrowed in last two months (BDT)	0.067*	8261.29	649.96	+7.9%	0.057		6169
Highest interest rate (%/month)	0.005	4.34	0.03	+0.7%	0.903		6169
Received remittances (0/1)	0.029	0.09	0.01	+11.1%	0.284	0.800	9120
Earning potential index	0.021				0.371	0.800	9114
No crops lost/able to replant (0/1)	0.008	0.68	0.00	0.0	0.776		9114
Paid hours of work/adult last week	0.022	9.04	0.24	+2.7%	0.393		9114

Notes: The first column shows an intent-to-treat effect for pre-specified outcomes (bold) and sub-indices. The second column shows the control mean, followed by the non-standardised treatment effect (Δ) and percentage change relative to the control mean (% Δ). *p*-values are reported on all outcomes and sub-indices, with standard errors clustered at union level. False discovery rate *q*-values for eight hypotheses are calculated over the main outcomes following the sharpened two-stage procedure of Benjamini et al. (2006). Covariates include age, gender, education level, household size, dependency ratio, house structure, UNFPA/FAO recipient status and land type. Includes union fixed effects. * p<0.10, ** p<0.05, *** p<0.01

1.7.3 Correcting for differential non-response rates

In this section, we show that our results are unlikely to be driven by differential non-response rates. We identify three types of differential non-response comparing treated to control households: (1) differential non-response conditional on the call being received by the beneficiary (0.6%); (2) differential non-response conditional on the call being received by anyone, beneficiary or not (4.3%); (3) differential non-response conditional on a call made by the survey firm to a phone number

(24%). We consider the relevant non-response rate to be the second category outlined above (cases in which a phone call was successfully received by someone, beneficiary or otherwise).

We rule out one explanation for why the third category of non-response might cause us to over-estimate the treatment effects. We consider the case that control households were less likely to respond to the phone surveys, because lack of cash support caused them to experience higher asset loss, including mobile phones, or be more likely to migrate out of mobile coverage. In this case, we would conclude that our treatment effects are likely to be lower bounds as a result.

We correct for all three types of differential non-response using Lee Bound (Lee, 2009). This method relies on the monotonicity assumption that the likelihood of non-response is monotonically related to receiving the treatment. In our case, the assumption implies that the treated observations who did not respond in the survey would also not have responded if they were in the control group - and vice versa. Take up of the cash transfer was high and those who receive the cash are more likely to respond to the survey than the control group, especially if they were better able to cope with the flood impacts. Thus, the monotonicity assumption is likely to be reasonable in this context. However, this assumption would be violated if the cash transfer induced a wide range of responses to the flood that in turn threatens the likelihood of responding to the survey, such as some households choosing to evacuate or migrate, and other households choosing not to.

In Table 1.8, we show the lower bound (from trimming households with outcomes roughly above the 95th percentile) and upper bound (from trimming treated households with outcomes below the 5th percentile) in this order. Correcting for the second type of differential non-response, the effects on child food consumption, asset loss, and costly borrowing are robust and remain highly significant for the most part. The result on child food consumption is robust to correcting even for the most conservative type of differential non-response.

Table 1.8: Lee bounds to correct for differential non-response

		Lee bounds for non-response – by definition of contacted households		
	Standardised mean treatment effect [N]	Call received by beneficiary Δ n-r = 0.6% [N]	Call received Δ n-r = 4.3% [N]	Call made Δ n-r = 24% [N]
Child food consumption	0.070*** [7563]	0.070***/0.070*** [7563/7563]	0.070***/0.070*** [7563/7563]	0.070***/0.502*** [7563/6568]
Adult food consumption index	0.075* [8951]	0.053/0.082** [8909/8924]	-0.030/0.141*** [8670/8678]	-0.376***/0.395*** [7394/7484]
Life satisfaction	0.121** [8941]	0.121**/0.121** [8941/8941]	-0.015/0.121** [8702/8941]	-0.334***/0.121** [7470/8941]
Pre-emptive actions	0.055 [8947]	0.045/0.055 [8920/8947]	0.013/0.055 [8837/8947]	-0.154***/0.055 [8362/8947]
Asset loss index	-0.078** [8950]	-0.100***/-0.078** [8916/8950]	-0.190***/-0.078** [8678/8950]	-0.456***/0.191*** [7470/7340]
Costly borrowing index	-0.100*** [6061]	-0.117***/-0.094*** [6041/6035]	-0.201***/-0.059* [5899/5878]	-0.462***/0.119*** [5095/5155]
Remittances	0.046 [8950]	0.046/0.046 [8950/8950]	0.046/0.046 [8950/8950]	-0.286***/0.046 [8325/8950]
Earning potential index	0.077** [8945]	0.052*/0.077** [8903/8945]	-0.030/0.077** [8675/8945]	-0.324***/0.445*** [7358/7463]

Notes: Lee (2009) bounds (lower/upper) for estimated treatment effects are shown for different definitions of contacted households and corresponding differential non-response rates. Higher non-response was recorded in the control group in all cases. Reasons considered as non-response under alternative definitions: (a) call received by participant – partial interview, refused to interview; (b) call received – partial interview, refused to interview, rescheduled – call received by other person who was staying far from the home during call, participant absent (visiting relatives), not interviewed for other reasons; (c) call made – as for call received and phone switched off, did not answer call. The *unreachable* control category and incorrect phone number cases are excluded from all three non-response groups. Covariates include age, gender, education level, household size, dependency ratio, house structure, UNFPA/FAO recipient status and land type. Union fixed effects are included. Standard errors are clustered at union level for inference. The number of observations included when calculating high and low bounds are shown in square brackets. * p<0.10, ** p<0.05, *** p<0.01

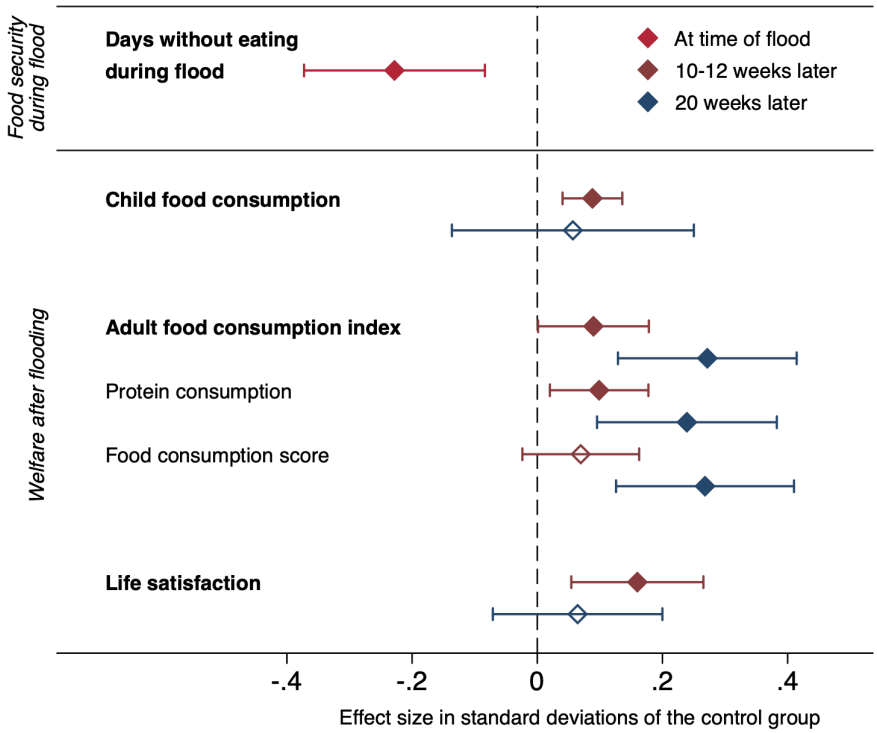
1.8 Results five months post-intervention

We explore the effects of the cash transfer on household welfare in the subset of unions covered by the second round of phone surveys five months after the intervention, noting that this survey has less power and coverage relative to our first round of data collection. We employ the same empirical strategy as outlined in Section 3.4. For comparability, we also present results from the first round using only the panel of unions covered in both rounds of data collection.

Figure 1.8 illustrates the standardised treatment effects on our three welfare outcomes for the two rounds of data (results also reported in Appendix Table 1.28). Focusing first on the treatment effects from the first round of data collected 10-12 weeks after the intervention, we see that the welfare effects of the cash transfer

mirror those noted in Section 1.5.1, if not even stronger in magnitude. Our panel of unions are thus not too dissimilar to the full sample of unions considered in the earlier analysis.

Figure 1.8: Welfare effects over time within panel unions



Notes: Markers indicate the standardised mean treatment effect of receiving the cash transfer on welfare outcomes during the flood (red), 10-12 weeks later (maroon), and 20 weeks later (blue), with 90% confidence intervals shown. The sample is restricted to unions included in both surveys, and upazilas (the administrative unit above a union) with at least 10 treatment and control households. Solid markers indicate significance at the 10% level. Covariates include age, gender, education level, household size, dependency ratio, house structure, UNFPA/FAO recipient status and land type. Union fixed effects are included. Standard errors are clustered at union level.

Five months after the intervention, we find that child food consumption has improved for both treated and control households, relative to the earlier round of data collection. However, as illustrated in Figure 1.9, control households have converged to treated households, such that the difference is no longer statistically significant; 86% of control households reported that children had consumed at least three meals in the day prior to the survey in December 2020 compared to only 80% in late September. This convergence is reassuring and perhaps unsurprising, given

that the child food consumption measure is blunt and captures an extreme coping strategy. Similarly, life satisfaction improves for both treated and control households, with some convergence for control households. We cannot reject the null hypothesis that the cash transfer has had no effect on life satisfaction five months after the intervention relative to the control group. Table 1.9 reports the *p*-values and sharpened *q*-values for the standardised treatment effects on the welfare outcomes and the percentage change relative to the control mean. We do not observe significant treatment effects on any other outcomes.

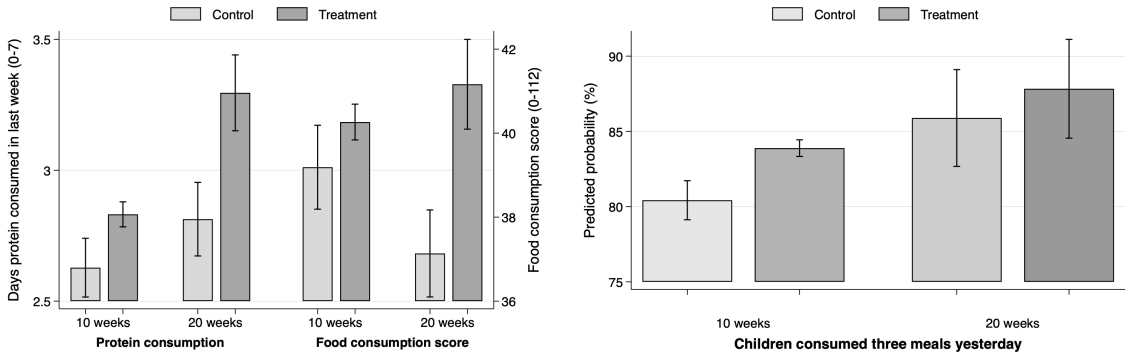
Table 1.9: Standardised mean treatment effect on welfare outcomes within panel unions (second round only)

	Std. treatment effect	Control mean	Δ	% Δ	*p*-value	*q*-value	N
Day without eating during flood (0/1)	-0.222**	0.28	-0.10	-35.7%	0.030	0.047	1284
Children consumed three meals (0/1)	0.057	0.87	0.02	+2.3%	0.625	0.455	1057
Adult food consumption index	0.272***				0.002	0.010	1284
Days protein products consumed (0–7)	0.239***	2.97	0.48	+16.2%	0.007		1284
Food consumption score (0–112)	0.268***	38.22	4.03	+10.5%	0.003		1284
Life satisfaction (0–10)	0.066	2.41	0.13	+5.4%	0.421	0.391	1283

Notes: The first column shows the standardised mean treatment effect for welfare outcomes measured in the second survey (for households in panel unions). The second column shows the control mean, followed by the non-standardised treatment effect (Δ) and percentage change relative to the control mean (% Δ). *p*-values are reported on all outcomes and sub-indices, with standard errors clustered at union level. False discovery rate *q*-values for four hypotheses are calculated over the main outcomes following the sharpened two-stage procedure of Benjamini et al. (2006) . Covariates include age, gender, education level, household size, dependency ratio, house structure and land type. Union fixed effects are included. * p<0.10, ** p<0.05, *** p<0.01

In contrast, we observe a strong positive treatment effect on adult food consumption of 0.272 standard deviations relative to the control group (p=0.002). This effect is statistically significant at the 1% level after correcting for the false discovery rate (Table 1.9). As shown in Figure 1.9, improvements in adult food consumption are driven by a strong increase in the number of days in which expensive proteins are consumed (+16%, p=0.007) and in general nutrient in-take or FCS (+11%, p=0.003) in the week prior to the survey. Within the FCS, there is an increase in almost all forms of nutrients consumed by treated households relative to control households. Whereas protein and general nutrient intake increased over time for treated households, protein consumption increased by a lesser extent for control

Figure 1.9: Child and adult food consumption 10 and 20 weeks after the intervention



(a) Changes in adult food consumption (b) Changes in child food consumption

Notes: Panel (a) shows the predicted protein consumption and FCS for adults. Panel (b) shows the predicted probability that children consumed three meals yesterday. Both were estimated using the main specification. Samples vary from 10 weeks to 20 weeks. Covariates include age, gender, education level, household size, dependency ratio, house structure, UNFPA/FAO recipient status and land type. Union fixed effects are included. Standard errors are clustered at union level. 90% confidence intervals are shown.

households, while general nutrient intake significantly decreased over time. We test whether control households substitute away from adult food consumption to protect child food consumption over time but reject this hypothesis.[36]

1.9 Conclusion

In the face of increasing climate volatility and stretched aid budgets, a better understanding of how to support households effectively in times of crisis is needed. This paper examines the welfare impact of a one-off anticipatory cash transfer provided to households affected by extreme floods in Bangladesh.

A cash transfer provided seven days on average before local peak flooding allowed households to spend more on food in anticipation of the flood. This had an immediate and large impact on food consumption. Households with the cash transfer were 36% less likely to go a day without eating compared to households that did not receive the transfer. The transfers also encouraged households to take

[36]Child and adult food consumption are complementary in panel control households (93% of control households in the second survey were included in the first survey). The decrease in the adult FCS is significant for control households without children (p=0.078), but not for the majority (77%) of these control households with children (see Appendix Table 1.29).

action to reduce the impact of the flood. Households that received the cash transfer were 12% more likely to evacuate household members and 17% more likely to evacuate their livestock. As a result, transfer-receiving households were 8% less likely to lose small livestock and 5% less likely to lose poultry during and after the flooding.

The effects of the transfer were still present three months later. Children in households that received the transfers were 4% more likely to have consumed three meals in the day prior to the survey. Adults in the treatment households had higher food security; for example, the number of days that meat, fish or eggs was consumed was higher by 7%. Stated wellbeing was 12.5% higher for transfer-receiving households (using Cantril's ladder of life satisfaction measure). Those receiving the transfer borrowed less (they were 3% less likely to borrow and borrowed 6% less when they did) and did so more cheaply (monthly interest rates were 12% lower). They were more likely to be working (they were 6% more likely to work and hours of work were 7% higher). The effect sizes after three months are small, but are surprisingly broad, impacting both current and future welfare.

There are five important take-aways for crisis response from these results. First, one-off cash transfers provided to households in a natural disaster make a difference to food security and stated wellbeing. They also reduce behaviours that have a cost on future wellbeing (such as reducing child consumption and costly borrowing) and help households recover quicker, pointing to long-run welfare benefits.

Second, early action in a crisis is essential. Failing to act early has real welfare costs. All the welfare impacts estimated in this analysis were measured before the typical arrival date of humanitarian assistance. During the floods of 2019, households started to receive support 100 days after the flood. These results show that providing support early – in this case, even in advance of the flood – can alleviate losses that are known to have scarring effects, such as lower child consumption and costly borrowing. As a result, there are significant advantages from early action, even if later support provides some potential for recovery.

Third, speed matters, even when acting early. A small one-off cash transfer

is more effective if received earlier in shock response. Households that received a cash transfer one day earlier during the floods had a small, but statistically significant difference in adult food consumption. This highlights the importance of investing in preparing and targeting well in advance of disasters, such that the intervention can scale immediately with the activation of triggers, and the value of using earlier (10-day rather than five-day) triggers to release funding. It also highlights the need to innovate on protocols for immediately enrolling potential beneficiaries in mobile money platforms (as showcased in Aiken et al. (2021) in Togo).

Fourth, more work is needed to determine the ideal size and frequency of the transfers, as a small one-off transfer appears not to be enough. The transfers provided were equivalent to two weeks' worth of support, but the crisis lasted much longer and the needs of households were much larger than this. Unsurprisingly then, the size of the impacts we estimate are small. It is possible that larger and longer lasting impacts will require larger or repeated transfers. This is something that should be explored in future work.

Fifth, there seems to be some potential in utilising geographic variables in targeting crisis response. We find that the cash transfer was more beneficial for households located in vulnerable lands and in areas where households are on average poorer. While we do not know what exactly is driving the differential results across space, the fact that higher impacts are seen in areas that could be identified as more vulnerable highlights the potential for using more geographic variables to target crisis response.

There are also some clear lessons from this work on evaluating humanitarian interventions. This paper is an unusual large-scale evaluation of the impact of humanitarian cash transfers provided in response to a sudden onset climate disaster. The unprecedented scale-up of cash transfers during the Covid-19 crisis will likely provide other valuable learning opportunities.

Much more learning on emergency response is needed. There are real challenges to conducting rigorous impact evaluations in these settings. The absence of baseline data requires large ex-post surveys. The impacts of one-off interventions

are immediate, which in turn requires speed in data collection. This may mean using data collected via phone surveys. However, perhaps most significant, is the challenge of identifying a valid control group. This requires a stronger commitment from humanitarian partners to learning and to using rules-based processes to target beneficiaries. When these rules are set out clearly in advance of a crisis and adhered to during implementation, it is possible to conduct rigorous evaluations of impact without withholding time-sensitive support to vulnerable households. We hope that this work motivates a stronger commitment by documenting the welfare benefits resulting from the hard work of humanitarian partners and identifying clear lessons on the timing and targeting of support.

Appendices of Chapter 1

Table 1.10: Balance across treatment groups by transfer date

	14 July mean	15 July mean	16 July mean	30 July mean	Joint orth. p-value	14=15=16 July p-value
Individual characteristics						
Age	38.04	39.72	39.59	38.46	0.004	0.003
Female respondent	0.97	0.97	0.97	0.96	0.269	0.141
Household head	0.22	0.20	0.22	0.21	0.704	0.456
Completed primary school	0.34	0.26	0.25	0.30	0.000	0.000
Household characteristics						
Household size	4.73	4.74	4.74	4.72	0.060	0.047
Dependency ratio	0.76	0.74	0.75	0.77	0.064	0.047
Raw material house	0.25	0.28	0.29	0.29	0.751	0.953
Distance to large water body (m)	1419.68	1060.98	1081.00	1165.52	0.565	0.374
Protected mainland	0.38	0.27	0.28	0.26	0.392	0.615
Unprotected mainland	0.26	0.30	0.27	0.29	0.476	0.454
Char land	0.36	0.43	0.45	0.45	0.916	0.806
Anticipatory action						
Received WFP cash transfer	1.00	1.00	1.00	1.00	.	.
Received dignity kit from UNFPA	0.17	0.12	0.11	0.11	0.327	0.174
Received feed or storage from FAO	0.08	0.07	0.05	0.08	0.390	0.348
Technology						
Used digital wallet in last six months	0.48	0.48	0.49	0.38	0.023	0.237
Own mobile	0.79	0.82	0.80	0.79	0.635	0.453
Uses someone else's mobile	0.19	0.16	0.19	0.20	0.310	0.185
Uses mobile at least once a week	0.96	0.96	0.97	0.97	0.679	0.406
Observations	3312	1218	1085	670		

Notes: The tests for equivalence across means report the p-value controlling for union fixed effects and clustering standard errors at union level as in our main specification.

Table 1.11: Balance in sample from second round of surveys

	Control mean	Treatment mean	Δ	Norm. Diff.	*p*-value
Individual characteristics					
Age	37.00	37.89	0.88	0.07	0.665
Female respondent	0.99	0.99	0.00	0.03	0.965
Household head	0.22	0.26	0.04	0.09	0.948
Female head	0.25	0.27	0.02	0.05	0.422
Completed primary school	0.36	0.36	-0.01	-0.01	0.804
Household characteristics					
Household size	4.47	4.53	0.05	0.03	0.558
Dependency ratio	0.77	0.78	0.01	0.02	0.919
Raw material house	0.25	0.24	-0.01	-0.03	0.820
Household asset categories (June)	6.02	5.90	-0.12	-0.06	0.491
Protected	0.54	0.54	0.01	0.01	0.819
Unprotected	0.26	0.28	0.02	0.05	0.355
Char	0.20	0.17	-0.03	-0.08	0.155
Anticipatory action					
Received WFP cash transfer in July	0.00	1.00	1.00	.	.
Received WFP cash transfer in October	0.00	0.00	0.00	.	.
Technology					
Used digital wallet in last six months	0.61	0.49	-0.12	-0.25	0.030
Own mobile	0.86	0.87	0.00	0.00	0.860
Uses someone else's mobile	0.13	0.13	-0.00	-0.01	0.959
Uses mobile at least once a week	0.99	0.99	-0.00	-0.00	0.218
Observations	659	632			

Notes: Δ reports the treatment mean minus the control mean. Norm. Diff. reports the normalised difference between the treatment and control group means, following Imbens and Rubin (2015). The last column reports the *p*-value from ordinary least squares regressions of each variable on the treatment dummy to test equivalence of means, controlling for union fixed effects and clustering standard errors at union level as in our main specification.

Table 1.12: Variable construction for welfare measures

Index	Sub-scale	Question(s)
1. Children's food consumption	Dummy variable for whether children consumed three or more meals in the previous day.	How many meals did children (younger than 15 years old) eat yesterday? [Number of meals]
2. Adult food consumption	Number of days meat products were consumed over the last seven days (the selection of meat will be confirmed by looking at the variation in frequency of consumption of all food items listed).	How many days over the last seven days, did adult members (15 years or older than 15 years) of your household eat meat, fish, eggs (goat, beef, chicken, buffalo, fish, including tuna, dry fish, and/or other seafood, eggs)?
	Food consumption score (FCS)	The following question will be used to construct the FCS. The FCS will then be calculated according to the standard formula: $FCS = (starches * 2) + (pulses * 3) + vegetables + fruit + (meat * 4) + (dairy * 4) + (fats * .5) + (sugar * .5)$ How many days over the last seven days, did adult members (15 years or older than 15 years) of your household eat the following food items, prepared and/or consumed at your home? 1. Cereals, excluding rice (pasta, bread, sorghum, millet, maize, fonio, potato, yam, cassava, white sweet potato, parched rice (muri), chira) 2. Legumes/nuts (beans, peas, peanuts, lentils, mascalai, mung beans, khesari, ankar, arahar pulses, nut, soy, and / or other nuts) 3. Milk and other dairy products (fresh milk/sour, yogurt, cheese, other dairy products) (exclude margarine/butter or small amounts of milk if used in tea/coffee) 4. Meat, fish, eggs (goat, beef, chicken, buffalo, fish, including tuna, dry fish, and/or other seafood, eggs) 5. Vegetables and leaves (various spinach, onion, tomatoes, carrots, peppers, green beans, lettuce, etc.) 6. Fruits (banana, apple, lemon, mango, papaya, peach, etc.) 7. Oil, fat, butter (vegetable oil, palm oil, shea butter, margarine, other fats/oil) 8. Sugar or sweet (sugar, honey, jam, cakes, candy, cookies, pastries, cakes and other sweets including sugary drinks)
3. Wellbeing	Life satisfaction	Please imagine a ladder with steps numbered from 0 at the bottom to 10 at the top. The top of the ladder represents the best possible life for you and the bottom of the ladder represents the worst possible life for you. Which step of the ladder best represents the way you personally feel you stand these days? [0–10]

Note that for questions on food consumption, respondents were encouraged to pass the phone to someone in the household who could respond the questions about food consumption with sufficiently good recall.

Table 1.13: Variable construction for other outcome measures

Outcome	Construction	Question(s)
1. Actions taken to reduce the impact of the flood	Number of preventative actions taken	Which actions did you take to prepare for the flooding? (count of the following actions taken): 1. Protect valuable assets 2. Evacuate household members/moved 3. Purchase food 4. Evacuate livestock 5. Protect roof/walls 6. Warn others
2. Household asset loss or damage	A standardised index of the following variables:	
	Number of livestock that died over the past two months	How many cows, calves and buffalo that you owned died during the past two months (from July 15 to September 15)? How many goats, sheep and pigs that you owned died in the past two months(from July 15 to September 15)?
	Number of categories of household assets that were lost or damaged	Other than damage to your house and animals; were any assets damaged or lost due to the flooding? What assets were damaged/lost? Enter all that apply. 1. Poultry 2. Crop (stock in home) 3. Irrigation pump 4. Fruit plantation 5. Fish 6. Equipment for fishing i.e., fishing net 7. Vehicle by any animal 8. Boat 9. Rickshaw, van, or cycle etc. 10. Shop 11. Sewing machine 12. Furniture 13. Clothes 14. Household appliances like home utensils, mobile phone, television etc. 15. Ornaments (gold and silver) 16. Others: (specify)
	Area of cultivated plots lost [in decimal]	Have you lost cultivated crops in the past two months (from July 15 to September 15) due to the flooding? [Yes/No] If yes: How much have you lost in cultivated plots in decimal?

Variable construction for outcome measures (cont.)

Outcome	Construction	Question(s)
3. Costly borrowing	A standardised index of the following two variables: 1. How much was borrowed in the last two months (in BDT)? 2. The highest interest rate charged (percent per month)	In the past two months (from July 15 to September 15), has your household borrowed any money from friends/family/credit institutions or groups – both formal and informal – to cover for basic needs? [Yes/No] If yes: How much did you borrow in the past two months (from July 15 to September 15)? [BDT] What is the highest interest rate you were charged on the loan(s) you received in the past two months (from July 15 to September 15)? [%] Was this interest rate per month or per year? [monthly/yearly]
4. Remittances	Dummy variable for whether household received remittances in the last two months.	Did you receive any remittances in the past two months (from July 15 to September 15)? [Yes/No] If yes: How much did you receive in the past two months (from July 15 to September 15)? [BDT]
5. Earnings potential	A standardised index of the following two variables: 1. Able to replant (dummy variable taking the value of 1 if the household reported replanting) 2. Number of hours worked for an income in the last seven days (hours)	Have you lost cultivated crops in the past two months (from July 15 to September 15) due to the flooding? [Yes/No] If yes: Have you been able to replant? [Yes/No] How many hours did you or someone in your household work towards an income in the past seven days?

Table 1.14: False Discovery Rate q-values and wild bootstrap inference

	Standardised mean treatment effect	p-value	q-value	Wild bootstrap p-value	N
Child food consumption	0.070***	0.006	0.027	0.007	7563
Adult food consumption index	0.075*	0.072	0.046	0.084	8951
Life satisfaction	0.121**	0.018	0.028	0.016	8941
Pre-emptive actions	0.055	0.240	0.099	0.266	8947
Asset loss index	-0.078**	0.027	0.033	0.033	8950
Costly borrowing index	-0.100***	0.005	0.027	0.010	6061
Remittances	0.046	0.155	0.072	0.162	8950
Earning potential index	0.077**	0.012	0.027	0.021	8945

False discovery rate q-values for eight hypotheses are calculated following the sharpened two-stage procedure of Benjamini et al. (2006). Covariates include age, gender, education level, household size, dependency ratio, house structure, UNFPA/FAO recipient status and land type. Union fixed effects are included. Standard errors clustered at union level. * p<0.10, ** p<0.05, *** p<0.01

Table 1.15: Treatment effect on secondary outcomes

	(1) Early warning (0/1)	(2) Any action to prepare (0/1)	(3) Evacuated household (0/1)	(4) Evacuated livestock (0/1)	(5) Lost small livestock (0/1)	(6) Lost poultry (0/1)	(7) Borrowed money (0/1)	(8) Worked for wage (0/1)
Transfer	0.009	0.040*	0.034**	0.030**	-0.023*	-0.032*	-0.024*	0.039***
	(0.019)	(0.020)	(0.014)	(0.013)	(0.014)	(0.018)	(0.013)	(0.012)
Control mean	0.61	0.53	0.29	0.17	0.30	0.60	0.70	0.68
% Δ	1.4%	7.4%	11.7%	17.3%	-7.6%	-5.4%	-3.4%	5.8%
Controls	✓	✓	✓	✓	✓	✓	✓	✓
Union fixed effects	✓	✓	✓	✓	✓	✓	✓	✓
N	8951	8947	8947	8947	8952	8950	8946	8950
R^2	0.07	0.10	0.10	0.07	0.06	0.17	0.06	0.06

Notes: Covariates include age, gender, education level, household size, dependency ratio, house structure, UNFPA/FAO recipient status and land type. Union fixed effects are included. Standard errors are clustered at union level. * p<0.10, ** p<0.05, *** p<0.01

Table 1.16: Treatment effect on main outcomes for 14,15,16 July transfers vs 30 July transfers

	(1) Child food consumption	(2) Adult food consumption index	(3) Life satis-faction	(4) Pre-emptive actions	(5) Asset loss index	(6) Costly borrowing index	(7) Remit-tances	(8) Earning potential index
Transfer on	0.144***	0.126***	0.101*	0.094*	-0.051	-0.050	0.065*	0.050
14–16 July	(0.033)	(0.038)	(0.057)	(0.054)	(0.036)	(0.045)	(0.035)	(0.037)
Transfer on 30	0.126**	0.033	0.116*	0.001	-0.122***	-0.205***	0.076	0.099**
July	(0.049)	(0.049)	(0.067)	(0.061)	(0.046)	(0.053)	(0.067)	(0.049)
p-value: Before = After	0.686	0.019	0.756	0.083	0.089	0.001	0.853	0.247
q-value: Before	0.001	0.007	0.106	0.106	0.162	0.162	0.106	0.162
q-value: After	0.032	0.252	0.106	0.328	0.029	0.002	0.162	0.091
Controls	✓	✓	✓	✓	✓	✓	✓	✓
Union fixed effects	✓	✓	✓	✓	✓	✓	✓	✓
N	7217	8552	8542	8548	8551	5785	8551	8547
R^2	0.04	0.10	0.10	0.11	0.13	0.10	0.04	0.12

Notes: The standardised mean treatment effect of receiving cash before the flood peak (14, 15, 16 July) and after the flood peak (30 July) is shown in the main panel, with standard errors clustered at union level in parentheses. The second panel shows p-values from a test of equivalence comparing before and after flood transfer coefficients. The false discovery rate q-values for 16 hypotheses are calculated over the main outcomes for before and after flood transfers, following the sharpened two-stage procedure of Benjamini et al. (2006). Covariates include age, gender, education level, household size, dependency ratio, house structure, UNFPA/FAO recipient status, land type and wallet use category. We also control for survey date. Union fixed effects are included. * p<0.10, ** p<0.05, *** p<0.01

Table 1.17: Marginal effect of an earlier cash transfer on secondary outcomes

	(1) Early warning (0/1)	(2) Any action to prepare (0/1)	(3) Evacuated household (0/1)	(4) Evacuated livestock (0/1)	(5) Lost small livestock (0/1)	(6) Lost poultry (0/1)	(7) Borrowed money (0/1)	(8) Worked for wage (0/1)	(9) Replanted lost crops (0/1)
Transfer	0.013	0.039*	0.004	0.009	-0.046***	-0.012	-0.004	0.012	-0.030
	(0.021)	(0.021)	(0.020)	(0.015)	(0.016)	(0.020)	(0.017)	(0.019)	(0.029)
Transfer × days	0.001	0.001	0.003**	0.002	0.002*	0.001	-0.001	0.001	-0.004**
before flood peak	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Control mean	0.61	0.53	0.29	0.17	0.30	0.60	0.70	0.68	0.24
% Δ transfer	2.1%	7.3%	1.3%	5.4%	-15.1%	-1.9%	-0.6%	1.8%	-12.7%
% Δ/day vs transfer	0.2%	0.2%	1.1%	0.9%	0.9%	0.2%	-0.2%	0.1%	-2.1%
Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓
Mauza fixed effects	✓	✓	✓	✓	✓	✓	✓	✓	✓
N	8354	8349	8349	8349	8354	8352	8348	8353	3405
R^2	0.15	0.18	0.18	0.15	0.12	0.24	0.13	0.14	0.18

Notes: Covariates include age, gender, education level, household size, dependency ratio, house structure, UNFPA/FAO recipient status and land type. We also control for survey date. Union fixed effects are included. Standard errors are clustered at union level. * p<0.10, ** p<0.05, *** p<0.01

Table 1.18: Marginal effect of an earlier cash transfer - IV

	(1) Child food consumption	(2) Adult food consumption index	(3) Life satis-faction	(4) Pre-emptive actions	(5) Asset loss index	(6) Costly borrowing index	(7) Remit-tances	(8) Earning potential index
Transfer	0.128***	0.077**	0.064	0.036	-0.074**	-0.125***	0.052	0.067*
	(0.035)	(0.039)	(0.050)	(0.051)	(0.035)	(0.044)	(0.042)	(0.037)
Transfer × days	-0.000	0.006**	-0.001	0.006*	0.004	0.010***	0.001	-0.004
before flood peak	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)
Weak-id test: F-stat	2686	2515	2515	2518	2520	1853	2516	2532
Endog. test: *p*-value	0.253	0.113	0.250	0.144	0.062	0.003	0.740	0.270
Over-id test: *p*-value	0.180	0.993	0.139	0.341	0.063	0.137	0.531	0.369
Controls	✓	✓	✓	✓	✓	✓	✓	✓
Union fixed effects	✓	✓	✓	✓	✓	✓	✓	✓
IV	✓	✓	✓	✓	✓	✓	✓	✓
N	7117	8442	8432	8438	8441	5710	8441	8437
R^2	0.01	0.03	0.01	0.01	0.03	0.03	0.01	0.07

Excluded instruments for the number of days before the flood peak include dummy variables for receiving the transfer on the 14, 15, and 16 of July, respectively (30 July is redundant). Estimated using the two-step efficient generalised method of moments.
Weak-id test reports the Kleibergen-Paap F-statistic for weak identification (the Stock-Yogo weak ID critical value for 10% maximal IV size=19.93). Endog. test reports the *p*-value from an Anderson-Rubin test for the significance of endogenous regressor (H0: days before flood peak=0 and orthogonality conditions are valid). Over-id test reports the *p*-value from a Hansen J overidentification test of all instruments (H0: the instruments are valid).

Table 1.19: Marginal effect of an earlier cash transfer - excluding 30 July

	(1) Child food consumption	(2) Adult food consumption index	(3) Life satis-faction	(4) Pre-emptive actions	(5) Asset loss index	(6) Costly borrowing index	(7) Remit-tances	(8) Earning potential index
Transfer	0.088	0.036	0.125	0.133*	-0.019	-0.116	0.035	-0.000
	(0.066)	(0.069)	(0.095)	(0.072)	(0.068)	(0.073)	(0.069)	(0.064)
Transfer × days	0.006	0.010	-0.002	-0.006	-0.003	0.009	0.003	0.004
before flood peak	(0.007)	(0.007)	(0.009)	(0.007)	(0.007)	(0.007)	(0.007)	(0.006)
Controls	✓	✓	✓	✓	✓	✓	✓	✓
Mauza fixed effects	✓	✓	✓	✓	✓	✓	✓	✓
N	6469	7695	7685	7691	7694	5142	7694	7692
R^2	0.11	0.15	0.18	0.20	0.21	0.17	0.11	0.18

Notes: Covariates include age, gender, education level, household size, dependency ratio, house structure, UNFPA/FAO recipient status and land type. We also control for survey date. Union fixed effects are included. Standard errors are clustered at union level. * p<0.10, ** p<0.05, *** p<0.01

Table 1.20: Summary statistics by land type

	Unprotected mean	Char mean	Protected mean	Joint orth. p-value
Individual characteristics				
Age	38.10	39.57	37.51	0.751
Female respondent	0.98	0.93	1.00	0.881
Household head	0.22	0.22	0.18	0.531
Completed primary school	0.34	0.26	0.34	0.289
Household characteristics				
Household size	4.72	4.85	4.58	0.428
Dependency ratio	0.76	0.77	0.73	0.828
Raw material house	0.29	0.29	0.24	0.145
Distance to large water body (m)	1397.50	458.19	2025.82	0.000
Satellite measure of flood intensity				
Maximum flood extent (% mauza)	0.35	0.31	0.28	0.347
Flood duration (days above 50% max. extent)	47.91	54.37	38.11	0.022
Anticipatory action				
Received WFP cash transfer	0.73	0.78	0.68	0.019
Received dignity kit from UNFPA	0.09	0.12	0.15	0.629
Received feed or storage from FAO	0.04	0.08	0.07	0.004
Technology				
Used digital wallet in last six months	0.51	0.47	0.47	0.471
Own mobile	0.80	0.79	0.83	0.230
Uses someone else's mobile	0.18	0.19	0.16	0.399
Uses mobile at least once a week	0.96	0.97	0.97	0.692
Observations	2405	3345	3204	

Notes: The tests for equivalence across means report the p-value controlling for union fixed effects and clustering standard errors at union level as in our main specification.

Table 1.21: Treatment effect on secondary outcomes by land type

	(1) Early warning (0/1)	(2) Any action to prepare (0/1)	(3) Evacuated household (0/1)	(4) Evacuated livestock (0/1)	(5) Lost small livestock (0/1)	(6) Lost poultry (0/1)	(7) Borrowed money (0/1)	(8) Worked for wage (0/1)
Transfer	0.017	0.048	0.031	0.045**	-0.042*	-0.006	-0.021	0.014
	(0.026)	(0.031)	(0.020)	(0.020)	(0.025)	(0.030)	(0.026)	(0.021)
Char	0.012	-0.026	-0.037	0.003	-0.035	0.003	0.000	-0.081**
	(0.046)	(0.042)	(0.037)	(0.030)	(0.032)	(0.038)	(0.033)	(0.036)
Transfer × char	-0.009	0.022	0.010	-0.016	0.035	-0.027	0.030	0.069**
	(0.037)	(0.041)	(0.030)	(0.028)	(0.035)	(0.034)	(0.030)	(0.030)
Protected	0.023	0.028	-0.046	0.039	-0.039	0.057	0.065	-0.015
	(0.041)	(0.052)	(0.041)	(0.029)	(0.033)	(0.044)	(0.042)	(0.040)
Transfer × protected	-0.015	-0.050	-0.003	-0.028	0.019	-0.050	-0.044	-0.000
	(0.042)	(0.049)	(0.033)	(0.025)	(0.032)	(0.041)	(0.033)	(0.033)
Control mean: Unprotected	0.59	0.50	0.26	0.12	0.30	0.55	0.68	0.67
Treat effect: Unprotected	0.017	0.048	0.031	0.045**	-0.042*	-0.006	-0.021	0.014
% Δ: Unprotected	2.9%	9.5%	11.7%	36.1%	-14.1%	-1.1%	-3.0%	2.0%
p-value: Unprotected	0.523	0.131	0.122	0.029	0.098	0.843	0.430	0.524
Control mean: Char	0.65	0.53	0.27	0.15	0.28	0.50	0.69	0.65
Treat effect: Char	0.007	0.070**	0.041*	0.029	-0.007	-0.033*	0.009	0.083***
% Δ: Char	1.1%	13.2%	15.3%	19.4%	-2.5%	-6.7%	1.3%	12.7%
p-value: Char	0.796	0.017	0.090	0.144	0.754	0.095	0.622	0.000
Control mean: Protected	0.60	0.56	0.32	0.22	0.32	0.69	0.73	0.70
Treat effect: Protected	0.002	-0.002	0.028	0.016	-0.023	-0.056*	-0.064***	0.014
% Δ: Protected	0.3%	-0.4%	8.8%	7.4%	-7.1%	-8.1%	-8.8%	2.0%
p-value: Protected	0.953	0.954	0.275	0.344	0.337	0.064	0.001	0.561
p-value: Unprot. = Char.	0.797	0.597	0.733	0.573	0.321	0.428	0.317	0.022
p-value: Unprot. = Prot.	0.721	0.311	0.931	0.251	0.553	0.224	0.186	1.000
p-value: Char = Prot.	0.888	0.146	0.694	0.576	0.637	0.428	0.006	0.028
F-test *p*-value:								
Control $\Delta = 0$	0.836	0.604	0.439	0.247	0.445	0.237	0.222	0.076
Treatment effect $\Delta = 0$	0.935	0.344	0.910	0.504	0.604	0.472	0.021	0.030
Controls	✓	✓	✓	✓	✓	✓	✓	✓
Union fixed effects	✓	✓	✓	✓	✓	✓	✓	✓
N	8951	8947	8947	8947	8952	8950	8946	8950
R^2	0.07	0.10	0.10	0.07	0.06	0.17	0.06	0.07

Notes: Covariates include age, gender, education level, household size, dependency ratio, house structure, UNFPA/FAO recipient status and land type. Union fixed effects are included. Standard errors are clustered at union level. * p<0.10, ** p<0.05, *** p<0.01

Table 1.22: Standardised treatment effects on sub-samples by wallet-use

	(1) Child food consumption	(2) Adult food consumption index	(3) Life satis-faction	(4) Pre-emptive actions	(5) Asset loss index	(6) Costly borrowing index	(7) Remit-tances	(8) Earning potential index
Existing wallet users								
Transfer	0.070**	0.054	0.104	0.043	-0.078*	-0.088**	0.028	0.089**
	(0.031)	(0.052)	(0.070)	(0.053)	(0.043)	(0.041)	(0.048)	(0.037)
N	4606	5439	5435	5437	5438	3664	5439	5436
R^2	0.05	0.10	0.13	0.11	0.14	0.11	0.05	0.12
New wallet users								
Transfer	0.087*	0.124*	0.099*	0.096	-0.060	-0.102*	0.103*	0.042
	(0.048)	(0.065)	(0.058)	(0.081)	(0.056)	(0.058)	(0.058)	(0.048)
N	2842	3384	3377	3382	3384	2300	3383	3381
R^2	0.07	0.11	0.12	0.15	0.15	0.11	0.06	0.14
p-value: Exist.=New	0.774	0.378	0.959	0.568	0.799	0.835	0.337	0.418
Controls	✓	✓	✓	✓	✓	✓	✓	✓
Union fixed effects	✓	✓	✓	✓	✓	✓	✓	✓

Notes: Results are reported for two separate models corresponding to different mobile money users. Coefficients are interpreted as the standardised mean treatment effect relative to control households in the same wallet use category, with standard errors clustered at union level shown in parentheses. *p*-values from a seemingly unrelated regression test of equivalent coefficients from pairwise comparisons of the models are reported in the second last panel. Covariates for all models include age, gender, education level, household size, dependency ratio, house structure, UNFPA/FAO recipient status and land type. Union fixed effects are included. * $p<0.10$, ** $p<0.05$, *** $p<0.01$

Table 1.23: Standardised treatment effects on sub-samples by wallet-use

	(1) Child food consumption	(2) Adult food consumption index	(3) Life satis-faction	(4) Pre-emptive actions	(5) Asset loss index	(6) Costly borrowing index	(7) Remit-tances	(8) Earning potential index
Frequent wallet users								
Transfer	0.059	-0.013	0.097	0.014	-0.089	-0.127**	-0.106*	0.130***
	(0.051)	(0.059)	(0.068)	(0.065)	(0.058)	(0.061)	(0.061)	(0.047)
N	2298	2678	2676	2678	2677	1834	2678	2675
R^2	0.07	0.10	0.15	0.19	0.19	0.14	0.09	0.16
Infrequent wallet users								
Transfer	0.117**	0.100	0.091	0.049	-0.066	-0.035	0.140**	0.078
	(0.050)	(0.069)	(0.088)	(0.065)	(0.057)	(0.059)	(0.056)	(0.062)
N	2306	2758	2756	2756	2758	1820	2758	2758
R^2	0.09	0.15	0.15	0.11	0.13	0.13	0.07	0.13
New wallet users								
Transfer	0.087*	0.124*	0.099*	0.096	-0.060	-0.102*	0.103*	0.042
	(0.048)	(0.065)	(0.058)	(0.081)	(0.056)	(0.058)	(0.058)	(0.048)
N	2842	3384	3377	3382	3384	2300	3383	3381
R^2	0.07	0.11	0.12	0.15	0.15	0.11	0.06	0.14
p-value: Freq.=Infreq.	0.450	0.138	0.930	0.607	0.771	0.273	0.000	0.539
p-value: Freq.=New	0.693	0.139	0.983	0.465	0.747	0.782	0.012	0.203
p-value: Infreq.=New	0.657	0.776	0.941	0.597	0.930	0.397	0.658	0.628
Controls	✓	✓	✓	✓	✓	✓	✓	✓
Union fixed effects	✓	✓	✓	✓	✓	✓	✓	✓

Notes: Results are reported for three separate models corresponding to different mobile money users. Coefficients are interpreted as the standardised mean treatment effect relative to control households in the same wallet use category, with standard errors clustered at union level shown in parentheses. *p*-values from a seemingly unrelated regression test of equivalent coefficients from pairwise comparisons of the models are reported in the second last panel. Covariates for all models include age, gender, education level, household size, dependency ratio, house structure, UNFPA/FAO recipient status and land type. Union fixed effects are included. * $p<0.10$, ** $p<0.05$, *** $p<0.01$

Table 1.24: Robustness to alternate model specifications

	(1) Child food consumption	(2) Adult food consumption index	(3) Life satis-faction	(4) Pre-emptive actions	(5) Asset loss index	(6) Costly borrowing index	(7) Remit-tances	(8) Earning potential index
Main								
Transfer	0.070***	0.075*	0.121**	0.055	-0.078**	-0.100***	0.046	0.077**
	(0.025)	(0.042)	(0.050)	(0.046)	(0.035)	(0.035)	(0.032)	(0.030)
Controls	✓	✓	✓	✓	✓	✓	✓	✓
Union fixed effects	✓	✓	✓	✓	✓	✓	✓	✓
N	7563	8951	8941	8947	8950	6061	8950	8945
R^2	0.04	0.09	0.09	0.09	0.13	0.09	0.04	0.11
Exclude 30 July								
Transfer	0.078***	0.083*	0.127**	0.080*	-0.081**	-0.077*	0.046	0.079**
	(0.030)	(0.043)	(0.059)	(0.046)	(0.039)	(0.039)	(0.033)	(0.034)
Controls	✓	✓	✓	✓	✓	✓	✓	✓
Union fixed effects	✓	✓	✓	✓	✓	✓	✓	✓
N	6755	7998	7990	7994	7997	5398	7997	7995
R^2	0.04	0.09	0.10	0.10	0.13	0.09	0.04	0.11
Mauza fixed effects								
Transfer	0.074**	0.086***	0.124***	0.066**	-0.058*	-0.094***	0.052*	0.045
	(0.030)	(0.033)	(0.038)	(0.033)	(0.034)	(0.031)	(0.030)	(0.028)
Controls	✓	✓	✓	✓	✓	✓	✓	✓
Mauza fixed effects	✓	✓	✓	✓	✓	✓	✓	✓
N	7471	8862	8852	8858	8861	5954	8861	8856
R^2	0.10	0.14	0.16	0.18	0.20	0.17	0.10	0.17
Winsorised at p95								
Transfer	0.070***	0.078*	0.116***	0.068	-0.065**	-0.103***	0.046	0.078**
	(0.025)	(0.041)	(0.041)	(0.043)	(0.033)	(0.037)	(0.032)	(0.032)
Controls	✓	✓	✓	✓	✓	✓	✓	✓
Union fixed effects	✓	✓	✓	✓	✓	✓	✓	✓
N	7563	8951	8941	8947	8950	6061	8950	8945
R^2	0.04	0.09	0.10	0.10	0.13	0.09	0.04	0.10
Excl. UNFPA/FAO								
Transfer	0.065**	0.071*	0.116**	0.047	-0.077**	-0.094***	0.041	0.067**
	(0.029)	(0.041)	(0.055)	(0.047)	(0.036)	(0.035)	(0.036)	(0.031)
Controls	✓	✓	✓	✓	✓	✓	✓	✓
Union fixed effects	✓	✓	✓	✓	✓	✓	✓	✓
N	6366	7549	7539	7546	7548	5120	7549	7545
R^2	0.04	0.09	0.10	0.09	0.12	0.09	0.04	0.12

Notes: Covariates include age, gender, education level, household size, dependency ratio, house structure, UNFPA/FAO recipient status and land type. Union fixed effects are included. Standard errors are clustered at union level. * $p<0.10$, ** $p<0.05$, *** $p<0.01$

Table 1.25: Standardised mean treatment effects - welfare

	Child food consumption		Adult food consumption index		Life satisfaction	
	(1)	(2)	(3)	(4)	(5)	(6)
Transfer	0.053**	0.070***	0.070	0.075*	0.119**	0.121**
	(0.025)	(0.025)	(0.045)	(0.042)	(0.049)	(0.050)
Controls		✓		✓		✓
Union fixed effects	✓	✓	✓	✓	✓	✓
N	7565	7563	8953	8951	8943	8941
R^2	0.03	0.04	0.06	0.09	0.09	0.09

Notes: Covariates include age, gender, education level, household size, dependency ratio, house structure, UNFPA/FAO recipient status and land type. Union fixed effects are included. Standard errors are clustered at union level. * p<0.10, ** p<0.05, *** p<0.01

Table 1.26: Standardised mean treatment effects - livelihood, finance and actions

	Pre-emptive actions		Asset loss index		Costly borrowing index		Remittances		Earning potential index	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Transfer	0.061	0.055	-0.073**	-0.078**	-0.109***	-0.100***	0.047	0.046	0.071**	0.077**
	(0.047)	(0.046)	(0.035)	(0.035)	(0.039)	(0.035)	(0.033)	(0.032)	(0.031)	(0.030)
Controls		✓		✓		✓		✓		✓
Union fixed effects	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
N	8949	8947	8952	8950	6062	6061	8952	8950	8947	8945
R^2	0.09	0.09	0.10	0.13	0.07	0.09	0.03	0.04	0.05	0.11

Notes: Covariates include age, gender, education level, household size, dependency ratio, house structure, UNFPA/FAO recipient status and land type. Union fixed effects are included. Standard errors are clustered at union level. * p<0.10, ** p<0.05, *** p<0.01

Table 1.27: Robustness to controlling for satellite measures of flood intensity at mauza level

	(1) Child food consumption	(2) Adult food consumption index	(3) Life satis-faction	(4) Pre-emptive actions	(5) Asset loss index	(6) Costly borrowing index	(7) Remit-tances	(8) Earning potential index
Main								
Transfer	0.070***	0.075*	0.121**	0.055	-0.078**	-0.100***	0.046	0.077**
	(0.025)	(0.042)	(0.050)	(0.046)	(0.035)	(0.035)	(0.032)	(0.030)
R^2	0.04	0.09	0.09	0.09	0.13	0.09	0.04	0.11
Survey date dummies								
Transfer	0.112***	0.114***	0.089*	0.056	-0.049	-0.093**	0.046	0.064**
	(0.029)	(0.040)	(0.046)	(0.052)	(0.032)	(0.040)	(0.034)	(0.029)
R^2	0.04	0.09	0.10	0.10	0.13	0.10	0.04	0.11
Maximum flood extent								
Transfer	0.073***	0.075*	0.123**	0.054	-0.078**	-0.103***	0.046	0.079**
	(0.025)	(0.042)	(0.051)	(0.046)	(0.035)	(0.035)	(0.033)	(0.031)
Flood extent	0.207**	-0.085	-0.064	-0.043	0.191	0.133	-0.106	-0.108
	(0.104)	(0.134)	(0.181)	(0.182)	(0.161)	(0.121)	(0.128)	(0.144)
R^2	0.04	0.09	0.09	0.09	0.13	0.09	0.04	0.11
Flood duration								
Transfer	0.071***	0.076*	0.123**	0.055	-0.080**	-0.105***	0.048	0.080***
	(0.025)	(0.042)	(0.050)	(0.046)	(0.035)	(0.035)	(0.032)	(0.030)
Flood duration	-0.003	-0.001	-0.001	-0.000	0.001	-0.001	0.002	0.000
	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)
R^2	0.04	0.09	0.09	0.09	0.13	0.09	0.04	0.11
Flood extent and duration								
Transfer	0.074***	0.075*	0.123**	0.054	-0.078**	-0.103***	0.046	0.079**
	(0.025)	(0.042)	(0.051)	(0.046)	(0.035)	(0.035)	(0.032)	(0.031)
Flood extent	0.318***	-0.067	-0.052	-0.044	0.199	0.195	-0.176	-0.131
	(0.112)	(0.143)	(0.192)	(0.180)	(0.171)	(0.125)	(0.138)	(0.155)
Flood duration	-0.004**	-0.001	-0.000	0.000	-0.000	-0.002	0.003	0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)
R^2	0.04	0.09	0.09	0.09	0.13	0.09	0.04	0.11
Controls	✓	✓	✓	✓	✓	✓	✓	✓
Union fixed effects	✓	✓	✓	✓	✓	✓	✓	✓
N	7563	8951	8941	8947	8950	6061	8950	8945

Notes: Covariates include age, gender, education level, household size, dependency ratio, house structure, UNFPA/FAO recipient status and land type. Union fixed effects are included. Standard errors are clustered at union level. * p<0.10, ** p<0.05, *** p<0.01

Table 1.28: Welfare effects over time - panel unions

	(1) Days without food in flood	(2) Child food consumption	(3) Adult food consumption index	(4) Adult protein consumption	(5) Adult food consumption score	(6) Life satisfaction
Round 2 (20 weeks)						
Transfer	-0.228**	0.057	0.272***	0.239***	0.268***	0.064
	(0.087)	(0.116)	(0.086)	(0.086)	(0.085)	(0.081)
N	1284	1057	1284	1284	1284	1283
R^2	0.12	0.09	0.14	0.12	0.15	0.13
Round 1 (10-12 weeks)						
Transfer		0.088***	0.090*	0.099**	0.069	0.160**
		(0.029)	(0.053)	(0.047)	(0.056)	(0.063)
N		5289	6258	6258	6258	6250
R^2		0.03	0.09	0.06	0.10	0.08
Controls	✓	✓	✓	✓	✓	✓
Union fixed effects	✓	✓	✓	✓	✓	✓

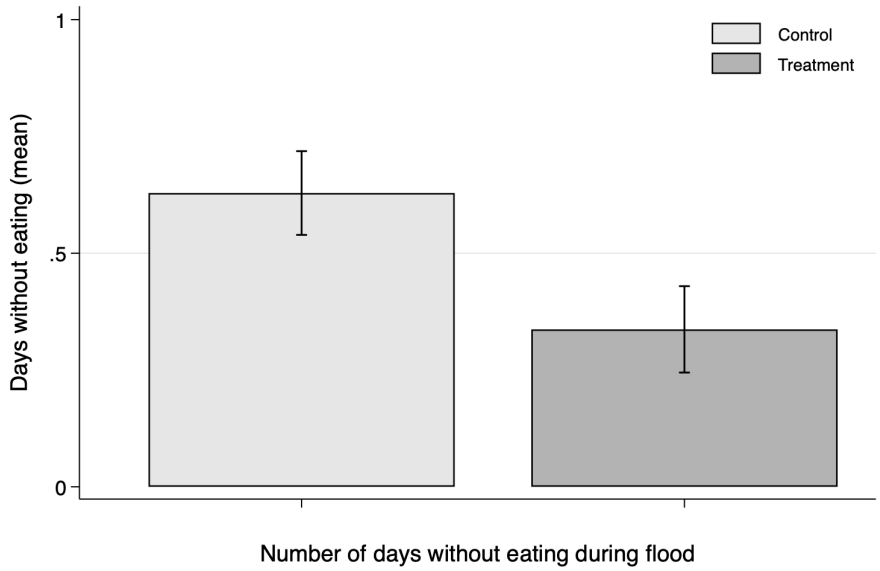
Notes: Covariates include age, gender, education level, household size, dependency ratio, house structure, UNFPA/FAO recipient status and land type. Union fixed effects are included. Standard errors are clustered at union level. * p<0.10, ** p<0.05, *** p<0.01

Table 1.29: Test for substitution of adult and child food consumption over time in panel control households

	Δ children meals yesterday			Δ children ate three or more meals		
	(1)	(2)	(3)	(4)	(5)	(6)
Δ adult meals yesterday	0.333***			0.268***		
	(0.037)			(0.040)		
Δ adult food consumption score		0.006***			0.003**	
		(0.002)			(0.001)	
Δ adult protein consumption			0.033***			0.014
			(0.011)			(0.009)
N	469	469	469	469	469	469
R^2	0.08	0.02	0.01	0.14	0.01	0.01

Notes: First difference model for panel control households that report child food consumption 10 weeks after flooding and 20 weeks after flooding. 93% of 659 control households in the second survey were also contacted in the first survey. 469 of these panel households (77%) report child food consumption. Positive significant coefficients indicate that changes in child and adult food consumption are complementary. Standard errors are clustered at union level. Note first differencing removes time-invariant confounders at household level. * p<0.10, ** p<0.05, *** p<0.01

Figure 1.10: Food security during the flood



Notes: 90% confidence intervals are shown. Covariates include age, gender, education level, household size, dependency ratio, house structure, UNFPA/FAO recipient status and land type. Union fixed effects are included. Standard errors are clustered at union level.

CHAPTER 2

Learning to See the World's Opportunities: The Impact of Imagery on Entrepreneurial Success

with **Nava Ashraf, Gharad Bryan, Alexia Delfino, Emily Holmes and Leonardo Iacovone**

Abstract We explore the importance of mental imagery for entrepreneurial success. An important skill for imagining the future, mental imagery is the process of mental experiencing and "seeing with the mind's eye", by accessing information from memory. Distressing past experiences can increase the cost of using imagery and undermine the ability to imagine the future in a balanced way, by inducing overly negative future imagery. We design a training curriculum to teach positive imagery in future decision making and test its impact on economic outcomes through a randomised control trial with roughly 2,000 would-be entrepreneurs in Colombia. When compared to a placebo entrepreneurship training, we find that imagery can be trained and those who receive the imagery training have significantly higher earnings, both before and after the start of the Covid-19 pandemic. For individuals with high levels of baseline trauma, imagery addresses their deficit in positive imagery. When we compare the imagery training to a no-intervention group, we find no results. Moreover, the placebo group have significantly lower earnings relative to the no-intervention group. Nevertheless, our results serve as a proof of concept that the ability to imagine the future more vividly and in a more balanced way is an important determinant of entrepreneurial success.

Acknowledgements *We thank Innovations for Poverty Action Colombia for their support in data collection and partnership management, and Moustafa El-Kashlan, Miguel Fajardo-Steinhäuser, George MacDonald and Misha Seong for excellent research assistance. We are grateful to Maria Paula Gómez Franco, Jose David Angel Castaño, Camilo Andres Hernandez Estrada, Mariana Medina and Manuel Pardo for their input in the curriculum design, training and supervision of the implementation team. We thank Stefan Dercon, Xavier Gabaix, Supreet Kaur, and David Laibson for very helpful comments. We gratefully acknowledge the financial support of IPA Peace and Recovery Program, J-PAL Post-Primary Education, J-PAL Jobs and Opportunity Initiative, the Rockefeller Foundation and the Institute of Global Affairs at the LSE. This RCT was registered in the American Economic Association Registry for randomised control trials under Trial number AEARCTR-0004695. It received official approval by the Innovation for Poverty Action Institutional Review Board USA (IPA IRB Protocol No. 14638) and the London School of Economics Research Ethics Committee (REC Ref No. 00769c).*

2.1 Introduction

Provision for the future makes no inconsiderable demands on our intellectual strength...The present always gets its rights. It forces itself upon us through our senses. To cry for food when hungry occurs even to a baby. But the future we must anticipate and picture. Indeed, to have any effect in the future, we must form a double series of anticipations. We must be able to form a mental picture of what will be the state of our wants, needs, feelings, at any particular point of time. And we must be able to form another set of anticipations as to the fate of those measures which we take at the moment with a view to the future. -Böhm-Bawerk, The Positive Theory of Capital and Interest (1889)

The ability to imagine the future has long been considered a possible foundation for inter-temporal decision making (Böhm-Bawerk, 1889; Loewenstein, 1992; Gabaix and Laibson, 2017). A myopic view of the future reduces the incentive to invest in the future. In contrast to a standard view of time discounting, however, the ability to imagine the future requires effort and practice to develop (Becker and Mulligan, 1997). An extensive body of work in neuroscience and psychology have focused on developing our understanding about this skill, in what is termed "mental imagery" (Holmes et al., 2019). Mental imagery is the process of mental experiencing or "seeing with the mind's eye" (visualising), by accessing information from memory (Pearson et al., 2015; Holmes et al., 2016). While humans think in words and images, images are faster to retrieve and has the potential to make imagined future scenarios more vivid, thus extending the time horizon over which decisions can be made (Holmes and Mathews, 2010). However, despite a vast literature on the importance of inter-temporal decision making, little empirical work has linked the ability to imagine the future with economic outcomes, such as income.

There is also large heterogeneity in how often and how well people imagine the future. By impacting the memory and thus the foundation for imagining the future (Schacter et al., 2007), distressing past experiences can significantly impede this capacity. Trauma induces an over-weighting of negative future scenarios and creates a deficit in the ability to imagine positive future scenarios (McNally et al.,

1995; Tulving, 2002; Moore and Zoellner, 2007; Kleim and Ehlers, 2008; Kleim et al., 2014). As a result, trauma increases the cost of using imagery in decision making and distorts the way in which people imagine the future. Exposure to trauma is common throughout the world, with over 70% of people estimated to have experienced at least one traumatic event in their lifetime (Benjet et al., 2016; Kessler et al., 2017). The prevalence of trauma is only likely to be exacerbated with the Covid-19 pandemic (Holmes et al., 2020). Shedding light on how imagining the future relates to trauma is not only policy relevant, but also offers a first step in developing our understanding about how the memory can impact future-oriented decision making.

This paper aims at building the link between the ability to imagine the future, trauma and economic outcomes. We first use standard psychology survey measures on the ability to imagine vividly different scenarios to show that a more frequent use and higher quality of mental imagery correlate strongly with higher asset accumulation, earnings, and life satisfaction. This motivates the importance of mental imagery in individual decision-making. Consistently with clinical evidence, we also find that high trauma individuals imagine substantially more vivid and emotional negative future scenarios, but not positive future scenarios.

We then move beyond correlations to ask whether imagery is malleable and whether an improvement in the use and quality of imagery used leads to an improvement in earnings. To speak to this, we design a training curriculum to teach positive imagery in future decision making and test its impact on economic outcomes through a randomised control trial with roughly 2,000 would-be entrepreneurs in Colombia. The ten-session curriculum teaches entrepreneurs to use imagine future scenarios and the pathways for achieving these outcomes with vividness and emotion, as well as mentally practice useful behaviours, such as saving. The curriculum is also specifically designed to build up the ability to produce imagery of positive future scenarios for entrepreneurs who have experienced past trauma, following programmes developed in clinical psychotherapy (Holmes

et al., 2007, 2019).[1]

The randomised control trial was conducted in partnership with the local government of Bogotá and randomly assigned participants to our imagery curriculum, a placebo business training, or a no-intervention control group. Our focus is on the comparison between the imagery curriculum and the placebo training, which we designed to have an identical structure, entrepreneurship content and delivery method. The only difference between the two is that the placebo training features written and verbal group-based activities (such as role play) instead of imagery exercises. We conduct two follow-up phone surveys 8 and 14 months after the end of the intervention and during the Covid-19 pandemic.

When comparing the imagery training to a placebo entrepreneurship training, we find that imagery can be trained: we observe a 0.18 standard deviations improvement ($p < 0.1$) on our imagery index measuring the frequency of imagery use, enhanced vividness (specificity) and emotional intensity of imagined future business scenarios. We also find that those who receive the imagery training have significantly higher earnings on average, both before and after the start of the Covid-19 pandemic. Relative to the placebo group, entrepreneurs report on average an increase of 0.19 and 0.14 standard deviations in earnings pre- and post-pandemic lockdown respectively. These findings are both statistically significant at the 5% level. Imagery participants report significantly higher business survival during the pandemic, but not before the pandemic. When we further unpack the heterogeneity in our results, we find that women are driving these improvements; women are also less experienced entrepreneurs and carriers of higher trauma burden at baseline. Moreover, we show that the imagery training differentially benefits individuals with high levels of baseline trauma symptoms by addressing their deficit in positive imagery and mediating negative imagery. Indeed, we find evidence that these effects translate through to higher earnings, although we fail to

[1] In clinical psychotherapy, mental imagery has also been used to rescript negative intrusive memories, in addition to encouraging positive imagery (Holmes et al., 2007, 2019). However, rescripting negative memories requires specialised clinical skills, which do not lend itself to scalable policies that can be delivered by lay trainers. Consequently, we focus on the development of positive imagery as a low-cost policy lever that could be used to complement for negative imagery, rather than directly address the negative imagery itself.

reject the hypothesis that there is a statistically significant difference with the low-trauma group.

Finally, we compare the imagery training to the no-intervention group to assess the impact of business training in general and find small to no differences across all outcomes. Relative to the no-intervention group, the placebo control group report lower earnings before the start and during the pandemic on average. There are several reasons that might explain these results. First, we observe differential attrition in the no-intervention group in the follow-up surveys during the pandemic, whereas balance is maintained across the two training arms. Although there is limited evidence that attrition may be selective on observable characteristics, there is suggestive evidence that there may be differential attrition based on unobservables. Using Lee (2009) bounds, we find evidence that the treatment effects of the imagery training compared to the placebo group are bounded away from zero, providing reassuring evidence that the imagery training has a positive impact relative to the placebo training. In contrast, the results arising from the comparison between both training groups and the no-intervention group are not robust to making weak assumptions on non-respondent outcomes. Second, biased reporting in the surveys may have caused the no-intervention group to diverge from the two training arms. Finally, the placebo training that drew on standard training curricula may have had a negative impact on our sample of entrepreneurs. In this case, imagery training served to mitigate the negative effects of the placebo training. We currently do not have enough data to pin down the exact reason for the negative outcomes observed in the placebo group compared to the pure control, although we believe the first and third reasons to be most likely. Even in the worst case scenarios, our conclusion that imagery matters for economic decision making still holds.

Our paper speaks to three bodies of literature. First, our results highlight a critical role for the ability to imagine future scenarios and the future consequences of current actions more vividly, contributing to the literature advancing micro-foundations for inter-temporal decision making. Becker and Mulligan (1997) and Gabaix and Laibson (2017) have made the theoretical case for the importance of

imagining the future as a means to micro-found inter-temporal discounting. However, only two papers in economics have empirically tested whether visualising the future affects economic choices. In a primary school setting, Alan and Ertac (2018) show that students who learn to imagine their future selves and to build counterfactual scenarios in their minds make more patient inter-temporal choices in incentivised experiments and perform better in the classroom. A recent working paper by John and Orkin (2021) compares the impact of imagery with a standard planning intervention and a pure control to encourage the take-up of water chlorination. For their primary behavioural outcome of interest, the presence of chlorine within household drinking water, they find large effects of the imagery treatment compared to the control group, although these effects are not statistically distinguishable from the planning training after 12 weeks. In other work in psychology, imagining a future version of the self has been found to increase savings (Hershfield et al., 2018). In contrast to these papers, we explore the role of imagery in making forward-orientated economic decisions and measure the impact of imagery on a wider range of economic outcomes, such as earnings and business survival.

Our results also contribute to the literature on the long-lasting effects of trauma and conflict by showing that the ability to imagine the future in a more balanced way is an important determinant of entrepreneurial success. Temporary shocks have already been shown to shift social, risk and time preferences in the longer term. For example, exposure to violence has been shown to result in a change in risk aversion (Voors et al., 2012; Callen et al., 2014), most pertinently in a Colombian context where the decades-long civil conflict has affected a large share of the population (Moya, 2018). In this paper, we explore another channel through which past traumatic experiences can affect economic choices: the ability to imagine the future, by impacting memory. The role of memory in limiting decision making and imagining the future has been explored theoretically in economics (Mullainathan, 2002; Gennaioli and Shleifer, 2010; Bordalo et al., 2016, 2018). Although we cannot draw a direct link between memory and future imagery, we show how an imagery training can rebalance the ways in which past experiences affect future-oriented decision making, by mediating overly negative future imagery and increasing the quality of positive future imagery.

Lastly, we apply a structured imagery programme to the domain of entrepreneurship and showcase another example of a business training focused on promoting soft skills. Efforts to train entrepreneurs in general have had modestly positive effects on average (McKenzie, 2021). There is, however, increasing evidence that using lessons from psychology to promote “soft” skills could enhance the effectiveness of entrepreneurship training for subsistence-level enterprises (McKenzie et al., 2021). For instance, a personal initiative training advocating a “proactive" mindset in Togo increased micro-enterprise profits by 30% on average over two years (Campos et al., 2017), although the long-term effects were not replicated in other contexts (Alibhai et al., 2019; Ubfal et al., 2019). Business training programmes for the youth have also combined soft and hard skill development, with success in promoting self-employment (Alaref et al., 2020; Chioda et al., 2021). Similarly, we postulate that entrepreneurs require the skill to imagine the future: imagine new products, customers, competitors – "to form an adequate picture of the forms which goods will take, of the quantity of them, and of the time when they will come to maturity as result of those productive or commercial activities which we are now commencing", as Böhm-Bawerk (1889) describes. We demonstrate that combining hard entrepreneurship skills training with this soft skill can improve earnings in the short term. With aid agencies and governments spending more than a billion US$ on entrepreneurship training annually (McKenzie et al., 2021), integrating psychological methods into these programmes offers a clear pathway to impact at scale.

The next section presents a short primer on mental imagery and trauma, and a motivational framework for conceptualising how imagery shapes future-oriented thinking. Section 2.3 presents correlational evidence to motivate that imagery matters for economic outcomes. We describe our curriculum innovation in Section 2.4. Section 2.5 outlines our experimental design. We summarise our approach to measurement and empirical strategy in Sections 2.6 and 2.7 respectively. We present results in Section **??**, while Section 2.10 concludes.

2.2 A primer on imagery, future thinking and trauma

2.2.1 Evidence from neuroscience and psychology

We define imagery as a process of mental experiencing or visualising resulting in an experience akin to "seeing with the mind's eye" (Holmes et al., 2016). While humans think either words or images, images are faster to retrieve and offer a heightened sense of realism. This heightened sense of realism occurs through two channels: increased specificity or vividness and increased emotive content of the imagined scenarios. From a neuroscience perspective, the perceptual equivalence between images and reality arises from the common activation of brain systems that are involved in perception, such as the amygdala, early visual cortex, frontal and parietal areas (Holmes and Mathews, 2010).[2]

Yet imagery is not used uniformly across individuals; rather, its use varies, along two dimensions, with significant implications for decision making (Nelis et al., 2019). First, there may be variation in how much mental imagery is used.[3] Second, the content of mental imagery is not homogeneous and there exists stark differences in the nature of the projections it gives rise to. In turn, variation in use and content has prompted questions of whether imagery should be thought of as an innate trait or a malleable skill that can be trained.

An individual's past experiences of trauma may be a driver of both types of variation. Research demonstrates that imagining the future and reflecting on the past use the same neural machinery (Schacter et al., 2007). Put differently, we draw on our memory to imagine the future.[4] Unsurprisingly, therefore, distressing circumstances which bring about trauma can influence future mental imagery both in

[2]Mental imagery is multi-sensory, leading it to act as an "emotional amplifier" as termed in psychology. Physiologically speaking, this is the result of mental images activating brain systems underlying emotion more directly than symbolic representations that do not use sensory prompts (Holmes and Mathews, 2010). Lab experiments have demonstrated imagery's more powerful impact on emotions than verbal cognition (Holmes et al., 2016).

[3]Less than 1% of the population who are otherwise healthy are estimated to have aphantasia, which is the inability to produce visual imagery (Zeman et al., 2015; Pearson, 2019). However, this condition typically applies to the visual domain, so this small subset of the population may still be able draw on other senses during mental imagery.

[4]Neuroimaging studies, including fMRI scans that look at blood flows to the brain to detect areas of activity, find that the same neural machinery used to store information is activated during mental simulations about the future and perspective of others (Schacter et al., 2007).

terms of the quality and nature of imagery. Trauma may decrease specificity or hijack mental projections in such a way that all mental images become characterised by only negative outcomes. A substantial literature supports the idea of trauma inducing either an overly negative or "over-general" personal memory (McNally et al., 1995; Moore and Zoellner, 2007; Kleim and Ehlers, 2008). Indeed, evidence suggests that those suffering from post-traumatic stress disorder (PTSD) experience decreased specificity, compared to those without the condition. However, this decreased specificity is present in response to positive, but not negative, cues suggesting that positive imagery is especially hampered by trauma (Kleim et al., 2014). Furthermore, those who have experienced past trauma may struggle to use imagery at all, with studies showing deficits in memory (Douglas Bremner et al., 1995; Jenkins et al., 1998; Buckley et al., 2000) and hence the ability to mentally project into the future.[5] When imagery involves reliving distressing circumstances of the past due to intrusive negative imagery, people may avoid imagining altogether. In many ways, deficiencies in terms of the content and the use of imagery, brought about by trauma, cannot be disentangled. Collectively, they result in a sense of a "foreshortened future", whereby individuals struggle to project themselves into future experiences.[6]

Mental experiencing has been shown to be closely linked to decision making. Vivid and emotional images can be motivational and induce behavioural change, more so than verbal descriptions of the same scenarios (Mathews et al., 2013; Pearson et al., 2015). Specifically, imagery can be used to identify a concrete and emotionally motivating future outcome and then mentally map out the possible pathways towards achieving that outcome, both in terms of the steps and potential obstacles. Empirical evidence from psychology supports the notion that imagining one's own future behaviour can increase the likelihood of enacting that behaviour in reality (Gregory et al., 1982; Koehler, 1991; Conway et al., 2004; Rutchick et al.,

[5]Changes to important brain areas, such as the hippocampus, amygdala, and medial prefrontal cortex, that support the memory function have been shown to occur (Bremner, 2006).

[6]A similar phenomenon occurs in depression. Imagery of past negative events and suicidal imagery of the future coexist in depressed individuals (Holmes et al., 2007). Imagining predominantly negative scenarios may lead to persistent negative mood, which hinders active behaviour. Moreover, depressed individuals are more likely to avoid imagining altogether in the attempt to stop negative intrusive memories, even when compared to non-depressed individuals with a similar frequency of intrusive memory (Newby and Moulds, 2011).

2018; Renner et al., 2019). For instance, imagining the process of voting has been shown to increase voter turnout (Libby et al., 2007). In an economics domain, an imagery intervention that encouraged women to visualise the future increased take-up of water chlorination and reduced child diarrhoea (John and Orkin, 2021). The frequency, speed and emotional intensity with which a person can imagine certain scenarios also increases the subjective probability that the scenario will occur, through a form of an "availability heuristic" (Carroll, 1978), in turn also driving behavioural change. Naturally, deficits in the ability to generate vivid and emotional mental images can hinder decision making, especially for those who have experienced past trauma. For instance, feeling anxiety or fear, in response to a threatening or negative image, increases the decision maker's awareness of the danger involved in that scenario (Arntz et al., 1995; Forgas, 1995). The subjective probability attached to the feared outcome increases avoidant, and ultimately sub-optimal, behaviour.

To improve future-oriented decision making, therefore, requires rebalancing the content of future imagery among those who have experienced past trauma. Previous efforts to improve future-oriented decision making have operated in non-economic domains, while those focused on addressing the effects of trauma have lacked scalability. For example, mental imagery is already frequently adopted by professional athletes and musicians to mentally rehearse and avoid physical injury from over-practice, thereby reducing performance anxiety (Driskell et al., 1994; Munroe et al., 2000; Osborne et al., 2014).[7] On the other hand, in clinical psychotherapy, imagery has been used to re-script negative intrusive memories, in order to lessen their emotional content or build up the ability to produce positive imagery as a substitute for negative imagery (Holmes et al., 2007, 2019). The program required specialised clinical skills leading to high unit costs. This paper extends the application of this skill to the domain of business decision making while ensuring it remains a low-cost, scalable policy lever.

[7] "Repeatedly imagining a finger movement sequence (as in piano playing) not only improves performance but there is a corresponding change in motor cortex. Thus, imagery selectively activates those areas involved in processing related sensory information in reality, or in producing similar responses and actions" (Holmes and Mathews, 2010, p. 351).

Teaching mental imagery within a given domain - entrepreneurship - raises the question of whether we expect participants' learning to be domain-specific or to be extended more broadly into other decision-making areas. A substantial literature in developmental psychology has drawn the distinction between domain-general (Li et al., 2014) and domain-specific (Siegler, 2006) theories of learning and development. While we acknowledge that mental imagery has the potential to be domain-general, our intervention focuses on how imagery can be effective when applied to specific business-related scenarios. A few studies in the entrepreneurship literature have explored whether mental imagery can be linked to better business decision making. Anthony et al. (1993) find that mental imagery can be used to enhance long term strategic thinking whilst Neck et al. (1999) find that mental imagery can aid organisational leadership.[8] We take an agnostic approach in the paper on whether our imagery training is effective beyond the business domain. We let the data speak to this point by collecting measures on people's mental imagery in both business-specific and non-business scenarios.

2.2.2 A motivating theoretical framework on imagery and future thinking

In this section, we present a simple motivating framework to conceptualise how imagery can shape future thinking and the role of trauma in undermining this ability.

The key intuition is that mental imagery makes the downstream consequences of current and future actions more concrete and thus can improve choices in the pathway to achieving future outcomes. Consider a perfectly rational and risk-neutral entrepreneur making a choice between actions A and B. A has a known utility outcome of U_A and it is the default option. B has either a high utility of U_B^H or a low utility of U_B^L, but the true utility is unknown to the entrepreneur. We assume that $U_B^H > U_A > U_B^L$. The entrepreneur holds a prior belief over the

[8]The study by Zaleskiewicz et al. (2020) compares imagery ability among entrepreneurs and non-entrepreneurs, finding that there are no differences in the use of non-business imagery between the two groups. However, entrepreneurs in their study are more frequent users of business-related imagery which, in turn, is correlated with increased risk taking and improved economic outcomes. Evidence of this kind suggests that entrepreneurship may be a particularly effective domain in which to focus the use of imagery.

probability of obtaining a high utility outcome from B, represented by γ. The expected utility of B is given by $\gamma U_B^H + (1-\gamma)U_B^L < U_A$. The inequality follows from the notion that A is the default action in the status quo.

Mental imagery reveals the true utility with probability λ at a small mental cost of C. In other words, the entrepreneur can exert mental effort to experience each option and, thus, get a better idea of their utility. In the extreme case of perfect introspection, the entrepreneur will choose B in the high utility state and A if B has a low utility. Even if mental imagery imperfectly captures reality, imagery 'demystifies the future' by adding an additional level of specificity and emotion to future outcomes, and thus increases expected utility prior to imagery as follows:

$$\lambda(\gamma U_B^H + (1-\gamma)U_B^L) + (1-\lambda)(U_A) - C \underbrace{>}_{small\ C} U_A$$

In other words, when the mental cost of imagery is relatively low, the expected utility of deciding between A and B after imagining is higher than simply choosing the status quo.

Trauma undermines the ability to imagine the future through two channels. Firstly, trauma can increase the cost of mental imagery C if future imagery is overly negative or suppressed, such that $C_T > C$. Secondly, trauma can induce an over-general memory and hence non-specific future imagery. In other words, trauma can reduce the quality of mental imagery and thus lower the probability with which the true future utility is revealed by imagery, so $\lambda_T < \lambda$. As a result, some people who experienced past trauma might choose not to use imagery or if they do, the imagery could be overly negative or non-specific.[9]

A training that promotes positive mental imagery serves to reduce the mental cost of imagery C and increase the probability λ that imagery reveals the true utility of future outcomes. As a result, training can induce a greater use and quality of imagery. On quality, the training can help individuals to increase the level of specificity and emotional intensity of images about future outcomes.

[9]We do not model the asymmetry between imagery of positive or negative scenarios. One way to do so would be to define two different weights, λ_1 for positive scenarios and λ_2 for negative scenarios, with $\lambda_2 > \lambda_1$ for a traumatized decision maker.

2.3 Correlations between imagery and economic outcomes

What is the impact of imagery on economic outcomes? Using data on a sample of nearly 2,000 micro-entrepreneurs, we approach this question in three steps.[10] First, we show that mental imagery is correlated with better economic outcomes to motivate the importance of imagery in economic choices. Second, we ask whether imagery is malleable and can be taught through a training programme using a randomised control trial. Third, we explore whether an improvement in the quality and frequency of imagery leads to an improvement of economic outcomes, especially for entrepreneurs with higher baseline trauma.

Upon investigation of the importance of mental imagery in determining economic outcomes, we document four new empirical observations.

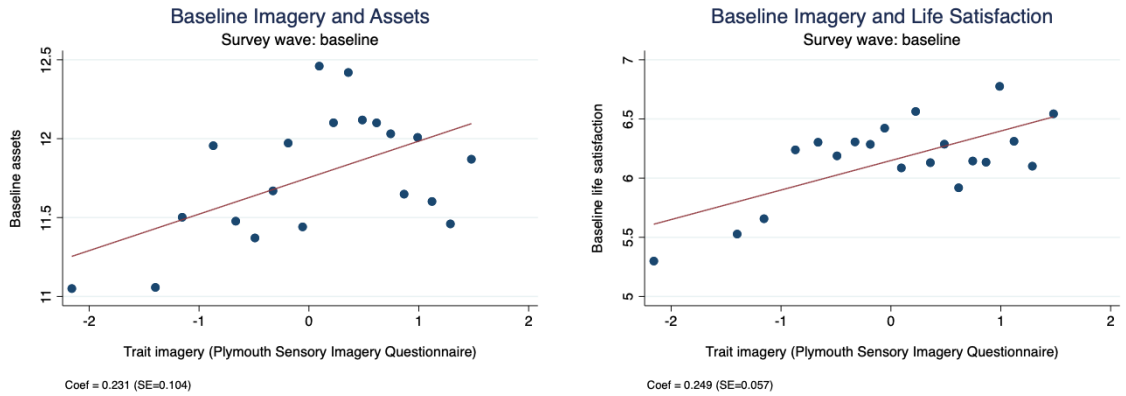
> *1. People who have a higher capacity to use imagery invest more in durable goods and are more satisfied with their life.*

We hypothesise that higher quality mental imagery as an innate trait should be related with better long-term outcomes, such as asset accumulation and life satisfaction. We measure trait mental imagery using a contextually-adapted version of a well-known psychometric scale, the Plymouth Sensory Imagery Questionnaire or Psi-Q (Andrade et al., 2014). Akin to a personality trait, the Psi-Q is intended to capture a cognitive trait and in particular, the extent to which a person's use of imagery is richly developed. We show that trait mental imagery is positively related with both outcomes that typically unfold over time for our entire sample at baseline.

The left panel of Figure 2.1 shows that trait mental imagery (on the x-axis) is positively correlated with the total count of household assets (on the y-axis), both measured at baseline and controlling for age, gender, income categories, entrepreneurship status, and the time period and group in which the respondents participated in the programme. Including these controls allows us to test whether a

[10] Section 2.5 provides further details about the data and sample.

Figure 2.1: Baseline imagery, assets and life satisfaction



Notes: The figure shows the correlation between trait mental imagery and total count of household assets (left panel) and life satisfaction (right panel). Trait imagery is measured using the Plymouth Imagery Scale Questionnaire and life satisfaction using a Cantril Ladder with ten steps. Asset count is the total number of assets owned by a respondent's household out of a list of different asset types (e.g., TV, computer). All these measures are collected in our baseline survey prior to the intervention. The imagery index is standardised to be mean zero and unitary standard deviation in the no-intervention group. Regressions control for gender, age, business status and income category at baseline, and subdivision-wave interaction fixed effects.

correlation exists between imagery and wealth within groups of people who share these same characteristics. Quantitatively, the correlation between imagery and asset count is small, but statistically significant: an increase of one standard deviation in the Psi-Q score is correlated with an increase in assets of 0.23 ($p < 0.05$), which corresponds to a quarter of a television or a mobile phone, the latter being an asset owned by most of the sample.

The right panel of Figure 2.1 shows that trait imagery (on the x-axis) is positively correlated with self-reported life satisfaction using the Cantril ladder (y-axis), both measured at baseline and controlling for the same variables as above. More precisely, the scatter plot shows that going from one standard deviation below to one standard deviation above the mean in the Psi-Q score is correlated with an increase in life satisfaction of approximately half a step along the Cantril ladder ($p = 0.000$).

2. Individuals who use mental imagery with higher frequency and quality over a short time period achieve better earnings in the same time frame.

We also assess whether the use of mental imagery is correlated with better

economic outcomes in the short term. We build an imagery index that combines the extent to which people report using imagery in daily life by spontaneously producing images in their minds (Spontaneous Use of Imagery Scale or SUIS (Nelis et al., 2019)) and the vividness and emotional intensity with which respondents report imagining pre-defined future scenarios (Prospective Imagery Task or PIT (Holmes et al., 2008)). The index thus measures both the frequency and quality with which individuals use mental imagery in the immediate term (i.e. over a few months), rather than an innate trait that would be invariant over time. Figure 2.2 shows a positive correlation between maximal earnings reported in the six months before the first national lockdown imposed in response to the Covid-19 pandemic, and the use and quality of imagery measured three months after the start of the lockdown.[11] We restrict our sample to respondents that only received the placebo training[12], and again control for gender, age, business status, time period and group of programme participation.

We find that an increase of one standard deviation in the imagery index is related to an increase of approximately 0.2 standard deviations in the earnings index ($p = 0.001$). As a benchmark, this effect is comparable to the average impact of business training programmes that focus on teaching 'soft' or psychological skills, which typically amounts to approximately 10% on business sales in low- and middle-income countries (McKenzie, 2021; McKenzie et al., 2021). Overall, Observation 1 and 2 demonstrate that there is good reason to believe that the ability to use mental imagery is relevant for economic outcomes and wellbeing.

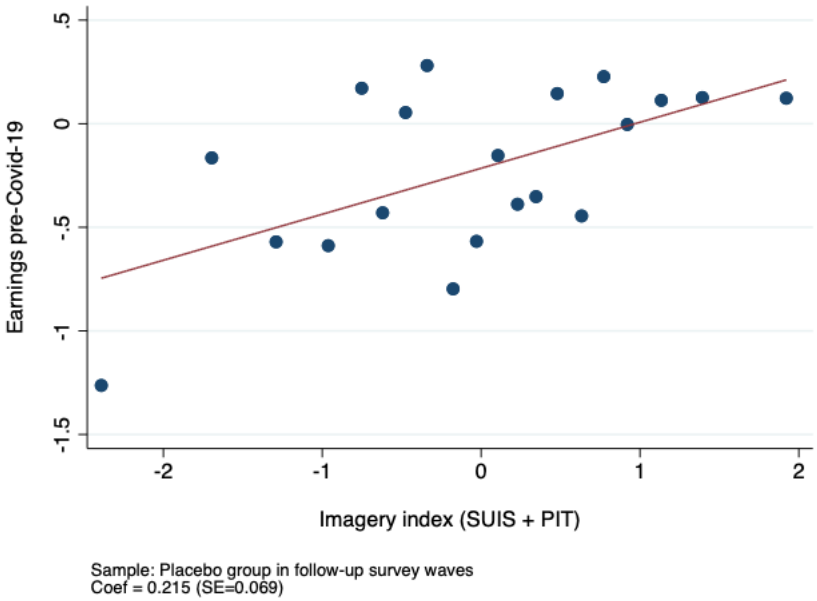
> *3. Attaching a higher subjective probability to mental imagery of future scenarios is positively correlated with realised earnings.*

Our third observation pertains to a particular channel through which mental imagery should improve inter-temporal decision making: by increasing the subjective probability attached to future scenarios. Figure 2.3 reports the correlation

[11] The earnings index includes both income and business sales and it is standardised to be mean zero and unitary standard deviation in the no-intervention control group. For more detail on the construction of this index, see Section 2.6 for more detail.

[12] If we restrict our sample to those who received no training at all, we find consistent results for Observations 2, 3 and 4, although measured with slightly more noise, due to the smaller sample size.

Figure 2.2: Imagery and earnings pre-Covid-19



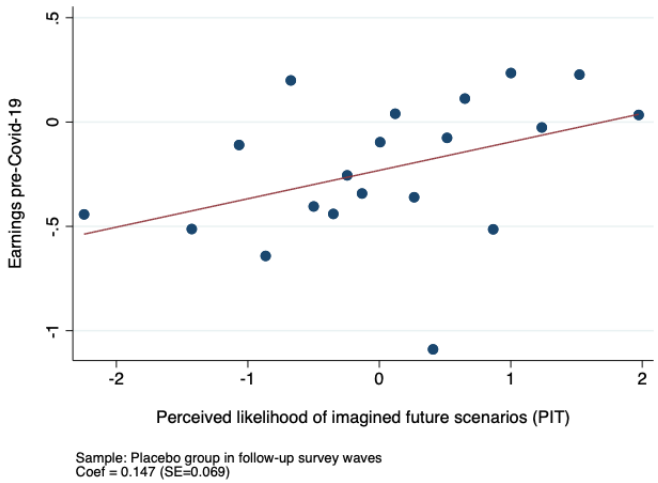
Notes: The figure shows the correlation between imagery and an index of earnings in the six months prior to the first national lockdown in response the Covid-19 pandemic. The imagery index combines the Prospective Imagery Task (PIT) and the Spontaneous Use of Imagery Scale (SUIS) for general scenarios. The index of earnings pre-Covid-19 combines business sales (conditional on business survival) and take-home income of the respondent. Both measures are collected in our first follow-up survey and are standardised to be mean zero and unitary standard deviation in the no-intervention group. We restrict the sample to respondents that received placebo training. Regressions control for gender, age, business status and income category at baseline, and subdivision-wave interaction fixed effects.

between an individual's subjective belief that their mental imagery of pre-defined future scenarios will be realised (using the PIT scale) and the pre-Covid-19 realised earnings introduced above. We find that an increase of one standard deviations in the likelihood attached to the imagined scenarios increases realised earnings by 0.14 standard deviations ($p < 0.05$).

Observations 2 and 3 are both consistent with the idea that imagery may improve decision making and economic performance by making the future more vivid and more easily available in one's mind eye. Observation 2 highlights that this greater availability is achieved through higher vividness and emotional arousal when imagining future scenarios, while Observation 3 suggests that individuals rationally update their subjective probability that certain future events will

happen, leading to higher earnings. However, most of the economic studies on the relationship between beliefs about the future and economic choices focus on the subjective likelihood that people attach to positive relative to negative scenarios (Eil and Rao, 2011; Brunnermeier and Parker, 2005). In contrast, the index of subjective probability in Figure 2.3 averages the perceived probability across positive and negative scenarios, suggesting that mental imagery increases the realism of the future beyond through the sole channel of heightened optimism.

Figure 2.3: Subjective probability of imagined scenarios and earnings in the months before Covid-19



Notes: The figure shows the correlation between the subjective probability attached to both general scenarios imagined in the Prospective Imagery Task (PIT) and an index of earnings pre-Covid-19, which combines sales (conditional on business survival) and take-home income. Both measures are collected in our first follow-up survey. They are standardised to be mean zero and unitary standard deviation in the no-intervention group. We restrict the sample to respondents that received only the placebo training. Regressions control for gender, age, business status and income category at baseline, and subdivision-wave interaction fixed effects.

4. Stronger symptoms of trauma are positively related with negative future imagery.

Finally, we document that trauma can cause a distortion in the way in which people imagine future scenarios, by creating disproportionately powerful negative future imagery, often to the detriment of positive future imagery. In Figure 2.4, we examine the relationship between trauma symptoms and the self-reported

quality of positive and negative future scenarios imagined using the PIT scale. We measure symptoms of trauma using the Impact of Event Scale-Revised, which is a 22-item self-reported measure that assesses subjective distress caused by past traumatic events. Imagery quality is represented by an index that combines vividness, emotional intensity and likelihood of the imagined non-business scenario occurring in real-life from the PIT scale. For example, we asked participants to imagine that the Covid-19 pandemic is over and they are struggling to make ends meet (negative future scenario) or they have saved enough money to buy an asset they really want (positive future scenario). We again restrict the sample to the placebo group.[13]

The left panel of Figure 2.4 shows that individuals with higher trauma symptoms report significantly more vivid and emotionally intense negative future scenarios using the PIT scale (slope of 0.014; $p = 0.000$), but there is no relationship between trauma symptoms and positive future scenarios (slope of -0.004; $p = 0.219$), as shown in the right panel.

2.4 Intervention

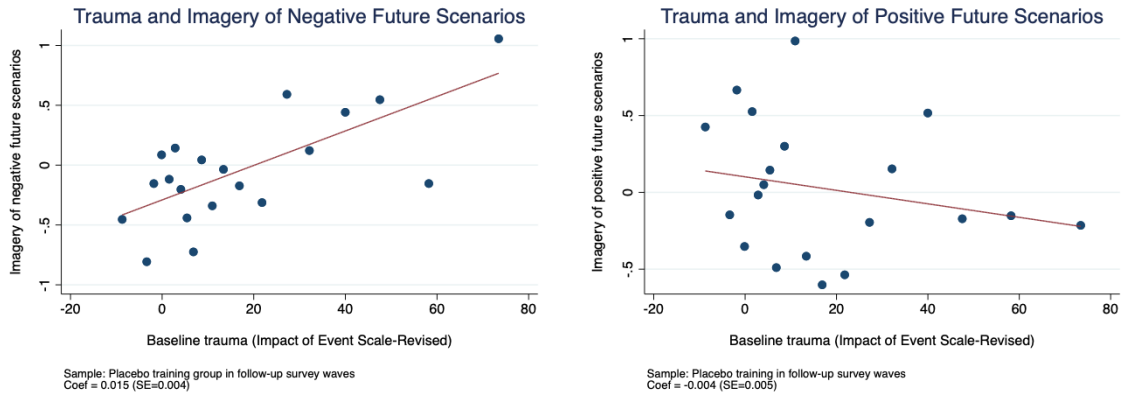
2.4.1 Description of intervention

To test whether mental imagery can be taught and whether it improves economic outcomes, we designed an entrepreneurship training overlaid with imagery techniques (termed 'imagery curriculum') and compared its impact to a traditional entrepreneurship training (termed 'placebo curriculum').[14] We designed the imagery and placebo curricula to be as similar as possible in terms of structure, content and delivery. Both programmes were implemented by the local government of Bogotá,

[13]We focus on non-business scenarios so that we can include the entire sample, as business scenarios were only asked of respondents with a business at the time of the survey.

[14]We conducted extensive qualitative data collection with marginalised entrepreneurs and victims of conflict to inform the curriculum content and language. In order to adapt and refine the sessions to the needs of the target population, we piloted the curriculum twice in November 2018 and May 2019 with a total of 60 entrepreneurs. We also obtained ongoing feedback throughout the curriculum development process from various government ministries that were mandated to develop and deliver programmes that benefit vulnerable populations. In particular, we worked closely with the lead psychologist team at the National Victims Unit (Unidad para las Victimas), Alta Consejería for Victims' Rights within the local government, and the National Training Service (Servicio Nacional de Aprendizaje or SENA).

Figure 2.4: Quality of positive and negative imagery by baseline trauma



Notes: The figure shows the relationship between imagery quality and baseline level of trauma for defined positive and negative future scenarios. Baseline trauma is measured using the Impact of Event Scale-Revised, which is a 22-item self-reported measure that assesses subjective distress caused by past traumatic events. Imagery quality is represented by an index that combines vividness, emotional intensity and likelihood of the imagined scenario occurring in real-life from the PIT scale. Both measures are collected in our first follow-up survey. They are standardised to be mean zero and unitary standard deviation in the no-intervention group. We restrict the sample to respondents that received only the placebo training. Regressions control for gender, age, business status and income category at baseline, and subdivision-wave interaction fixed effects.

specifically the District Department of Social Integration (SDIS) and District Department of Economic Development.

Both curricula took the form of ten three-hours sessions covering the same entrepreneurship themes in identical order. We structured the ten sessions to follow the journey of an aspiring entrepreneur, whereby each session centres on a particular topic deemed relevant from extensive qualitative interviews with our population of interest. These themes include product development, customer experience, marketing, competition, savings, accessing finance, productivity and managing employees.[15] At the beginning of each session in both curricula, we motivated the relevance of the content through real-life stories and relatable examples.

Both treatment arms also had an identical delivery format. The curricula were implemented in weekly sessions with participants assigned to a particular time slot. Each session had a class size of between 15 and 25 participants and venues were located in the same neighbourhoods across Bogotá. Two trainers led each

[15]The curricula drew on materials from existing entrepreneurship training programmes, including the International Labour Organization's Start and Improve Your Business programme.

session, with one trainer experienced in psychology or social work and the second in entrepreneurship. All trainers were recruited through the same call using a well-defined selection criteria. Moreover, to minimise the impact of trainer heterogeneity, trainers were required to follow a standardised scripted manual and presentation deck prepared by the research team. Trainers received over three weeks of training. They also rehearsed before each week's session in group workshops led by four supervisors who had co-designed and piloted the curricula together with the research team. For all activities, trainers were kept separate by treatment arm. Participants received a certificate if they attended six or more sessions.

The key feature that distinguishes the imagery curriculum from the placebo curriculum is the inclusion of three to four imagery exercises in each session. In the placebo curriculum, we replace the imagery exercises with group discussions, role play and written work of the same content time length. For example, in the imagery curriculum, participants were encouraged to imagine their product or service in detail and how they felt towards the product, before imagining the day in the life of their business and noting the emotions that arise. In contrast, participants in the placebo curriculum were asked to think about their product and write down the ways in which their business idea matches their passions and skills.[16] The following subsection describes the imagery exercises in more detail.

2.4.2 Teaching imagery

As highlighted by Böhm-Bawerk, imagery is like a muscle. While individuals might differ in the extent to which they use imagery naturally (a physical endowment), it ultimately takes costly effort to learn how to visualise the future. We designed our imagery curriculum with the goal to teach imagery in business decision making and to make subjects aware that this is a portable skill that needs to be trained and strengthened with practice. In particular, we exemplify how imagery might be used to increase the specificity of future scenarios and goals, and

[16] As a second example, participants in the imagery treatment were asked to imagine their product or service from the shoes of their target customer in the third session and empathise with the customer to determine whether the product satisfied their customers' needs. In contrast, participants undertaking the placebo training filled out a table that summarised the characteristics and needs of their customer and then list the ways in which their product fulfilled these needs. Participants subsequently discussed their tabulated response with another participants.

thus help entrepreneurs make business decisions with medium to long-term consequences (e.g., when choosing the product to sell, when asking for a loan in a bank). The curriculum seeks to demonstrate the value of imagery in various ways, including showing how performance might be improved with the use of imagery within the sessions and linking the exercises to real-life behaviours and outcomes. Moreover, providing a safe space to practice imagery in a guided way should also reduce the cost of future imagery, which is particularly fitting for participants for whom imagery may be unpleasant due to the lack of specificity in imagery or involuntary reminders of a traumatic past. Under the common goal of teaching imagery as a portable skill, imagery activities fall into one of three categories: (1) thinking about the future, (2) mental practice and (3) adopting the perspective of others.

The first category of exercises required participants to imagine future scenarios, different counterfactuals and think through the downstream consequences of their actions. In doing so, the imagery exercises encouraged participants to become more specific about these future scenarios. The future scenarios encompass both those that arise as a result of behaviours taken in the present (such as savings) or plausible scenarios that may occur beyond one's control. In many exercises, participants are asked to set a goal, which ranged from a big picture aspiration to concrete weekly goals. Once the goal had been defined, participants were encouraged to use imagery to pave concrete steps for achieving these goals.[17] For instance, we designed a session on savings, after learning that many entrepreneurs shared the view that saving was perceived to be impossible for them during our qualitative study. In this session, participants first performed basic accounting to establish the minimum amount of money they could realistically commit to saving every week and then imagined their savings contribution accumulating over weeks, months and years to appreciate the long-term power of small habits. To motivate the importance of savings, participants subsequently visualised the consequences of experiencing a personal negative shock in the instance where they

[17]Our extensive qualitative study highlighted the need to concretise the pathway of actions for our population of interest. Many victims of conflict who aspired to start a business envisioned exporting their products without being able to describe how they plan to make their first sale.

had saved versus not.[18] Moreover, they established a use for their savings and developed a simple savings strategy that entailed a set of actions triggered in response to a pre-identified reminder.

By becoming more specific about future scenarios through imagery, the curriculum also sought to enable participants to simulate the utility they attach to these future scenarios. Imagery can produce a strong emotional response in participants, which carries information about the utility associated with the content of the imagery. Consequently, in the second session, we motivate the value of imagery by asking participants to decide whether they would like to start a business as a fishmonger. After making a provisional decision, participants are asked to imagine the day in the life of a fishmonger in the most intimate details, including smelling the fish upon arriving at the fish market. Details are provided in the following excerpt:

> *Now, I will take you through a day in Julián's life. Your alarm clock rings at 3 am. It is still dark and cold. You travel in your truck for an hour to get to Paloquemao's market. You arrive to the fish market. You can see and smell fresh fish piled up on top of the tables, with noisy salesmen. Think about this scene for a moment. I would like for you to paint an image or a video in your head, as if you were an artist or a film director trying to capture this moment. What do you see? What do you smell? How do you feel right now?*

Participants then imagine describing their day and business to a relative or friend and noting how they feel during this conversation. Upon completing the exercise, participants are asked to reflect on what imagery has taught them and reconsider whether they would like to become a fishmonger after all. In contrast, the

[18] The following excerpt showcases how imagery is used to imagine a counterfactual future scenario:

> *Imagine your entrepreneurial business works well and you receive a constant flow of profits. You feel satisfied and use your profits to cover personal expenses that you could not afford before. As you pay all these expenses, you forget to save. Six months later, you discover that one of your relatives or close friends gets sick. You need an emergency treatment that is awfully expensive. But you realize that you have no savings. What do you do? How do you feel in this situation? Is this a situation you could imagine happening in the future? How would this event affect your life in one year from now?*

placebo curriculum did not place any emphasis on emotion and instead encouraged participants to rationalise their utility by writing lists. During this particular exercise, participants were asked to write down the ways in which becoming a fish monger matched or did not align with their passions and skills, before discussing these reflections in pairs.

If the first category of imagery exercises can be thought of as experimentation between different pathways of action and identifying the utility associated with those actions, the second category reinforces these insights through mental practice. In other words, participants use imagery to practice a plan chosen through mental trial and error. While mentally practising their plan in a safe space, participants identify potential obstacles and how to overcome them. Mental rehearsing should improve the concreteness and realism in their plan and thus increase their propensity to choose the better option more frequently. For example, participants visualise making an origami step-by-step before attempting in reality. We intend for this illustrative example to highlight how practising plans improves performance and reduces errors. Similarly, we extend this idea to the entrepreneurship domain by asking participants to mentally practise conducting their customer survey, their savings strategy or a plan to become more productive and reduce waste, to name a few examples. Participants are encouraged to implement these plans as homework and report back on their experiences, with the hope that these reflections should serve to improve the quality of imagery over time.

The final category of exercises entails adopting the perspective of others. We hypothesise that mental simulations should allow people to predict the choices and behaviour of others more precisely and react accordingly. Adopting the perspective of others is particularly important for entrepreneurs who should cater to their customers, while differentiating themselves from their competitors. Moreover, with high credit constraints, our sample of micro-entrepreneurs may be required to request funding from informal and formal networks and adapt their pitch to sell their proposal most effectively. In addition, our sample hire approximately two part or full time employees on average; managing and motivating these employees will thus be necessary. Our curriculum is designed to develop

perspective taking for this wide range of situations. For instance, over several sessions, participants visualise their product or service from the shoes of their target customer and identify how best to improve their product so to satisfy their needs[19], differentiate from their closest competitors and minimise customer dissatisfaction.

All imagery exercises were implemented following a standardised structure. The facilitator first introduces the purpose of the exercise and prepares participants with the correct posture and breathing. The facilitator subsequently follows a script to guide participants through the exercise, which lasts five to ten minutes on average. The scripts are designed to be sufficiently general in nature, such that participants could apply the imagery to their own business case. After the exercise, participants are asked to reflect on a few debriefing questions in pairs and as a group, which aim to provide feedback on their experience and motivate how they may use the technique in their business decision making. The structure was designed to teach imagery as a general skill that could be used more broadly in the domain of business decision making. Through debriefs and homework tasks, participants were repeatedly provided the opportunity to practise the three applications outlined above and reminded of the value of imagery in connection with real-life behaviours and outcomes.

2.5 Experimental design

This section provides an overview of the treatment arms, the timeline of activities, a description of the sample, take up rates and attrition.

2.5.1 Treatment arms

1967 participants were randomised into three treatment arms across two waves of implementation in a randomised control trial. To isolate the effect of imagery, we randomised 956 participants into the imagery programme and 558 participants into the placebo training. Of these, 770 and 431 participants in the imagery and

[19] In place of this particular imagery exercise, the placebo curriculum asked participants to describe their target customer in detail to a partner and reflect on their needs. Their partner then evaluated whether the product met the needs of the respective customer on a scale of 1 to 5. Upon sharing the score, the pair discussed ways in which to improve the product such that it better meets the needs of the customer.

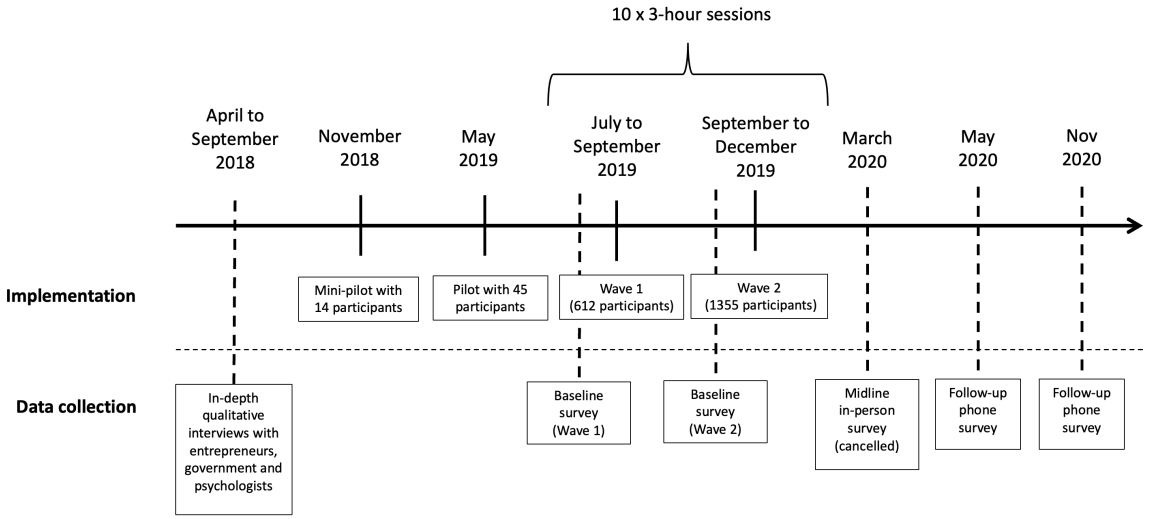
placebo training arms respectively agreed to participate in the training prior to knowing their treatment status. To test the effect of the training, we randomised a further 453 people into a no-intervention group that received nothing. We doubled the number of participants in the imagery training relative to both the placebo and no-intervention control groups in order to meet the budget reporting requirements of our government implementing partner.

2.5.2 Timeline of activities and data collection

For ease of implementation, the intervention was delivered in two waves, the first taking place from July to September 2019 and the second from September to December 2019. Figure 3.8 shows the timeline of the intervention and data collection. We conducted an extensive qualitative study and two pilots to refine the curriculum and RCT design in the year prior to implementation. Before starting the training within both waves, we first screened interested applicants to determine eligibility, randomised eligible participants into the three treatment arms and conducted a baseline survey. We conduct the first follow-up phone survey 6 to 8 months and the second follow-up phone survey 12 to 14 months after the end of the training programme.[20]

[20]Our in-person follow-up survey was scheduled to commence in early March 2020. However, the survey was cancelled shortly thereafter in response to the Covid-19 pandemic and the stringent nationwide lockdown in Colombia. We implemented two phone surveys commencing in May and November 2020 instead.

Figure 2.5: Timeline of implementation and data collection



To reduce the correlation between treatment assignment and social desirability bias in survey responses, we used surveyors from Innovations for Poverty Action (IPA) Colombia, an independent nonprofit research organisation. IPA Colombia communicated to respondents that they were conducting a survey on entrepreneurs in Bogotá in collaboration with our local government partner. All respondents were compensated COP $10,000 ($2.8) for their time, irrespective of whether they completed the survey.

2.5.3 Sample

We designed our curricula to target a low-income population in a country where there are over eight million registered victims of conflict, of whom 85% were forcibly displaced, and more than 1.2 million Venezuelan migrants. Our government partners recruited participants through a multi-channel media campaign by advertising an entrepreneurship programme promoting soft skills, including through community centres and social media platforms. Interested applicants were required to fill out a short application form online or in-person, which we

used to screen for eligibility.[21] To meet their mandate, the District Secretariat of Social Integration was required to train numerous sub-populations, all of whom can be defined as low income.[22] When we randomly select and randomise eligible applicants into our three treatment arms, we stratified by subdivision, sex, age, entrepreneurship status (existing business or idea), sales for business owners or income for people without a business.[23]

Our final sample include 1,967 would-be entrepreneurs with either an existing business (51%) or a stable business idea only (49%). Table 2.1 provides descriptive statistics for this sample. In line with the mission of our government partners to serve "vulnerable" populations, 85% reported baseline exposure to traumatic past experiences, including armed conflict and forced displacement. 58% of participants are women. The average entrepreneur is 32 years old and has completed secondary school. Participants earn an average monthly income of COP $693,077 (approximately USD $200) at baseline, which is approximately 80% of the Colombian monthly minimum wage of COP $877,802.

As presented in Table 2.1, participants appear well balanced on baseline characteristics across all treatment and control groups. This balance extends across a wide range of characteristics, including variables that were not used for stratification in the randomisation procedure.

[21] 3553 individuals applied to attend our programme, of which 2337 were deemed eligible according to our predefined criteria. To be eligible, participants needed to demonstrate entrepreneurship potential by reporting having a business or plans to launch a business in the following three months. We further defined entrepreneurship potential to encompass those who could describe their business or business idea in a few words and classify them by sector. In the second wave, we also prioritised applicants who reported taking at least three steps towards starting or growing their business in the last six months. In order to account for the need to self-fund transportation costs to the training centres, we limited eligibility to those applicants who reported either non-zero income or business sales in the past six months. Eligible applicants were also required to be literate, over the age of 18 years and to provide three points of contact.

[22] The sub-populations include: victims of conflict; Venezuelan migrants; low-income youth; LGBTQ groups; entrepreneurs with disabilities or carers thereof; the formerly homeless; and the elderly. Appendix Table 2.13 summarises the breakdown of the sample by subdivision and treatment status.

[23] Participants who reported living in the same address were assigned to the same treatment status to avoid spillovers. Moreover, formerly homeless participants affiliated with the same shelter were also assigned to the same treatment status. For our analysis, we aggregate the data to household level.

Table 2.1: Baseline Balance Table

	(1) Imagery Treatment		(2) Placebo Treatment		(3) Pure Control		T-test Difference			F-test for joint
Variable	N	Mean/SE	N	Mean/SE	N	Mean/SE	(1)-(2)	(1)-(3)	(2)-(3)	orthogonality
Female	906	0.578	531	0.580	415	0.571	-0.002	0.007	0.009	0.043
		(0.016)		(0.021)		(0.024)				
Age (18-28)	906	0.604	531	0.571	415	0.629	0.033	-0.025	-0.058*	1.704
		(0.016)		(0.022)		(0.024)				
Age (29-45)	906	0.240	531	0.250	415	0.272	-0.011	-0.033	-0.022	0.815
		(0.014)		(0.019)		(0.022)				
Age (46-59)	906	0.092	531	0.113	415	0.094	-0.021	-0.002	0.019	0.919
		(0.010)		(0.014)		(0.014)				
Years Education	720	13.132	429	13.046	327	13.366	0.086	-0.234	-0.320	0.869
		(0.125)		(0.177)		(0.173)				
Only Business Owner	906	0.242	531	0.245	415	0.236	-0.003	0.006	0.009	0.048
		(0.014)		(0.019)		(0.021)				
Only Have a Business Idea	906	0.491	531	0.488	415	0.494	0.003	-0.003	-0.006	0.018
		(0.017)		(0.022)		(0.025)				
Have Business and Business Idea	906	0.267	531	0.266	415	0.270	0.002	-0.003	-0.004	0.011
		(0.015)		(0.019)		(0.022)				
Income Strat: Sales $\leq$ 1 month salary	906	0.667	531	0.663	415	0.641	0.004	0.026	0.022	0.434
		(0.016)		(0.021)		(0.024)				
Income Strat: Sales $>$ 1 month salary	906	0.240	531	0.230	415	0.248	0.010	-0.009	-0.018	0.221
		(0.014)		(0.018)		(0.021)				
Refused to Provide Income	906	0.094	531	0.107	415	0.111	-0.014	-0.017	-0.003	0.590
		(0.010)		(0.013)		(0.015)				
Assets Owned	720	11.681	429	11.794	327	11.835	-0.114	-0.155	-0.041	0.225
		(0.144)		(0.189)		(0.209)				
Household Size	719	3.484	428	3.579	326	3.650	-0.095	-0.165	-0.070	1.303
		(0.058)		(0.077)		(0.095)				
Save Monthly Y/N	716	0.552	429	0.531	326	0.580	0.020	-0.028	-0.048	0.873
		(0.019)		(0.024)		(0.027)				
Had Access to Credit	716	0.156	428	0.143	326	0.138	0.014	0.018	0.004	0.379
		(0.014)		(0.017)		(0.019)				
No. of Traumatic Events	906	2.123	531	2.295	415	2.308	-0.172	-0.185	-0.013	1.279
		(0.076)		(0.105)		(0.125)				
Reported Experiencing Trauma (Y/N)	719	0.828	429	0.837	326	0.844	-0.009	-0.016	-0.007	0.228
		(0.014)		(0.018)		(0.020)				
Impact Event Score	710	17.945	428	17.650	324	17.785	0.295	0.159	-0.136	0.024
		(0.842)		(1.092)		(1.164)				
Kessler Score	719	13.170	429	13.232	326	13.138	-0.062	0.032	0.094	0.054
		(0.157)		(0.194)		(0.220)				
Wave 1 Participant	906	0.294	531	0.392	415	0.304	-0.098***	-0.010	0.088***	7.918***
		(0.015)		(0.021)		(0.023)				

Notes: The table presents a balance test at baseline for all three treatment arms. Columns 1, 2, and 3 show the sample size, mean, and standard errors for the imagery treatment, placebo treatment, and no-intervention groups, respectively. The following 3 columns show the differences between treatment arms and p-values for t-tests measuring differences between treatment arms. The final column presents an F-test across all three treatment arms. * p<0.10, ** p<0.05, *** p<0.01

2.5.4 Take up and attrition

We find that take up of the treatment is substantial, with about 67% of individuals attending at least one session in the imagery training across both waves, conditional on confirming attendance. By way of comparison, attendance rates in the McKenzie and Woodruff (2014) review of business training programmes in developing countries range from 39 to 92%, with a mean of 64%. Of those attending at least one session, 48% attended seven or more of the ten sessions in the im-

agery treatment and qualified for a certificate. Similarly, 63% attended at least one session in the placebo training across both waves, conditional on confirming attendance. Of those attending at least one session, 53% attended seven or more sessions in the placebo training.[24]

We surveyed 74% of all participants in our first follow-up survey and 63% of participants in our second follow-up survey. We test whether attrition is differential across all three treatment groups. We find that 64% of participants from our imagery and placebo groups responded to our endline survey, so there is no differential attrition. However, only 57% of the no-intervention group were interviewed in the second follow-up survey. The difference is statistically significant at the 10% level.[25]

Table 2.2 shows that there is little evidence to suggest that there is selective attrition on observable characteristics at endline, especially between imagery and placebo groups. Attrition tables at the stage of training acceptance and at midline are provided in the Appendix Tables 2.25 and 2.26.

[24]Take up rates are 58% and 55% in the imagery and placebo group, respectively, if we do not condition on confirmed attendance. Our sample is highly resource-constrained and mobile. In collaboration with SDIS, we developed several strategies to maximise compliance among those who were randomised to both training treatments. For instance, participants were assigned to venues that matched their location preferences to the best extent possible. Our supervisors ensured that participants were reminded on a weekly basis over WhatsApp groups, text messages and phone calls. Participants who attended seven or more sessions received a certificate signed off by SDIS and our World Bank collaborator and invited to attend a graduation ceremony in the case of the second wave.

[25]Although the differential attrition is more pronounced in the second follow-up survey, we see early evidence of selective attrition in the first follow-up survey. While we interviewed 74% and 75% of respondents from the placebo and imagery treatments respectively, only 70% of the no-intervention group responded to the first follow-up survey. A t-test reveals that the difference across the no-intervention group and imagery/placebo treatment is marginally significant with $p = 0.1$.

Table 2.2: Endline Attrition Table

	(1) Imagery Treatment		(2) Placebo Treatment		(3) Pure Control		T-test Difference			F-test for joint
Variable	N	Mean/SE	N	Mean/SE	N	Mean/SE	(1)-(2)	(1)-(3)	(2)-(3)	orthogonality
Female	579	0.573	341	0.610	251	0.594	-0.037	-0.020	0.016	0.610
		(0.021)		(0.026)		(0.031)				
Age (18-28)	579	0.604	341	0.540	251	0.610	0.065*	-0.005	-0.070*	2.212
		(0.020)		(0.027)		(0.031)				
Age (29-45)	579	0.237	341	0.258	251	0.291	-0.021	-0.054*	-0.033	1.373
		(0.018)		(0.024)		(0.029)				
Age (46-59)	579	0.100	341	0.147	251	0.092	-0.046**	0.009	0.055**	2.997*
		(0.012)		(0.019)		(0.018)				
Years Education	497	12.949	294	12.885	216	13.216	0.063	-0.267	-0.331	0.658
		(0.148)		(0.210)		(0.222)				
Only Business Owner	579	0.254	341	0.258	251	0.239	-0.004	0.015	0.019	0.150
		(0.018)		(0.024)		(0.027)				
Only Have a Business Idea	579	0.491	341	0.504	251	0.490	-0.014	0.000	0.014	0.095
		(0.021)		(0.027)		(0.032)				
Have Business and Business Idea	579	0.256	341	0.235	251	0.271	0.021	-0.015	-0.036	0.528
		(0.018)		(0.023)		(0.028)				
Income Strat: Sales $\leq$ 1 month salary	579	0.705	341	0.674	251	0.689	0.030	0.015	-0.015	0.466
		(0.019)		(0.025)		(0.029)				
Income Strat: Sales $>$ 1 month salary	579	0.211	341	0.220	251	0.207	-0.009	0.004	0.013	0.083
		(0.017)		(0.022)		(0.026)				
Refused to Provide Income	579	0.085	341	0.106	251	0.104	-0.021	-0.019	0.002	0.692
		(0.012)		(0.017)		(0.019)				
Assets Owned	497	11.468	294	11.532	216	11.634	-0.065	-0.166	-0.102	0.147
		(0.170)		(0.225)		(0.249)				
Household Size	496	3.504	293	3.635	215	3.715	-0.130	-0.211	-0.080	1.536
		(0.068)		(0.090)		(0.120)				
Save Monthly Y/N	494	0.559	294	0.527	215	0.563	0.031	-0.004	-0.036	0.454
		(0.022)		(0.029)		(0.034)				
Had Access to Credit	494	0.160	294	0.156	215	0.149	0.003	0.011	0.008	0.069
		(0.017)		(0.021)		(0.024)				
No. of Traumatic Events	579	2.389	341	2.567	251	2.597	-0.179	-0.208	-0.029	0.897
		(0.099)		(0.134)		(0.165)				
Reported Experiencing Trauma (Y/N)	496	0.837	294	0.864	215	0.842	-0.027	-0.005	0.022	0.540
		(0.017)		(0.020)		(0.025)				
Impact of Event Score	490	18.320	294	18.459	215	18.276	-0.139	0.043	0.183	0.005
		(1.023)		(1.339)		(1.442)				
Kessler Score	496	13.149	294	13.287	215	13.093	-0.138	0.056	0.194	0.162
		(0.188)		(0.231)		(0.280)				
Wave 1 Participant	579	0.306	341	0.370	251	0.319	-0.064**	-0.013	0.051	2.037
		(0.019)		(0.026)		(0.029)				

Notes: The table presents a balance test at the second follow up survey for all three treatment arms. Columns 1, 2, and 3 show the sample size, mean, and standard errors for the imagery treatment, placebo treatment, and no-intervention groups, respectively. The following 3 columns show the differences between treatment arms and p-values for t-tests measuring differences between treatment arms. The final column presents an F-test across all three treatment arms. * p<0.10, ** p<0.05, *** p<0.01

2.6 Measurement

In constructing our outcome measures, we follow a detailed pre-analysis plan, registered in the American Economic Association Registry.[26] We collect all data via phone surveys and therefore our measures are limited by this format.

[26] https://www.socialscienceregistry.org/trials/6576

2.6.1 Primary outcome measures

We are interested in three sets of primary outcomes (or "families"): imagery, psychological resilience and economic outcomes. The imagery outcomes are considered "first stage" outcomes, used to check that the curriculum combining entrepreneurship and imagery is having their intended impact. The psychological resilience and economic outcomes are considered "second stage" outcomes that we hypothesise are affected by imagery.

Mental imagery

Our intervention was designed to increase the frequency with which participants make use of imagery in their business lives and to improve the quality of the images generated, conditional on using imagery. We construct a mental imagery index that captures both dimensions through adapted versions of two widely-used imagery scales in the psychology literature. We are primarily interested in the channel through which imagery might affect business decision making, so we adapt these scales to the entrepreneurship domain.

To measure the frequency of imagery usage, we adapt the Spontaneous Use of Imagery Scale (Reisberg et al., 2003) to the entrepreneurship domain. Our scale consists of eight statements designed to assess the propensity of an individual to make use of imagery in business-related scenarios. Specifically, respondents are asked to consider business-related statements and determine the extent to which they agree with the statement using a five-point Likert scale. For example, we ask respondents to consider "When I need to go to a meeting, I picture the route in my mind before going."

To assess the quality of imagery in an entrepreneurship domain, we adapt the Prospective Imagery Task (MacLeod, 1996; Stöber, 2000). This exercise asks respondents to imagine three positive and three negative business-related and non-business-related scenarios. For example, respondents are asked to imagine a scenario in which "the Covid-19 pandemic is over, and you are struggling to make ends meet" or in which "the Covid-19 pandemic is over, and your business is doing well". Respondents are then asked to assess the vividness or level of detail

of each image, and the emotional intensity generated by the image. We measure vividness and emotional intensity to pick up on the extent to which imagery is an "emotional" and "motivational amplifier".

From this scale, we build four sub-indices: two indices for imagery vividness and two indices for the emotional intensity of the images associated with the positive and negative scenarios respectively. We construct separate sub-indices for positive and negative imagery, as we expect the intervention to have differential effects on each type, especially for participants with a past history of trauma. In particular, participants with post-traumatic stress disorder tend to lack vividness in autobiographical positive imagery and overcompensate in the emotional intensity of negative imagery (Kleim et al., 2014).

The full list of questions used in the construction of the imagery index can be found in Appendix Table 2.14.

Psychological resilience

We explore whether the imagery training has downstream effects on mental wellbeing during the pandemic by building two indices. The first index captures psychological resilience, defined as the ability to respond well in the face of adversity. The second index reflects psychological distress, as measured by the Kessler K6 non-specific distress scale. The full list of questions used in the construction of the psychological resilience indices can be found in Appendix Table 2.15.

Economic outcomes

In addition to mental wellbeing, we assess the downstream impact of imagery training on economic outcomes. Our methodology allows us to unpack whether the imagery training improves economic decision making and outcomes relative to a more traditional entrepreneurship training or no training at all. Given the onset of a Covid-19-induced national lockdown in Colombia three months after the end of our intervention, we are well placed to evaluate whether imagery improves economic resilience in response to a negative shock. We construct economic indices to capture these downstream treatment effects and distinguish between two time

periods: (1) economic activity period prior to the Covid-19 induced lockdown (i.e. before 24 March 2020) and (2) economic activity during the lockdown.

Our first primary economic outcome is an index of sales and take-home income, conditional on owning a business, or simply take-home income if no business exists at the time of the survey. We ask participants to recall both outcomes for the period prior to the national lockdown and in the week prior to the survey. Business status is our second economic outcome of interest and in particular, whether a business is open or only temporarily closed (versus permanently closed) during the lockdown. Our third economic outcome of interest is business investment undertaken both pre-Covid and during the pandemic. We construct the investment index by identifying whether respondent newly acquired or significantly improved existing business assets in a predefined list. We are furthermore interested in whether our imagery training helped entrepreneurs in developing better safety nets, proxied by savings in the pre-Covid period and their perceptions of informal support networks during the pandemic. Finally, we assess whether treated entrepreneurs adapt their business more effectively to the pandemic and the associated government restrictions. To measure this, we ask participants to respond with yes/no answers to a list of thirteen business behaviours, ranging from setting up a safe work environments to identifying alternative supply chains to diversifying their products. We combine the index of behaviours with a sub-index of reverse-coded "unsafe" working hours during the Covid-19 pandemic, defined as the proportion of total hours worked "unsafely" in which social distancing, frequent hand washing, the use of face masks or home working were not adhered to (the government Covid-19 guidelines of the day).

The full list of questions used in the construction of the economic indices for the period prior to the Covid-19 national lockdown and during the lockdown can be found in Appendix Tables 2.16 and 2.17 respectively.

Trauma

Research in clinical psychology and neuroscience suggests that trauma can increase the cost of mental imagery by inducing distressing images and alter the

quality of imagery by undermining the specificity of images. Accordingly, we expect our imagery programme to have differential treatment effects for those individuals who reported higher baseline levels of trauma.

We measure trauma at baseline by first conducting a contextually-relevant trauma history checklist that captures whether respondents experienced or witnessed a traumatic event. To proxy for the level of trauma, we use the 22-item Impact of Event Scale-Revised to assess subjective distress in the past month caused by the most traumatic event ever lived, conditional on experiencing a traumatic event. Symptoms include "trouble concentrating", "troubles removing them from memory" and "trying not to talk or think about them". We construct a dummy variable for high trauma by assigning a value of one to participants with an Impact of Event Scale-Revised score of above 33 at baseline. A score of 33 is considered the threshold above which post-traumatic stress symptoms may be considered to be a probable clinical concern (Creamer, Bell and Failla, 2003). All respondents who reported that they did not experience a traumatic event in the past were assigned a score of zero.

Variable and index construction

To construct the variables of interest, we follow the procedure outlined in our pre-analysis plan. Given the presence of outliers, we winsorise the main economic measures (e.g. revenues, income, savings) at the 99th percentile. To address the right-skewed nature of their distributions, we use the inverse hyperbolic sine transformation of the raw measures.

We use the methodology of Kling et al. (2007) to construct indices and adopt the following steps: (i) we first ensure that all variables are consistently signed (e.g. higher value associated with higher imagery ability); and (ii) then standardise each component of the index by subtracting the no-intervention group mean and dividing by its standard deviation. In the case where there are multiple sub-scales, we take two additional steps: (iii) we take the sum of the standardised components and (iv) standardise the sum again using the control group mean and standard deviation. For the psychological scales, we first sum the individual response items within a scale prior to standardising the indices.

2.7 Empirical strategy

The Covid-19 pandemic heightened the noise inherent in business revenue and household income. Given the low auto-correlation in these measures, we conducted multiple follow-up surveys to average out the noise and increase statistical power. There were two surveys in total: the first commencing in May 2020 and the second commencing in November 2020. Hence, our analysis combines these two phone surveys, with clustering at the household level and the inclusion of survey fixed effects.

For each primary outcome, we estimate the following specification at the household level:[27]

$$y_{iws} = \beta_0 + \beta_1 \cdot T_{iw} + \theta m_{iw} + \eta_w + \delta_s + \mu_{iw}, \tag{2.1}$$

where y_{iw} is the outcome of interest for household i in wave w and survey s; T_{iw} is a dummy capturing whether the household was offered a given training; m_{iw} is the vector of randomisation strata dummy variables; and η_w and δ_s are wave-subdivision pair and survey fixed effects respectively. The strata variables include the subdivision, gender, age group, entrepreneurship status (business, business idea or both) and income/sales group (low, medium or high) of the individual included in the randomisation.

We first compare the results for the imagery training relative to the placebo training to isolate the effect of teaching imagery. We then compare results for both the imagery and placebo training to the no-intervention group in order to evaluate the effect of business training more broadly. Our estimates measure the intent-to-treat impacts of the interventions relative to the defined control group. For every primary outcome, we test the null hypothesis that the treatment has no impact. For inference, we present standard errors clustered at the household level and false discovery rate q-values calculated within each family of primary outcomes using the sharpened two-stage procedure of Benjamini et al. (2006). We present results without the inclusion of baseline covariates x_{iw} as controls for now.

[27] The majority of participants reflect single-unit households. We collapse the data to the household level for those individuals who reported living in the same address as another selected participant at screening and were thus randomised into the same treatment status.

Following our pre-analysis plan, we restrict the sample for analysis to people who confirmed their participation in the imagery or placebo programmes, which is approximately 78% of the sample for these treatment groups. Participants were blind to which programme they would receive when they had to confirm their intention to participate. Moreover, staff members who called potential participants to ask for their confirmation were following guidelines and scripts that did not contain any reference to the exact treatment to which people were assigned. Nevertheless, we collected outcomes on the full sample (including people who did not confirm participation) to check for differences between groups.

We also present heterogeneous treatment effects by trauma and gender in separate specifications. In each of them, we interact the treatment indicator variable T_{iw} with an indicator H_{iw} for being a woman or having experience a high level of trauma symptoms in the past month. This means that we use low-trauma individuals and men as the base categories for each specification respectively. To conduct heterogeneity analysis, we estimate a fully interacted model as follows:

$$y_{iws} = \beta_0 + \beta_1 \cdot T_{iw} + \beta_2 \cdot H_{iw} + \beta_3 \cdot T_{iw} * H_{iw} + \theta m_{iw} + \eta_w + \delta_s + \mu_{iw}, \tag{2.2}$$

where β_3 captures the additional average treatment effect for high-trauma individuals relative to low-trauma individuals or women relative to men, depending on the specification.

2.8 Results

2.8.1 Imagery versus placebo training

We begin by evaluating whether imagery can indeed be taught and the downstream effects of teaching imagery on psychological and economic outcomes. To isolate the impact of teaching imagery, we compare the results of the imagery treatment relative to the placebo training. With the exception of the imagery exercises in the imagery treatment, the two interventions were otherwise identical in terms of content and delivery. The following results are also robust to the inclusion of participants who did not confirm participation to the training when still blind to treatment status (Appendix Tables 2.28 and 2.29).

Is imagery malleable?

We first consider whether the imagery training had its intended impact: to develop imagery as a skill in business decision making. We combine the SUIS, PIT vividness and emotional intensity measures to create an overall imagery index. We focus on business scenarios as the relevant domain in which imagery was trained and thus most likely to have an impact. Given that the business-specific imagery was only asked to individuals with a business at the time of the surveys, we restrict our sample to households that had a business at baseline (51% of the sample). We find that there is no differences in business status across treatment arms in both follow-up surveys, lending support to the assumption of monotonicity.[28]

In support of the hypothesis that imagery can be trained, we find a 0.181 standard deviation difference in our imagery index for business scenarios between the imagery and placebo groups, as reported in Table 2.3. Even though we are considering half our sample, this difference is statistically significant at the 10% level ($p = 0.058$). As shown in Table 2.3, we see no difference in our imagery index between the imagery and placebo groups when we consider non-business scenarios for this sub-sample ($p = 0.934$). There are two possible reasons underlying the stronger imagery treatment effects in business-specific scenarios relative to non-business scenarios. First, imagery is domain specific and thus can only be trained for a specific type of decision making, while recognising that business decision making encompasses a rather broad cluster of decisions. Second, the sub-sample of entrepreneurs who had a business already at baseline were better able to train in imagery, as they had a greater opportunity to practice the skill within the domain of business decision making. However, there are null treatment effects on non-business imagery for both the business sub-sample (0.007 standard deviations) and the overall sample (-0.006 standard deviations).[29]

[28] 84.5% of households in the imagery treatment that had a business at baseline also had a business in the first follow-up survey, compared to 81% in the placebo group (difference not statistically significant at $p = 0.286$). These proportions shift to 75.5% in the imagery treatment and 73.7% in the placebo training in the second follow-up survey (difference not statistically significant at $p = 0.660$). The effect of imagery training on business status is likely driven by participants who started a business during the programme.

[29] See Appendix Table 2.19 for imagery results for the overall sample.

Table 2.3: Imagery vs Placebo Training: Effect on Imagery

	Combined imagery index		Business-specific imagery sub-components				
	(1) Imagery (non-business)	(2) Imagery (business)	(3) SUIS	(4) Positive PIT Vivid	(5) Positive PIT Emotion	(6) Negative PIT Vivid	(7) Negative PIT Emotion
Imagery training	0.007	0.181*	0.086	0.056	0.042	0.174**	0.158*
	(0.080)	(0.095)	(0.091)	(0.093)	(0.088)	(0.086)	(0.085)
FDR q-values			0.522	0.607	0.607	0.186	0.186
Strata Controls	✓	✓	✓	✓	✓	✓	✓
Survey Wave FE	✓	✓	✓	✓	✓	✓	✓
Wave-Subpop FE	✓	✓	✓	✓	✓	✓	✓
Mean DV in Placebo	-0.01	-0.09	-0.04	-0.04	-0.06	-0.10	-0.02
R-squared	0.038	0.047	0.033	0.056	0.025	0.023	0.036
Total N	936	798	797	792	792	789	789
N in Placebo	341	285	285	283	283	282	281
N in Treat	595	513	512	509	509	507	508

Notes: The table presents the difference-in-means in the standardised imagery measures between the imagery and placebo arms for the subsample of participants with a business at baseline. Standard errors clustered at the household level are presented in parenthesis. Columns 1 and 2 present results for non-business and business-specific imagery respectively. Differences in observation numbers between columns 1 and 2 reflect that some participants with a business at baseline closed their business during the Covid-19 pandemic and were excluded from business-specific scenarios. Columns 3 to 7 decompose the overall business imagery index into its constituent parts and we calculate the false discovery rate q-values across these five hypotheses following the sharpened two-stage procedure of Benjamini et al. (2006). Column 3 presents the standardised SUIS index, which captures the frequency with which people use imagery. Column 4 to 7 present results on the self-reported quality of scenarios imagined using the PIT scale, with columns 4 and 5 focusing on positive future scenarios and columns 6 and 7 on negative future scenarios. We average across the two follow-up surveys and collapse the data to household level in the case where more than one household member participated in the training. The sample consists of people who confirmed participation in the training prior to knowing their treatment status. * p<0.10, ** p<0.05, *** p<0.01

Table 2.3 further decomposes the business-specific imagery index into its five constituent parts for the subset of individuals who had a business at baseline.[30] Consistent with our theory of change, we observe that imagery participants reported more frequent use of imagery using the SUIS scale (column 3) and higher quality images of future scenarios using the PIT scale (columns 4 to 7). Only the results on the quality of negative future imagery are statistically significantly different from zero. Relative to the placebo training, the imagery treatment increased the vividness of negative future business scenarios by 0.174 standard deviations ($p = 0.043$) and their emotional intensity by 0.158 standard deviation ($p = 0.063$) on average. None of these results remain statistically significant once we correct for multiple hypothesis testing across these five hypotheses, following Benjamini et al. (2006). Imagery participants in turn were more likely to report a higher subjective likelihood that the future business scenarios could occur in their lives (0.048 standard deviations), although this difference is also not statistically significant

[30] Appendix Tables 2.18 and 2.19 decompose the imagery index into its constituent parts for non-business scenarios for the sub-sample of households with a business at baseline and the entire sample respectively, as all participants were asked to imagine future non-business scenarios.

($p = 0.567$).

It is unlikely that respondents who received the imagery training will have simply learnt how to respond to our survey questions on business imagery through the training, although we cannot fully rule out this possibility. First, if respondents did simply learn that vivid and emotional imagery is deemed to be desirable, we would similarly expect a systematic bias to appear for non-business imagery, given that participants were encouraged to use imagery in their daily lives during the training. Second, the majority of the PIT questions capturing vividness and emotional intensity in imagery refer to general business scenarios or non-business scenarios that were not specifically covered in the imagery training. As examples of general business scenarios, participants were asked to imagine that their business is doing well or closes after the Covid-19 pandemic, or that they are struggling to make ends meet. The exception to this is a question on savings[31] and results are robust to excluding this question. Lastly, a learnt response to our survey questions would not explain the null treatment effects we find when comparing the imagery training to the no-intervention response, as discussed in Section 2.8.2.

The impact of imagery training on downstream outcomes

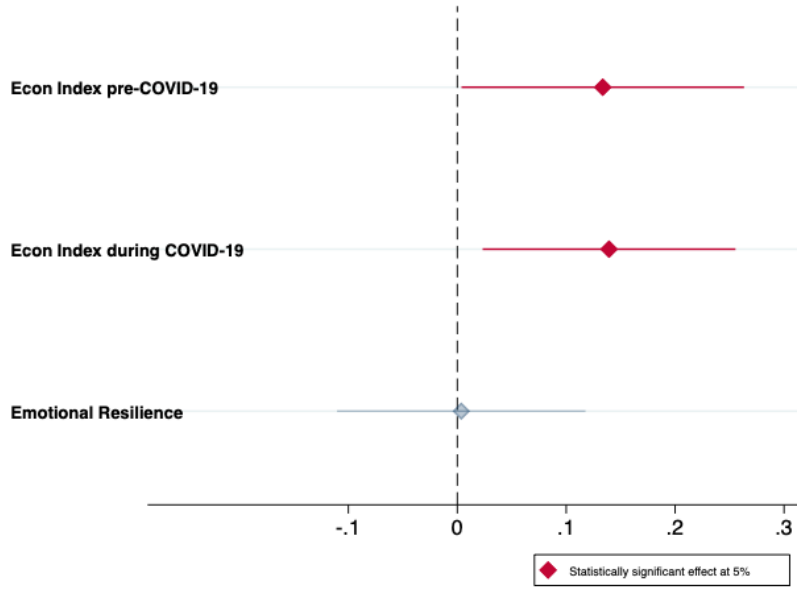
In this section, we show that imagery training improves economic outcomes, both before and after the start of the Covid-19 pandemic. We first review a set of results for condensed outcomes, where we combine our psychological outcomes into a single index measuring emotional resilience and combine our economic outcomes to create two indices: economic outcomes before the Covid-19-induced national lockdown and economic outcomes during the lockdown. Figure 2.6 shows the difference-in-means between the imagery and placebo training for the condensed outcomes with 95% confidence intervals.

Relative to the placebo group, those who received the imagery treatment score 0.13 standard deviations better on average ($p = 0.044$) on our combined pre-Covid-19 economic index. Moreover, during the Covid-19 pandemic, the imagery group

[31]The question is framed as follows: "I first want you to imagine that the Covid-19 pandemic is over, and you save enough money to buy an asset you really want."

score 0.14 standard deviations higher on average ($p = 0.019$) on our combined economic outcomes relative to the placebo group. In contrast, we observe no difference in emotional resilience measured at the time of the surveys (0.004 standard deviations; $p = 0.950$). Overall, the results are broadly consistent with our hypotheses that the use and quality of mental imagery is malleable and that imagery can improve economic outcomes.

Figure 2.6: Effects of receiving the imagery training on downstream outcomes



Notes: Markers indicate the difference-in-means on our standardised condensed outcome measures between the imagery and placebo training, with 95% confidence intervals. Red markers indicate statistical significance at the 5% level. We combine our psychological outcomes into a single index measuring emotional resilience, and combine our economic outcomes to create two indices: pre-Covid-19 economic outcomes and economic outcomes during the Covid-19 pandemic. Higher values of all variables indicates better outcomes. The sample consists of people who confirmed participation in the training. We average across our two follow-up surveys and collapse the data to the household level, the unit of randomisation. Standard errors are clustered at the household level and we control for survey wave and subdivision-wave interaction fixed effects. Monetary variables are winsorised at the 99th percentile.

Table 2.4 shows the full set of results for the economic and psychological outcomes, reporting both the p-values and sharpened q-values. Individuals receiving the imagery training report higher earnings, both before and after the start of the Covid-19 pandemic. Relative to the placebo group, imagery participants report on

average a 0.19 standard deviations increase in pre-Covid-19 earnings ($p = 0.022$) and a 0.14 standard deviations increase in earnings during the Covid-19 pandemic ($p = 0.020$). Both earnings results survive correcting for multiple hypothesis testing across our seven primary economic outcomes at the 10% level of significance ($q = 0.082$ for both results).

Imagery participants report significantly higher business survival of 0.048 standard deviations ($p = 0.058$) on average during the Covid-19 pandemic relative to the placebo group. In contrast, there is no observed difference in the average rate of business survival before the Covid-19 pandemic started ($p = 0.455$). Moreover, imagery participants report higher average levels of investment (0.066 standard deviations) and a more pro-active business behavioural response to the Covid-19 pandemic (0.084 standard deviations), although these results are not statistically significant with $p = 0.218$ and $p = 0.148$ respectively. There is no effect on the safety net index ($p = 0.669$). These broader set of results mirror the condensed results presented in Figure 2.6 above. Lastly, when we look at our pre-specified outcomes that make up the condensed emotional resilience index, we observe no difference in the Kessler Psychological Scale of Distress and general psychological resilience index between the imagery and placebo group.

Table 2.4: Imagery vs Placebo Training: Effect on Downstream Outcomes

	Pre-COVID		COVID					COVID	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Earnings	Business survival	Earnings	Business survival	Safety nets	Business behaviour	Investment	Kessler	Psych resilience
Imagery training	0.191**	0.020	0.141**	0.048*	0.024	0.084	0.066	0.017	-0.035
	(0.083)	(0.027)	(0.060)	(0.025)	(0.057)	(0.058)	(0.053)	(0.057)	(0.063)
FDR q-values	0.082	0.349	0.082	0.107	0.402	0.174	0.212	1.000	1.000
Strata Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓
Survey Wave FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Wave-Subpop FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
R-squared	0.073	0.181	0.045	0.106	0.043	0.026	0.058	0.052	0.021
Mean DV (Placebo)	-0.20	0.72	-0.14	0.65	0	-0.07	-0.05	0	-0.03
Total N	1,045	1,071	1,757	1,804	1,784	1,799	1,826	1,780	1,783
N in Placebo	380	392	642	659	650	657	667	647	649
N in Imagery	665	679	1,115	1,145	1,134	1,142	1,159	1,133	1,134

Notes: The table presents difference-in-means for downstream economic and psychological outcomes when comparing the imagery to the placebo training arm. Standard errors clustered at the household level are presented in parenthesis. False discovery rate q-values over the family of economic outcomes and the family of psychological outcomes are calculated following the sharpened two-stage procedure of Benjamini et al. (2006). We average across the two follow-up surveys and collapse the data to household level (the unit of randomisation) in the case where more than one household member participated in the training. We control for survey wave and subdivision-wave interaction fixed effects. The sample consists of people who confirmed participation in the training prior to knowing their treatment status. * p<0.10, ** p<0.05, *** p<0.01

We conduct an exploratory analysis to shed light on which behaviours might

be driving the increase in earnings. Table 2.5 reports difference-in-means results for a range of secondary outcomes when comparing the imagery to the placebo training. We first observe that imagery participants have better access to financial resources in the pre-Covid-19 period.[32] For instance, we note that, relative to the placebo group, imagery participants are 6% more likely to obtain credit ($p = 0.043$) and conditional on obtaining credit, acquire 0.13 standard deviations higher amounts ($p = 0.036$). Moreover, imagery participants reported a 4.7% higher perception of having enough savings at the start of the pandemic ($p = 0.094$), but they were not significantly more likely to save before the pandemic started ($p = 0.185$). Similarly, imagery participants appear to be more resourceful with their time. Consistent with the observed higher rates of business survival, we see that the imagery group worked 2.15 fewer hours for a wage ($p = 0.066$), although not in general. These results support the notion that the increase in earnings could have been driven by business improvements. We do not observe statistically significant differences in any other behaviour or outcomes of interest, including hiring more employees.

2.8.2 Comparison with the no-intervention group

We compared results from the imagery and placebo interventions to isolate the effect of imagery training on economic outcomes. We now explore the impact of business training more generally by comparing the no-intervention group to both the imagery and placebo interventions. Figure 2.7 illustrates the difference in means and confidence intervals for our pre-specified standardised primary outcomes between imagery training and the no-intervention group, whereas Figure 2.8 shows the difference in means between the placebo training and no-intervention group. Appendix Tables 2.20 and 2.21 tabulate these results.

The results from the imagery training relative to the no-intervention group are much smaller in size and we cannot reject that these treatment effects are statistically significantly different from zero across all outcomes of interest. Nevertheless, we observe small positive effects on business imagery and several economic

[32] We only measured access to financial resources for the period before the start of the national lockdown and not during the lockdown itself.

Table 2.5: Imagery vs Placebo Training: Effect on Secondary Outcomes

	(1) Credit > 0	(2) Credit	(3) Savings > 0	(4) High savings perception	(5) Hours worked	(6) Hours for wage	(7) Workers > 0
Imagery training	0.061**	0.131**	0.031	0.047*	-0.413	-2.150*	0.025
	(0.030)	(0.062)	(0.023)	(0.028)	(1.483)	(1.168)	(0.027)
Strata Controls	✓	✓	✓	✓	✓	✓	✓
Survey Wave FE	✓	✓	✓	✓	✓	✓	✓
Wave-Subpop FE	✓	✓	✓	✓	✓	✓	✓
Mean DV in Placebo	0.28	-0.14	0.84	0.46	31.25	14.97	0.54
R-squared	0.030	0.031	0.036	0.030	0.108	0.062	0.159
Total N	1,051	1,051	1,032	1,697	1,788	1,778	1,579
N in Placebo	382	382	384	623	653	647	562
N in Imagery	669	669	648	1074	1135	1131	1018

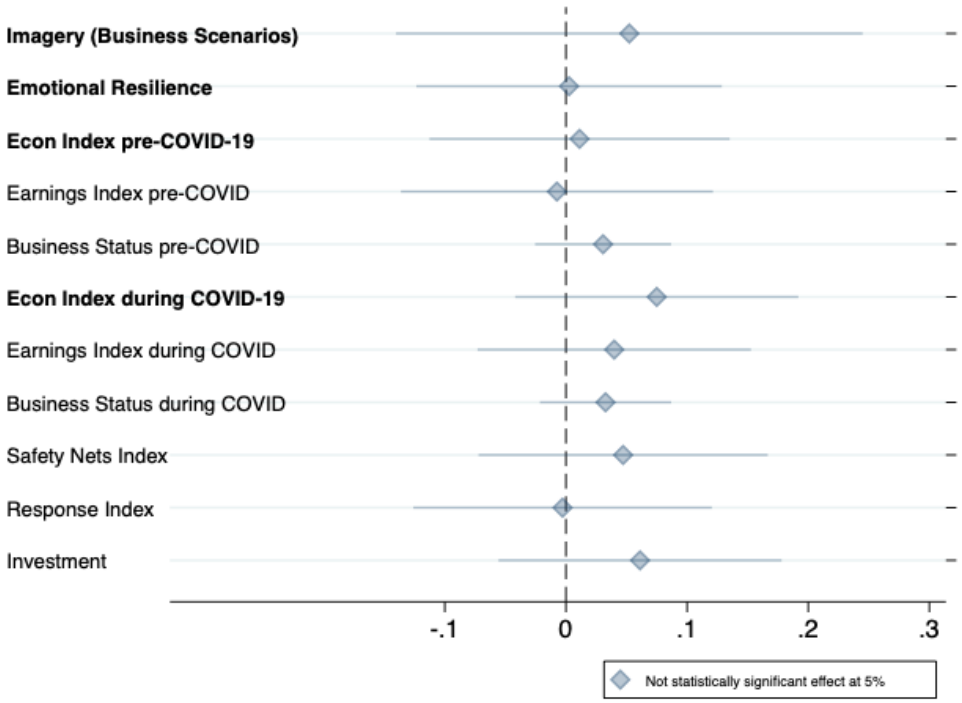
Notes: The table presents difference-in-means results on secondary outcomes when comparing the imagery training to the placebo training. Standard errors clustered at the household level are presented in parenthesis. We average across the two follow-up surveys and collapse the data to household level (the unit of randomisation) in the case where more than one household member participated in the training. We control for survey wave and subdivision-wave interaction fixed effects. The sample consists of people who confirmed participation in the training prior to knowing their treatment status. We measure the secondary outcomes as follows. Credit > 0 (column 1) is a dummy variable indicating whether the household borrowed money from any formal or informal contacts to start or grow their business in the six months prior to the Covid-19 pandemic. The credit amount obtained is standardised and reported in column 2. Savings > 0 (column 3) is a dummy variable indicating whether a household saved in the six months prior to the start of the pandemic-induced lockdown. Participants were also asked to reflect on whether they had an adequate amount of personal savings to ensure their safety for the first two months of the lockdown; high savings perception (column 4) indicates an above-median response on a 5-point Likert agreement scale. Column 5 and 6 show total hours worked and hours worked for a wage for another person, winsorised at the 99th percentile. A dummy variable indicating whether the business has hired any employee post-intervention. * p<0.10, ** p<0.05, *** p<0.01

outcomes during the Covid-19 lockdown period, including earnings, business survival, safety nets and investment. Similarly, we mostly do not observe statistically significant differences between the placebo and no-intervention group, with the exception of earnings. Relative to the no-intervention group, the placebo group report lower earnings on average, both in the six months before the start of the Covid-19 pandemic and during the pandemic itself. Specifically, the placebo group earned 0.222 standard deviations less in the pre-Covid-19 period ($p = 0.020$) and 0.144 standard deviations less during the pandemic ($p = 0.043$), relative to the no-intervention group.[33]

There are several reasons that might explain these surprising results. The first reason is the selective attrition we observe between the no-intervention group and the two treatment groups in our follow-up surveys. In contrast, there is no differential attrition between the two training arms in our follow-up surveys. Although we took great care to distinguish the surveys from the interventions, the

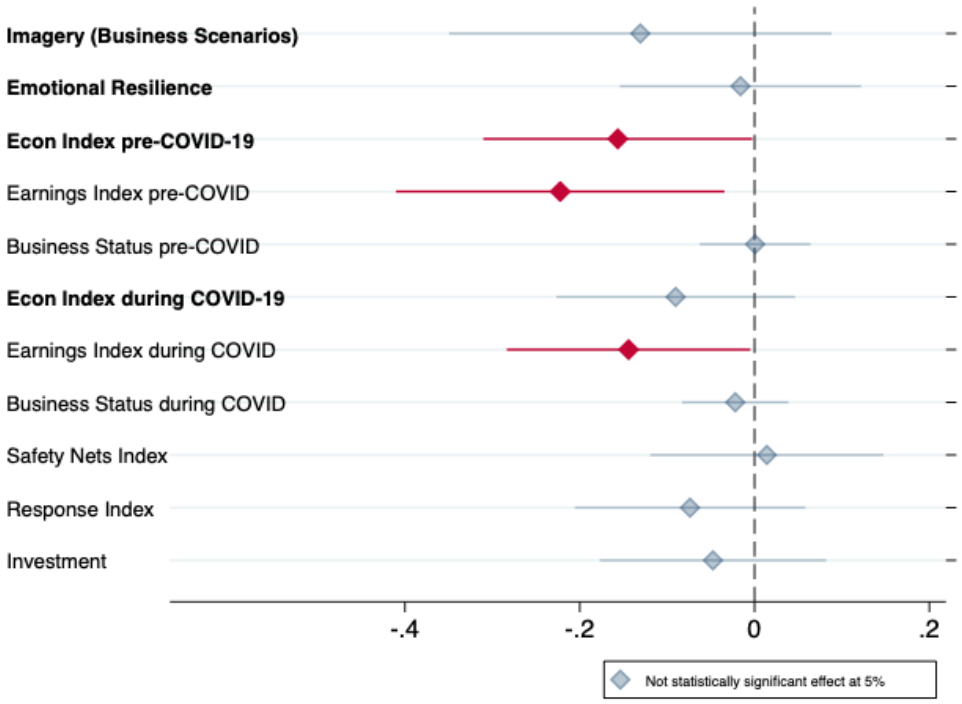
[33]The lower earnings in the placebo group withstands correcting for multiple hypothesis testing across the primary economic outcomes and winsorising at the 95th percentile.

Figure 2.7: Effect of receiving the imagery training relative to the no-intervention group



Notes: Markers indicate the standardised mean treatment effect of receiving the imagery training on pre-specified outcomes (bold) and sub-indices (normal), with 95% confidence intervals shown. Solid markers indicate significance at the 5% level. Higher values of all variables indicates better outcomes. We use the panel dataset, collapsed to the household level. Standard errors are clustered at the household level. Sample consists of people who confirmed participation in the training. Monetary variables are winsorised at the 99th percentile.

Figure 2.8: Effect of receiving the placebo training relative to the no-intervention group



Notes: Markers indicate the standardised mean treatment effect of receiving the imagery training on pre-specified outcomes (bold) and sub-indices (normal), with 95% confidence intervals shown. Cranberry markers indicate significance at the 5% level, whereas light red markers indicate significance at the 10% level. Blue markers indicate that the result is not statistically significant at conventional levels. Higher values of all variables indicates better outcomes. We use the panel dataset, collapsed to the household level. Standard errors are clustered at the household level. Sample consists of people who confirmed participation in the training. Monetary variables are winsorised at the 99th percentile.

no-intervention group were less likely to respond to our second follow-up survey in November and December 2020, relative to the imagery and placebo groups.[34] The Covid-19 pandemic significantly impacted our sample of low-income micro-entrepreneurs and their livelihoods and there is reason to believe that the no-intervention group may have been more resistant to participating in surveys in the absence of government support. However, as noted in Table 2.2 above, there is little evidence to suggest that there is selective attrition on observable characteristics at endline.

To speak to the question of selective attrition on unobservable characteristics, Table 2.6 reports Lee bounds on the main imagery and economic results (Lee, 2009). This method relies on the monotonicity assumption that the likelihood of non-response is monotonically related to receiving the treatment. In our case, the assumption implies that the treated observations who did not respond in the follow-up survey would also not have responded if they were in the no-intervention group - and vice versa. With this caveat in mind, we present lower and upper bounds for each pairwise comparison between experimental groups. The bounds are computed such that the additional percentiles of non-respondents in the relevant control group correspond to, respectively, the best and worst respective percentiles of observed outcomes in the treatment group. Table 2.6 shows that the treatment effects of the imagery training compared to the placebo arm are bounded away from zero, providing reassuring evidence that there is a positive impact of our imagery training compared to a traditional business training. Columns (4) and (6) of Table 2.6 tell a different story regarding the comparison between each training arm and the no-intervention group. For both sets of comparisons, the Lee bounds encompass zero for the great majority of outcomes. This implies that the negative results on earnings when comparing the placebo training with the no-intervention group is not robust to making weak assumptions on the non-respondent outcomes. These findings suggest that there could be differential attrition based on unobservables among the respondents in the no-intervention group.

[34] Whereas 64% of participants from our imagery and placebo groups responded to our surveys, only 57% of the no-intervention group were interviewed in the second follow-up survey. The difference is statistically significant at the 10% level.

Table 2.6: Lee bounds to correct for differential non-response

	Imagery-Placebo		**Imagery-Control**		**Placebo-Control**	
	Standardised Mean Treatment Eff. [N]	Diff. Attrition Δ n-r = 0.8% [N]	Standardised Mean Treatment Eff. [N]	Diff. Attrition Δ n-r = 4.5% [N]	Standardised Mean Treatment Eff. [N]	Diff. Attrition Δ n-r = 5.4% [N]
Non-Business Imagery	0.000	-0.012/0.036	-0.001	-0.083/0.130**	-0.046	-0.131*/0.124*
	[1985]	[1977/1973]	[1805]	[1751/1743]	[1280]	[1243/1232]
Business Imagery	0.181**	0.181**/0.207***	0.057	0.009/0.178**	-0.140	-0.216**/0.045
	[1403]	[1403/1398]	[1290]	[1265/1252]	[893]	[870/863]
Kessler	0.004	0.004/0.031	-0.011	-0.046/0.094	-0.006	-0.069/0.106*
	[1967]	[1967/1956]	[1791]	[1760/1741]	[1266]	[1235/1231]
Psych. Resilience	-0.019	-0.019/0.028	-0.036	-0.102/0.110*	-0.046	-0.102/0.096
	[1970]	[1970/1957]	[1792]	[1742/1730]	[1268]	[1240/1226]
Earnings pre-COVID	0.208**	0.198**/0.208**	-0.010	-0.060/0.126**	-0.218**	-0.286***/0.125**
	[1157]	[1152/1157]	[1057]	[1028/1031]	[746]	[724/718]
Business pre-COVID	0.011	0.011/0.011	0.018	0.018/0.018	0.003	0.003/0.003
	[1187]	[1187/1187]	[1084]	[1084/1084]	[769]	[769/769]
Earnings during COVID	0.164***	0.158***/0.164***	0.036	0.002/0.138***	-0.141**	-0.185***/0.090
	[1942]	[1933/1942]	[1765]	[1716/1731]	[1255]	[1223/1211]
Business during COVID	0.033	0.033/0.033	0.016	0.016/0.016	-0.018	-0.018/-0.018
	[1993]	[1993/1993]	[1812]	[1812/1812]	[1285]	[1285/1285]
Safety Index	0.048	0.030/0.073	0.063	-0.003/0.169***	0.013	-0.056/0.120*
	[1971]	[1958/1963]	[1793]	[1737/1739]	[1270]	[1237/1232]
Business Response	0.082	0.037/0.121**	-0.016	-0.177***/0.134**	-0.081	-0.223***/0.094
	[1988]	[1978/1978]	[1808]	[1753/1749]	[1282]	[1248/1236]
Investment	0.036	0.022/0.036	0.030	-0.068/0.030	-0.037	-0.125**/-0.037
	[2020]	[2014/2020]	[1840]	[1793/1840]	[1308]	[1287/1308]

Notes: Lee (2009) bounds (lower/upper) for estimated treatment effects are shown for each pairwise comparisons across treatment arms. The number of observations included when calculating high and low bounds are shown in square brackets. * p<0.10, ** p<0.05, *** p<0.01

The second reason for the negative results relates to demand effects in the survey measures, which may be causing the no-intervention group to diverge from the two training arms. Biased reporting may arise for two reasons. Firstly, the no-intervention group were not trained and obtained the knowledge that would enable them to respond accurately to our survey measures, such as reporting business sales. In contrast, record keeping and other business behaviours were taught and made salient to the imagery and placebo participants. Secondly, and more probable, the no-intervention group may have over-reported their success in order to attract greater financial support from the government.[35] However, the results in the comparison between the no-intervention group and the two training arms appear to move systematically across all outcomes of interest and it seems unlikely that there would be biased reporting in all survey measures.

A third reason may be that the placebo training actually had a negative impact on entrepreneurs (as seen in Oosterbeek et al. (2010) and Von Graevenitz et al. (2010)), but the imagery training mitigated the negative effects. Previous studies have discussed potential reasons explaining why traditional business training have had modestly positive results on economic outcomes at best, or at worst, null or even negative impacts (see, for instance, McKenzie et al. (2021)). The increase in earnings arising from the imagery training relative to the placebo arm almost exactly mirrors the decrease in earnings arising from the placebo training when compared to the no-intervention group.[36] When compared to the no-intervention group, the imagery training has a null treatment effect on pre-Covid earnings and we cannot reject that the 0.04 standard deviation increase in earnings during Covid-19 is statistically different from zero.

How could teaching imagery correct for such drawbacks in standard business training? On the one hand, imagery may stimulate participants' engagement with

[35] Alternatively, the placebo group may have been incentivised to under-report their successes if they felt frustrated with the programme.

[36] When we compare the imagery training to the placebo training, we observe a 0.19 standard deviation increase in earnings in the six-month period before the Covid-19 pandemic and 0.14 standard deviations increase in earnings during the pandemic-induced lockdown. As shown in Table 2.21, when compared to the no-intervention group, the placebo group report a 0.22 standard deviations decrease in pre-Covid earnings and a 0.14 standard deviations decrease in earnings during the pandemic. We observe similar patterns for business survival, business behavioural response and investment during the Covid-19 pandemic.

and attention to the content taught, leading to more effective learning. According to this view, imagery simply enhances the effectiveness of standard courses. On the other hand, imagery can be seen to appeal to completely different brain areas compared to standard teaching modes based on cognitive activation. The use of "as-if" experiences that activate all perceptions has an emotional effect on participants which is absent in structured programs based on planning or logical thinking. This is important given the high levels of trauma in our sample. Telling a depressed person logical reasons for why they need to do something will likely still make them feel depressed, unless there is also a way to act on their emotions through images. According to this latter view, imagery could unlock behavioral change in a completely different way compared to standard programs. For instance, imagery enables users to develop concrete plans and imagine the consequences of present behaviour, thereby ensuring that these plans are realistic and desirable. This may have been a critical component missing in the placebo training, leading to entrepreneurs set unrealistic goals and taking ill-informed risky decisions.

We currently do not have enough data to pin down the exact reason for the negative outcomes observed in the placebo group compared to the no-intervention control. However, we believe that differential attrition or a true negative impact of traditional training are the most likely explanations. We hope to shed more light on the differences across our three experimental groups through additional outcomes collected during a booster session to be completed in September 2021. Nevertheless, even in the worst case scenarios, our conclusion that imagery matters for economic outcomes still holds.

2.8.3 Heterogeneous treatment effects by trauma and gender

We hypothesise that imagery training should benefit people who have high costs to undertaking mental imagery or produce murky images when imagery is used. In this section, we provide evidence suggesting that the impact of the imagery treatment is heterogeneous across different groups, with strong positive treatment

effects for those with high baseline trauma and women.[37]

Heterogeneity by trauma

Distressing past experiences can increase the cost of adopting mental imagery by inducing more vivid negative imagery and undermining the quality of imagery when it is used, especially of positive future imagery. We designed the curriculum with the intention of boosting positive future imagery among populations with high levels of past trauma. Thus, we assess whether the effect of the imagery treatment is different for individuals with high levels of trauma relative to the rest of the sample. We define participants with high levels of trauma as those having an Impact of Event Scale-Revised (IES-R) score above 33, the threshold above which symptoms become clinically concerning. Just under a quarter of the sample can be identified as showcasing high levels of trauma using this definition. Previous studies leveraging IES-R scores as a measure of trauma symptoms demonstrate this to be a high level.[38]

We observe that the imagery treatment boosts the quality of positive future imagery and mediates negative future imagery for high-trauma individuals. Table 2.7 reports the heterogeneous effects of the imagery treatment on the imagery index and its sub-components by trauma. We also report the perceived likelihood of positive and negative imagined future scenarios occurring in participants' lives. We first note that although the overall imagery indices are larger for high-trauma individuals for both business and non-business imagery, the interaction terms between imagery and trauma are not statistically different from zero. Nevertheless, when we analyse the imagery sub-components, we observe strong evidence in support

[37] We pre-specified heterogeneity by trauma and defined high trauma as individuals with above 33 on the IES-R scale in our pre-analysis plan. We did not pre-specify gender; however, as we explain below, gender has a close relationship with trauma.

[38] Relative to the general population, our sample's experience of trauma is high. For example, Hunt and Evans (2004) randomly sample 414 individuals (298 happen to report experience of a traumatic event) and find a mean IES-R score of 22.3. In the context of studies which deliberately target those with known trauma experience, our sample mean IES-R remains moderately high. Ohtani et al. (2004) survey victims, of the 1995 Sarin gas attack in Tokyo subway, five years after the attack finding a mean IES-R score of 16.4.[39]. Meanwhile, for a sample exclusively characterised by motor vehicle accident experience, Beck et al. (2008) report sample mean IES-R scores of 34.98 - above the threshold for common PTSD diagnosis.[40]

of the notion that our imagery training addresses a deficit in positive future imagery. First, we note that, relative to low-trauma individuals in the placebo group, individuals with high trauma have significantly more vivid (0.48 SD; $p < 0.01$) and emotionally intense negative future imagery (0.48 SD; $p < 0.01$) and a very high perceived likelihood that these negative future scenarios could occur in their lives (0.63 SD; $p < 0.01$). High-trauma individuals also have less vivid positive imagery (-0.33 SD; $p = 0.135$) and they are less likely to use imagery at all (-0.065 SD, $p = 0.713$), although these effects are not statistically different from zero.

Second, high trauma participants who receive the imagery treatment report significantly more vivid positive imagery (a 0.56 standard deviations difference; $p = 0.032$) relative to low-trauma imagery participants and more likely to believe that these scenarios could occur in their lives (0.35 SD; $p = 0.112$). With low statistical power, these results do not survive correction for multiple hypothesis testing. There are no observed differences in the emotional intensity of positive future scenarios imagined or in the frequency of imagery used. In contrast, relative to the placebo group, low-trauma imagery participants do not show improvements in the quality of positive imagery, but they do show enhanced vividness in imagining negative future scenarios of 0.27 standard deviations ($p < 0.01$). The imagery treatment appears to mediate the vividness and emotional intensity of negative future imagery for high-trauma individuals, but these results are noisier and we cannot reject the hypothesis that the imagery training reduces the quality of negative imagery, nor that the difference between high and low trauma imagery participants is different from zero. However, we do observe that high-trauma individuals who receive the imagery treatment are 0.496 standard deviations less likely to perceive the negative future scenarios probable in their lives relative to low-trauma imagery participants ($p = 0.015$). These results suggest that the imagery training could rebalance the disproportionate weight that traumatised populations place on negative future imagery and boost their deficit in positive future imagery.

Not only do high-trauma individuals report higher quality positive imagery, but they also appear to have better economic outcomes in the pre-Covid period. Figure 2.9 plots the linear prediction of the effect of the imagery treatment relative to the placebo training on the condensed economic indices before and after the

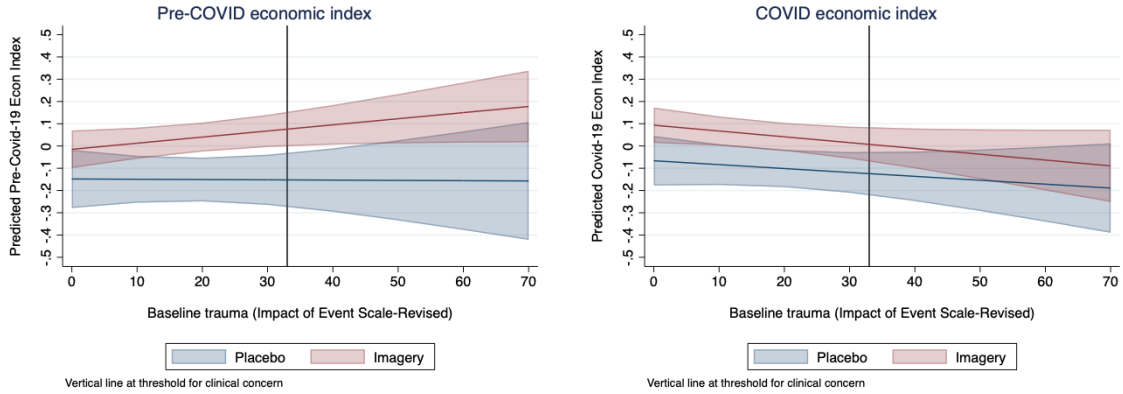
Table 2.7: Treatment effects on imagery, by trauma (business scenarios)

	Combined imagery index		Business-specific imagery sub-components					Perceived likelihood	
	(1) Imagery (non-business)	(2) Imagery (business)	(3) SUIS	(4) Positive PIT Vivid	(5) Positive PIT Emotion	(6) Negative PIT Vivid	(7) Negative PIT Emotion	(8) Positive Likelihood	(9) Negative Likelihood
Imagery training	-0.024	0.201*	0.142	-0.023	0.030	0.271***	0.150	-0.032	0.106
	(0.092)	(0.113)	(0.110)	(0.111)	(0.105)	(0.099)	(0.105)	(0.095)	(0.102)
High trauma	0.013	0.208	-0.065	-0.327	0.033	0.482***	0.476***	-0.146	0.628***
	(0.158)	(0.181)	(0.178)	(0.218)	(0.175)	(0.166)	(0.147)	(0.161)	(0.149)
High trauma*Imagery	0.209	0.080	0.124	0.556**	0.000	-0.211	-0.212	0.345	-0.496**
	(0.209)	(0.236)	(0.226)	(0.259)	(0.221)	(0.215)	(0.184)	(0.216)	(0.202)
Imag*Trauma=0 (p-value)	0.318	0.736	0.582	0.032	0.999	0.327	0.249	0.112	0.015
Imag*Trauma=0 (q-value)	0.504	0.962	0.962	0.152	0.962	0.504	0.504	0.353	0.152
Imag*Trauma+Imag=0 (p-value)	0.330	0.178	0.179	0.022	0.876	0.757	0.686	0.112	0.028
Imag*Trauma+Imag=0 (q-value)	0.474	0.334	0.334	0.145	0.638	0.609	0.609	0.334	0.145
Strata Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓
Survey Wave FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Wave-Subpop FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Mean DV	-0.03	-0.14	-0.05	0.02	-0.06	-0.20	-0.10	0.02	-0.10
R-squared	0.046	0.057	0.047	0.068	0.027	0.046	0.059	0.054	0.059
Total N	820	697	696	691	691	688	688	690	689
N in Placebo	341	285	285	283	283	282	281	283	282
N in Treat	595	513	512	509	509	507	508	508	508

Notes: The table presents heterogeneous treatment effects for the standardised imagery measures by trauma, when comparing the imagery and placebo arms for the sub-sample of participants with a business at baseline. Standard errors clustered at the household level are presented in parenthesis. We define high trauma as participants with an Impact of Event Scale-Revised score of above 33, the threshold above which symptoms become clinically concerning. Columns 1 and 2 present results for non-business and business-specific imagery respectively. Columns 3 to 7 decompose the overall business imagery index into its constituent parts and we calculate the false discovery rate q-values across these five hypotheses following the sharpened two-stage procedure of Benjamini et al. (2006). We also report the perceived likelihood of positive and negative imagined future scenarios occurring in participants' lives in columns 8 and 9 respectively. We average across the two follow-up surveys and collapse the data to household level in the case where more than one household member participated in the training. The sample consists of people who confirmed participation in the training prior to knowing their treatment status. * $p<0.10$, ** $p<0.05$, *** $p<0.01$

Covid-19-induced lockdown, by baseline level of trauma. We observe that high-trauma individuals have on average 0.30 standard deviations higher economic outcomes in the imagery treatment, relative to the placebo training in the pre-Covid period. However, we cannot reject the hypothesis that there is no difference between the imagery and placebo training, except for individuals with moderate levels of trauma symptoms. During the Covid-19 pandemic, we cannot reject the hypothesis that the imagery treatment had a differential impact on the condensed economic index compared to the placebo training.

Figure 2.9: Linear prediction of imagery training on economic outcomes, by baseline trauma



Notes: This figure plots the linear prediction of the effect of the imagery treatment on economic outcomes, by baseline trauma. The figure on the left plots the predicted condensed pre-Covid economic index for the imagery treatment (in pink) and placebo group (in blue) by baseline trauma, measured using the Impact of Event Scale-Revised score. The figure on the right plots the predicted condensed economic index during the Covid-19 pandemic for both treatment arms. The vertical line at the score of 33 indicates the level above which trauma symptoms become a clinical concern.

When we look more closely at the economic outcomes in Appendix Table 2.24, we observe a 0.36 standard deviation improvement in pre-Covid earnings ($p = 0.080$) and 0.096 standard deviations higher rate of business survival during Covid-19 ($p = 0.090$) for high-trauma individuals in the imagery group relative to the placebo group. The results for pre-Covid business survival and earnings during Covid are qualitatively consistent in their direction, pointing towards a stronger effect for those participants with higher baseline trauma, but these effects are noisier and imprecisely estimated. Moreover, these effects for high-trauma individuals are larger than for low-trauma individuals within the imagery treatment, but we cannot reject the hypothesis that there is no difference across the two

sub-groups. The results for all other outcomes for high-trauma individuals in the imagery treatment are imprecisely estimated and statistically insignificant.

Heterogeneity by gender

We also present heterogeneous treatment effects of the imagery training for women relative to men in Table 2.8. In Columns 1 and 2, we show that the impact of the imagery training on the imagery index for non-business scenarios is larger for women than for men, although we cannot reject the hypothesis that men and women have the same size treatment effect. Nevertheless, we observe a 0.3 standard deviation improvement in the business imagery index for women in the imagery treatment relative to the placebo arm ($p = 0.020$). Looking more closely at the subcomponents of the imagery index, we observe that women in the imagery treatment are significantly more likely to use imagery in their decision making (0.35 standard deviations; $p = 0.070$) compared to men. The measures proxying for imagery quality also move in the direction that suggests that women benefit more from imagery training relative to men, although these differences are imprecisely estimated and not statistically different from zero. With low statistical power, none of these results survive correcting for multiple hypothesis testing.

We then move to measures of economic performance and find that the imagery treatment has stronger effects on downstream outcomes for women compared to men, as shown in Table 2.9. On average, women in the placebo group have worse outcomes than men in the placebo group on most measures. In the pre-Covid-19 period (columns 1 and 2), we observe that women who receive the imagery treatment have 0.454 standard deviations higher earnings on average than male imagery participants ($p = 0.000$). In contrast, men who received the imagery treatment see roughly the same earnings as those in the placebo group. We observe a similar result during the Covid-19 pandemic: women earn 0.337 standard deviations more than men who receive the imagery treatment ($p < 0.01$). Relative to men, women in the imagery treatment also experience better safety nets (0.31 SD; $p < 0.01$).[41] With the exception for the results on business survival and

[41] Relative to women in the placebo group, women in the imagery treatment also report higher business survival rates during Covid-19 (0.06 SD; $p = 0.080$).

Table 2.8: Treatment effects on imagery, by gender (business scenarios)

	Combined imagery index		Business-specific imagery sub-components					Perceived likelihood	
	(1) Imagery (non-business)	(2) Imagery (business)	(3) SUIS	(4) Positive PIT Vivid	(5) Positive PIT Emotion	(6) Negative PIT Vivid	(7) Negative PIT Emotion	(8) Positive Likelihood	(9) Negative Likelihood
Imagery training	-0.072	0.017	-0.116	0.009	-0.119	0.147	0.105	-0.066	0.025
	(0.134)	(0.148)	(0.127)	(0.137)	(0.133)	(0.131)	(0.136)	(0.118)	(0.127)
Female	0.101	-0.048	-0.212	-0.058	-0.016	-0.011	0.138	-0.202	0.062
	(0.135)	(0.157)	(0.140)	(0.156)	(0.146)	(0.140)	(0.148)	(0.126)	(0.136)
Female*Imagery	0.134	0.282	0.349*	0.080	0.277	0.046	0.090	0.152	0.002
	(0.168)	(0.191)	(0.179)	(0.190)	(0.175)	(0.175)	(0.176)	(0.164)	(0.169)
Imag*Female=0 (p-value)	0.424	0.141	0.052	0.674	0.113	0.790	0.608	0.354	0.990
Imag*Female=0 (q-value)	1.000	0.732	0.732	1.000	0.732	1.000	1.000	1.000	1.000
Imag*Female+Imag=0 (p-value)	0.536	0.015	0.066	0.488	0.170	0.091	0.074	0.442	0.809
Imag*Female+Imag=0 (q-value)	0.442	0.162	0.222	0.442	0.257	0.222	0.222	0.442	0.562
Strata Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓
Survey Wave FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Wave-Subpop FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Mean DV (Placebo+Female)	-0.07	-0.07	0.09	-0.03	-0.05	-0.10	-0.10	0.08	-0.06
R-squared	0.039	0.050	0.039	0.056	0.028	0.023	0.037	0.047	0.035
Total N	936	798	797	792	792	789	789	791	790
N in Placebo	341	285	285	283	283	282	281	283	282
N in Imagery	595	513	512	509	509	507	508	508	508

Notes: The table presents heterogeneous treatment effects for the standardised imagery measures by gender, when comparing the imagery and placebo arms for the sub-sample of participants with a business at baseline. Standard errors clustered at the household level are presented in parenthesis. Columns 1 and 2 present results for non-business and business-specific imagery respectively. Columns 3 to 7 decompose the overall business imagery index into its constituent parts and we calculate the false discovery rate q-values across these five hypotheses following the sharpened two-stage procedure of Benjamini et al. (2006). We also report the perceived likelihood of positive and negative imagined future scenarios occurring in participants' lives in columns 8 and 9 respectively. We average across the two follow-up surveys and collapse the data to household level in the case where more than one household member participated in the training. The sample consists of people who confirmed participation in the training prior to knowing their treatment status. * p<0.10, ** p<0.05, *** p<0.01

safety nets, all results survive correcting for multiple hypothesis testing. Moreover, women benefit significantly more from the imagery training in boosting their psychological resilience relative to men, with a 0.315 standard deviation difference ($p < 0.019$).[42] It should be noted that this outcome is noisy, and we cannot reject the hypothesis that women do not see an increase in psychological resilience, with the difference across men and women created in part by a fall in psychological resilience for men.

Table 2.9: Effect on downstream outcomes, by gender

	Pre-COVID		COVID					COVID	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Earnings	Business survival	Earnings	Business survival	Safety nets	Business behaviour	Investment	Kessler	Psych resilience
Imagery training	-0.076	-0.023	-0.058	0.031	-0.158*	0.086	0.051	-0.033	-0.220**
	(0.109)	(0.039)	(0.083)	(0.038)	(0.085)	(0.103)	(0.093)	(0.084)	(0.095)
Female	-0.481***	-0.113***	-0.402***	-0.055	-0.401***	0.073	-0.231**	-0.384***	-0.331***
	(0.139)	(0.041)	(0.098)	(0.040)	(0.091)	(0.095)	(0.090)	(0.088)	(0.097)
Female*Imagery	0.454***	0.073	0.337***	0.029	0.310***	-0.004	0.026	0.086	0.315**
	(0.162)	(0.053)	(0.118)	(0.050)	(0.114)	(0.123)	(0.113)	(0.114)	(0.126)
Imag*Female=0 (p-value)	0.005	0.167	0.005	0.567	0.007	0.971	0.819	0.451	0.012
Imag*Female=0 (q-value)	0.021	0.200	0.021	0.480	0.021	0.760	0.693	0.431	0.021
Imag*Female+Imag=0 (p-value)	0.001	0.164	0.001	0.076	0.045	0.224	0.232	0.497	0.257
Imag*Female+Imag=0 (q-value)	0.007	0.208	0.007	0.155	0.118	0.208	0.208	0.382	0.208
Strata Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓
Survey Wave FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Wave-Subpop FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Mean DV (No-intervention)	0.08	0.78	0.10	0.67	0.24	-0.13	0.09	0.22	0.18
R-squared	0.081	0.182	0.050	0.106	0.048	0.026	0.058	0.052	0.025
Total N	1,045	1,071	1,757	1,804	1,784	1,799	1,826	1,780	1,783
N in Placebo	380	392	642	659	650	657	667	647	649
N in Imagery	665	679	1,115	1,145	1,134	1,142	1,159	1,133	1,134

Notes: The table presents standardised mean treatment effects of receiving the imagery training relative to the no-intervention group. Standard errors clustered at the household level are presented in parenthesis. False discovery rate q-values over the family of economic outcomes and the family of psychological outcomes are calculated following the sharpened two-stage procedure of Benjamini et al. (2006). We pool the two follow-up surveys and collapse the data to household level in the case where more than one household member participated in the training. The sample consists of people who confirmed participation in the training prior to knowing their treatment status. * p<0.10, ** p<0.05, *** p<0.01

The subsample analysis indicates that the effects of the imagery training concentrate on women. The divergence in results may be partially explained by the difference in key baseline characteristics for the sample randomised to imagery or placebo training, as shown in Table 2.10. At baseline, women reported higher

[42]The improvements in psychological resilience is partly driven by an increase in generalised self-efficacy: a difference of 0.26 standard deviations ($p = 0.036$) relative to men.

levels of trauma symptoms, with an average Impact of Event Scale score of 20.1 compared to 14.8 for men, although women did not report experiencing a greater number of traumatic events. 27.5% of women were identified to have high trauma with an IES-R Scale score above 33 - symptomatic of PTSD - compared to only 19% of men, and they reported a significantly higher score on the Kessler Psychological Distress Scale. This observation is consistent with the literature that women in general report experiencing higher levels of trauma (Holbrook et al., 2002). Moreover, women in our sample have less experience in business at baseline: women were less likely to own a business and took 0.3 fewer steps towards starting a business than men. Additionally, women tend to earn less income, work fewer hours, save less, and have lower access to credit than men. We hypothesise that limited past experience in business may in turn produce less specific or more murky images, as fewer details can be drawn from memory. Accordingly, it is reasonable to expect that women could respond positively to the imagery treatment, as they produce more business-specific images for their memory bank.

Nevertheless, it is also important to note that there is differential attrition by gender, which could bias the results. Approximately 63.3% of women completed the last follow-up survey, compared to only 57.9% for men (difference of 5.4%; $p = 0.015$). There is no differential attrition for imagery[43], but women were significantly more likely than men to complete the last follow-up survey if assigned to the traditional training.[44]

2.9 Robustness checks for attrition

To address the high rates of attrition across the sample, we show that our results using only a stratum with lower rates of attrition are robust. When we randomised eligible applicants into the three arms, we stratified by sub-population as listed in Appendix Table 2.13, among other strata[45]. This was required to meet the mandate of our implementing partner to train a wide range of low-income populations.

[43] Approximately 60% and 63% of men and women respectively who were randomized to the imagery training completed the last follow up surveys (difference of 2.7%; $p = 0.395$).

[44] 66% of women completed the last follow up survey, compared to only 58% of men (difference of 7.2%; $p = 0.082$).

[45] We also stratified by sex, age, entrepreneurship status, and income/sales.

Table 2.10: Baseline characteristics by gender

Variable	N	(1) Female Mean/SE	N	(2) Male Mean/SE	T-test Difference (1)-(2)
Age 18 to 28	832	0.552 (0.017)	605	0.646 (0.019)	-0.095***
Age 29 to 45	832	0.263 (0.015)	605	0.217 (0.017)	0.047**
Age 46 to 59	832	0.118 (0.011)	605	0.074 (0.011)	0.043***
Age older than 59	832	0.067 (0.009)	605	0.061 (0.010)	0.006
Household size	664	3.601 (0.061)	483	3.409 (0.072)	0.192**
Total years of education	664	13.156 (0.140)	485	13.023 (0.148)	0.133
Total assets count	664	11.757 (0.142)	485	11.676 (0.189)	0.081
Income/sales $\leq$ 1 month salary	832	0.688 (0.016)	605	0.635 (0.020)	0.053**
Income/sales $>$ 1 month salaryP	832	0.215 (0.014)	605	0.264 (0.018)	-0.049**
Business at baseline	831	0.491 (0.017)	605	0.536 (0.020)	-0.045*
Business idea at baseline	832	0.508 (0.017)	605	0.464 (0.020)	0.044
Steps taken for business idea	624	2.108 (0.053)	457	2.415 (0.072)	-0.307***
Total hours worked	664	116.611 (4.273)	485	143.440 (5.197)	-26.829***
Savings (yes/no)	661	0.520 (0.019)	484	0.576 (0.022)	-0.056*
Savings amount	655	1.01e+05 (10094.760)	483	1.44e+05 (13600.819)	-4.23e+04**
Credit (yes/no)	660	0.132 (0.013)	484	0.178 (0.017)	-0.046**
At least one trauma event	664	0.828 (0.015)	484	0.835 (0.017)	-0.006
Number of trauma events	832	2.209 (0.083)	605	2.156 (0.092)	0.053
Impact of Event Scale-Revised	659	20.072 (0.940)	479	14.754 (0.896)	5.317***
IES Score above 33	659	0.275 (0.017)	479	0.190 (0.018)	0.085***
Kessler Psychological Distress Scale	664	13.558 (0.172)	484	12.692 (0.167)	0.866***
Life satisfaction	663	6.181 (0.077)	484	6.157 (0.091)	0.024
Imagery Index from Baseline Plymouth	664	0.086 (0.036)	484	0.023 (0.043)	0.064
Completed baseline	832	0.798 (0.014)	605	0.800 (0.016)	-0.002

Notes: The table presents baseline characteristics by gender for the sample of participants randomized to the imagery or placebo training. Columns 1 and 2 show the sample size, mean, and standard errors for men and women respectively. The following column show p-values for t-tests measuring differences between men and women.

Of these sub-populations, the "territorial" group, largely comprised of victims of conflict and displaced Venezuelan migrants above the age of 28 years, is the second largest in size (N=501) after the "youth" group (N=1121). It also has the second lowest attrition rates after the "disability" group, which is much smaller in size with only 117 individuals in total. 65% of the "territorial" group completed the last follow-up survey, compared to only 61% in the entire sample. This group reflects our intended sample of interest when we first designed the study, given the high rates of trauma prevalent in this population.

We find that the imagery results are robust in magnitude to restricting the sample to the "territorial" substratum, although all results lose statistical significance with the small sample size. Table 2.11 presents the business and non-business imagery, and decomposes the business-specific imagery index into its five constituent parts for the subset of individuals who had a business at baseline. The effects on the overall business imagery index, the frequency with which imagery is reportedly used (SUIS), the vividness and emotional intensity of positive imagery are substantially larger in magnitude for the "territorial" subgroup relative to the entire sample, although we cannot reject that these effects are statistically significantly different from zero.

Table 2.11: Imagery vs Placebo: Effect on Imagery for "Territorial" Sub-Stratum

	Combined imagery index		Business-specific imagery sub-components				
	(1) Imagery (non-business)	(2) Imagery (business)	(3) SUIS	(4) Positive PIT Vivid	(5) Positive PIT Emotion	(6) Negative PIT Vivid	(7) Negative PIT Emotion
Imagery training	0.087	0.317	0.310	0.271	0.096	0.048	0.155
	(0.181)	(0.222)	(0.218)	(0.257)	(0.170)	(0.166)	(0.144)
Strata Controls	✓	✓	✓	✓	✓	✓	✓
Survey Wave FE	✓	✓	✓	✓	✓	✓	✓
Wave-Subpop FE	✓	✓	✓	✓	✓	✓	✓
Mean DV in Placebo	-0.29	-0.27	-0.23	-0.59	0.02	-0.02	0.08
R-squared	0.074	0.077	0.035	0.054	0.048	0.079	0.104
Total N	250	210	210	209	209	208	207
N in Placebo	95	78	78	77	77	76	75
N in Treat	155	132	132	132	132	132	132

Notes: The table presents the difference-in-means in the standardised imagery measures between the imagery and placebo arms for the subsample of participants in the "Territorial" sub-population and with a business at baseline. Standard errors clustered at the household level are presented in parenthesis. Columns 1 and 2 present results for non-business and business-specific imagery respectively. Columns 3 to 7 decompose the overall business imagery index into its constituent parts. We average across the two follow-up surveys and collapse the data to household level in the case where more than one household member participated in the training. The sample consists of people who confirmed participation in the training prior to knowing their treatment status. * p<0.10, ** p<0.05, *** p<0.01

Similarly, the economic effects of imagery training relative to the placebo

training are also relatively robust to restricting the sample to the "territorial" substratum, although they again lose some statistical significance. We find a strong positive effect for earnings pre-Covid: 0.243 standard deviations increase ($p = 0.099$). We also find a positive effect of 0.095 standard deviation increase ($p = 0.053$) on business survival during the Covid-19 pandemic. In contrast, we find no effect on earnings during the Covid-19 pandemic.

Table 2.12: Imagery vs Placebo: Effect on Economic Outcomes for "Territorial" Sub-Stratum

	Pre-COVID		COVID					COVID	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Earnings	Business survival	Earnings	Business survival	Safety nets	Business behaviour	Investment	Kessler	Psych resilience
Imagery training	0.244*	0.062	0.028	0.098*	-0.050	0.111	0.173*	-0.006	0.038
	(0.147)	(0.052)	(0.129)	(0.051)	(0.104)	(0.120)	(0.102)	(0.112)	(0.135)
Strata Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓
Survey Wave FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Wave-Subpop FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
R-squared	0.124	0.155	0.048	0.089	0.044	0.033	0.060	0.028	0.019
Mean DV (Placebo)	-0.19	0.74	-0.17	0.62	0.11	-0.10	-0.18	0	-0.16
Total N	262	268	450	463	458	461	470	458	458
N in Placebo	99	102	174	179	177	178	181	177	177
N in Imagery	163	166	276	284	281	283	289	281	281

Notes: The table presents difference-in-means for downstream economic and psychological outcomes when comparing the imagery to the placebo training arm for the "territorial" sub-population. Standard errors clustered at the household level are presented in parenthesis. We average across the two follow-up surveys and collapse the data to household level (the unit of randomisation) in the case where more than one household member participated in the training. We control for survey wave fixed effects. The sample consists of people who confirmed participation in the training prior to knowing their treatment status. * p<0.10, ** p<0.05, *** p<0.01

The downstream results on economic outcomes are qualitatively similar when comparing the imagery and placebo training to the no-intervention group for the "territorial" subgroup relative to the entire sample, as shown in Appendix Tables 2.22 and 2.23. However, we find strong negative effects on the imagery indices when comparing the placebo training to the no-intervention group, especially on the non-business imagery index. It is unclear whether results are being driven by differential attrition by treatment arm[46] or whether the business training truly made imagery worse for placebo training participants.

[46]72% of placebo respondents completed the last follow-up survey, compared to 63% of the no intervention group.

2.10 Concluding remarks

The ability to imagine the future is not just an innate trait, but it is also a skill that can be trained and invested in. Using data on nearly 2,000 micro-entrepreneurs in Colombia, we show that there is heterogeneity in how well and how often people use imagery in their economic decision making. Those who invest the effort to imagine the future also have better economic outcomes in the short and medium term.

We conjecture that the ability to imagine the future is malleable and can be taught. To explore this hypothesis further, we design an entrepreneurship training programme that teaches would-be micro-entrepreneurs the ability to imagine the future with vividness and leverage emotions constructively. Through a randomised control trial, we observe results that are strongly supportive of our theory. We show that the imagery training increases the use and quality of business imagery when compared to a placebo entrepreneurship training. Moreover, those who increase their imagery skills also have better earnings than participants who learn similar concepts in more abstract or verbal thinking. However, we find little evidence in favour of business training in general. When we compare the imagery training to a group who receives no training, we find small to no results across all outcomes of interest. When we turn to the placebo training group, we find that placebo participants report significantly lower earnings relative to the no-intervention group.

Our results also suggest that imagery training is particularly beneficial in addressing distortions about the way in which future scenarios are imagined. Several neuroscientific and psychology studies show that traumatic experiences can significantly impede the ability to imagine the future, by increasing the cost of using imagery and inducing overly negative future imagery. Indeed, we find that individuals with high levels of trauma symptoms have significantly more powerful negative imagery on average. We further conjecture that the programme focused on promoting positive imagery would be more effective for those exposed to significant past trauma. Consistent with this hypothesis, we find that those trained in imagery and who have experienced high baseline trauma have significantly more

vivid positive future imagery, compared to the placebo group. Moreover, there is preliminary evidence that their negative imagery is mediated by the training. High trauma individuals also experience larger improvements in pre-Covid-19 earnings compared to participants with similarly high levels of trauma in the placebo group. The results for the economic outcomes during the Covid-19 pandemic are again noisier and imprecisely estimated, but remain qualitatively consistent in terms of their direction, pointing towards a stronger effect for those participants who had experienced higher trauma. With higher baseline trauma and lower business experience, women are also more likely to benefit disproportionately from the imagery training, with significantly higher earnings both before and after the start of the Covid-19 pandemic compared to men.

Our paper presents evidence to suggest that imagery matters for economic decision making. Two open questions require further research. First, we are limited in our ability to shed light on the mechanisms and behaviours by which imagery affect earnings. The start of the Covid-19 pandemic at the launch of our in-person midline survey meant that we were constrained to using short phone surveys. To speak to these underlying mechanisms, we have implemented a short refresher session and collecting lab-in-the-field measures between July and September 2021, with the objective to capture the immediate-term effects of the training. Results will be available soon.

Second, further research is needed to understand why we observe limited and even negative effects for the business training when comparing the imagery and placebo training to the no-intervention control group. We currently do not have enough data to determine whether differential attrition, negative impacts of the placebo training or demand effects may be driving the results, although we have some reason to believe the first two explanations. More broadly, it would be useful to understand in which contexts imagery may be a useful complement to existing programmes, such as in schools. Imagery is most likely to be beneficial when used to make targeted future-oriented decisions, thereby motivating our decision to incorporate imagery within a training aimed to promote business performance. Similarly, imagery can be more easily scaled within the structure of group training

or video interventions focused on promoting a common goal that is subsequently applied to one's own situation, such as career choice.

Imagining the future is arguably foundational for making important economic decisions, ranging from human capital development to savings and investment. This is especially true in uncertain and risky domains, such as entrepreneurship. Our preliminary findings provide a strong proof of concept that investing the effort in imagining the future is a worthwhile endeavour and can be taught through public policies that can be scaled.

Appendices of Chapter 2

Table 2.13: Breakdown of sample by subdivision and treatment status

Subdivision	Imagery training	Placebo training	No-intervention group	Total
Youth	557	300	264	1121
Territorial	234	136	131	501
LGBT	12	12	12	36
Disabilities	59	41	17	117
Formerly homeless	39	35	29	103
Elderly	55	34	0	89
Total	956	558	453	1967

Notes: The "Territorial" subdivision includes victims of conflict and Venezuelan migrants. SDIS were unable to recruit enough participants for the smaller sub-divisions in some waves to construct a no-intervention group. Hence, we include wave-subdivision fixed effects in all our analysis.

Table 2.14: Measurement: Imagery

Family 1: Imagery	
Index	**Question(s)**
a. Spontaneous Use of Imagery Scale (SUIS)	Respondents were asked "On a scale of 1 to 5, how much do you agree with this statement?", where numbers 1 to 5 meant respectively: Strongly disagree; Disagree; Neither agree nor disagree; Agree; Strongly agree. The statements used were as follows: • When I need to go to a meeting, I picture the route in my mind before going. • When I think about a customer using my product or service, I imagine the customer's experience through pictures and sensations in my mind. • When I think about the day ahead, I create mental pictures of all the tasks I must do. • When I am faced with difficult situations, I mentally experience the actions I could take and the consequences of those actions before reacting. • When I think about the type of business I want to have, I live the experience of running that business in my mind. • When I feel overwhelmed, I find a mental place or time where I feel safe and calm. • When someone is upset with me, I live that person's experience in my mind to understand what might have caused the situation. • When I buy an asset for my business, an image of owning the asset pops up in my mind before buying it.
b. Adapted Prospective Imagery Task (PIT) Positive Emotional Valence	**Adapted Prospective Imagery Task (PIT) – positive statements and questions on emotional valence.** What is the intensity of the emotion produced in you by this image, using a scale from 1 to 5 where 1 means "no emotion at all", 2 "little, but weak emotions", 3 "moderate emotions", 4 "strong emotions" and 5 "extremely strong emotions"? • I first want you to imagine that the COVID-19 pandemic is over, and you save enough money to buy an asset you really want. • Now I want you to imagine the COVID-19 pandemic is over and you spend quality time with your family and friends. • Now I want you to imagine that the COVID-19 pandemic is over, and your business is doing well.
PIT Negative Emotional Valence	**Adapted Prospective Imagery Task (PIT) – negative statements and questions on emotional valence.** What is the intensity of the emotion produced in you by this image, using a scale from 1 to 5 where 1 means "no emotion at all", 2 "little, but weak emotions", 3 "moderate emotions", 4 "strong emotions" and 5 "extremely strong emotions"? • In the first scenario, I want you to imagine that the COVID-19 pandemic is over, and you are struggling to make ends meet. • Now I want you to imagine that the COVID-19 pandemic is over, and you have had a serious disagreement with someone close to you. • Now, I want you to imagine that the COVID-19 pandemic is over and your business closes
PIT Positive Vividness	**Adapted Prospective Imagery Task (PIT) – positive statements and questions on vividness of image.** Using a scale for the mental image where 1 means "no image at all", 2 means "vague and dim", 3 means "moderately clear and vivid", 4 means "reasonably clear and vivid" and 5 means "perfectly clear and vivid", how detailed is this image from 1 to 5? • I first want you to imagine that the COVID-19 pandemic is over, and you save enough money to buy an asset you really want. • Now I want you to imagine the COVID-19 pandemic is over and you spend quality time with your family and friends. • Now I want you to imagine that the COVID-19 pandemic is over, and your business is doing well.
e. PIT Negative Vividness	**Adapted Prospective Imagery Task (PIT) – negative statements and questions on vividness of image.** Using a scale for the mental image where 1 means "no image at all", 2 means "vague and dim", 3 means "moderately clear and vivid", 4 means "reasonably clear and vivid" and 5 means "perfectly clear and vivid", how detailed is this image from 1 to 5? • In the first scenario, I want you to imagine that the COVID-19 pandemic is over, and you are struggling to make ends meet. • Now I want you to imagine that the COVID-19 pandemic is over, and you have had a serious disagreement with someone close to you. • Now, I want you to imagine that the COVID-19 pandemic is over and your business closes.

Table 2.15: Measurement: Psychological resilience

Family 2: Psychological Resilience	
Index	**Question(s)**
a. Psychological Resilience	On a scale of 1 to 5, how much do you agree with the following statements? We will use the same scale as before, where 1 means "I strongly disagree" and 5 means "I strongly agree". **Brief Resilient Coping Scale (Sinclair and Wallston, 2004)** • I am looking for creative ways to alter difficult situations. • Regardless of what happens to me, I am controlling my reaction to it. • I am growing in positive ways by dealing with difficult situations. • I am actively looking for ways to replace the losses I am encountering in life. **Brief Resilience Scale (adapted subset) (Smith et al., 2008)** • I believe that I will bounce back quickly after the COVID pandemic. **Self-efficacy (Chen et al., 2004)** • I will be able to achieve most of the goals that I have set for myself. • When facing difficult tasks, I am certain that I will accomplish them. • In general, I think that I can obtain outcomes that are important to me. • I believe I can succeed at any endeavor to which I set my mind. • I will be able to successfully overcome many challenges. • I am confident that I can perform effectively on many different tasks. • Compared to other people, I can do most tasks very well. • Even when things are tough, I can perform quite well.
b. Anxiety and Depression Scale	**Kessler Scale (Kessler et al. (2002)-** Consider how often you have felt the following emotions over the past 30 days on a scale from 1 to 5, where 1 means "never", 2 means "rarely", 3 means "sometime", 4 means "almost always" and 5 means "always": • About how often during the past 30 days did you feel nervous? • About how often during the past 30 days did you feel hopeless? • About how often during the past 30 days did you feel restless or fidgety? • About how often during the past 30 days did you feel so depressed that nothing could cheer you up? • About how often during the past 30 days did you feel that everything was difficult? • About how often during the past 30 days did you feel worthless?

Table 2.16: Measurement: Economic Outcomes Pre-Covid

Family 3a: Economic Outcomes Pre-Covid	
Index	**Question(s)**
a. Business Status Pre-Covid	• Dummy equal to 1 if the person had a business pre-COVID, 0 otherwise
b. Index constructed as: (Sales + Income) if business + Income if no business	• Please reflect on your best month of sales in the six months prior to the start of the lockdown. How much revenue did your main business receive from sales in total that month? Revenues mean every peso received in the business in exchange for a product or service sold to a customer. • How much income did you take home during a typical week in February this year?
c. Investment Pre-Covid	In the past nine months, since September 2019, did you invest in the following categories for your business, to acquire a new asset or significantly improve an existing asset? (yes/no answers) A. Tools and utensils for manual work B. Machinery and equipment for production C. Vehicles used in your business D. Land, space in a shop or building E. Other physical assets F. Training for yourself G. Software or computer programs

Table 2.17: Measurement: Economic Outcomes During Covid

Family 3b: Economic Outcomes During Covid	
Index	**Question(s)**
a. Business Status During Covid	• Dummy equal to 1 if the person has a business which is NOT permanently closed, 0 if the person has no business or the business has permanently closed during COVID (since March 2020)
b. Earnings	Index constructed as: (Sales + Income) if business (currently open or temporarily closed) + Incomes if no business (permanently closed and/or no business pre-COVID) • In the last 30 days, how much revenue did your main business receive from sales? Revenues mean every peso received the business in exchange for a product or service sold to a customer. • In total, how much income did you personally take home last week?
c. Actual and perceived safety nets during COVID-19	Please tell us how much you agree with the following statements using the same scale as before from 1 to 5, where 1 means that "I strongly disagree" and 5 means that "I strongly agree". • I had an adequate amount of personal savings to ensure my safety for the first two months of the lockdown (scale 1-5) • I have enough cash to cover my ongoing expenses for the next week (scale 1-5) • If I had to find $200,000 pesos in the next month, I would be able to obtain them with ease (scale 1-5) *Savings* • Please reflect on your best savings week in the six months prior the lockdown. How much did you save during that week, from all your income-generating activities?
d. Business behavioural response to COVID-19	I am now going to ask you questions about your business response to the COVID- 19 pandemic. Please respond yes, no or doesn't apply, if the question does not apply to you. (yes/no answers) • Have you rearranged for yourself or workers to work from home? • Have you adapted your main business to meet the social distance criteria and adopt safety and sanitation measures so that your customers and workers are less exposed to COVID-19? • Have you identified alternative ways to access raw materials or alternate suppliers, should there be disruptions in your supply chain? • Do you consult your suppliers and customers more regularly to assess their situations? • Do you work out your operational costs more frequently than before the lockdown (such as rent, supplies, among others)? • Have you identified new business opportunities for your main business that could increase sales, which still adhere with government regulations? • Has your main business begun or increased the use of the Internet, online social networks, apps or digital platforms to sell your products or services? • Have you had conversations with the people or institutions who have lent your credit to assess flexibility on loan requirements? • Have you collaborated or talked with other entrepreneurs like you to share health and safety practices, stock or equipment, among others? • Has your main business requested (or is benefiting from) any government measure, either local or national, issued in response to the COVID-19 outbreak? • If open: Have you started selling new products or services in your main business, since March 24, 2020? • If closed temporarily: Do you plan to make changes to your business before reopening, such as changing your products or services or the way in which you produce or distribute? • Have you opened a new business since the lockdown started that is still operating, in other words, since March 24, 2020? **"Safe" working hours sub-index** *Sub-index of "safe" working hours constructed as the sum of total hours worked, multiplied by the share of hours worked in safety* • How many hours did you work specifically on your main business last week? • How many hours did you work specifically on your new business last week? • How many hours did you work in total last week for someone else, for a wage? • Of the hours you said you worked last week, how many hours did you work following social distancing, frequent handwashing, use of face mask and working from home if possible? (0, 1⁄4, 1⁄2, 3⁄4, 1)

Table 2.18: Effect on non-business imagery for business sample

		Non-business imagery sub-components				
	(1) Imagery (non-business)	(2) SUIS	(3) Positive PIT Vivid	(4) Positive PIT Emotion	(5) Negative PIT Vivid	(6) Negative PIT Emotion
Imagery training	0.007	-0.048	0.053	-0.012	0.034	-0.005
	(0.080)	(0.075)	(0.079)	(0.082)	(0.073)	(0.073)
Strata Controls	✓	✓	✓	✓	✓	✓
Survey Wave FE	✓	✓	✓	✓	✓	✓
Wave-Subpop FE	✓	✓	✓	✓	✓	✓
Mean DV in Placebo	-0.01	-0.05	-0.03	0	0.08	-0.02
R-squared	0.038	0.033	0.038	0.044	0.044	0.030
Total N	936	929	929	931	931	935
N in Placebo	341	339	339	339	339	341
N in Imagery	595	590	590	592	592	594

Notes: The table presents standardised mean treatment effects on non-business imagery of receiving the imagery training relative to the placebo training for the subsample of participants with a business at baseline. Standard errors clustered at the household level are presented in parenthesis. Columns 1 present the treatment effect for the overall non-business imagery index. Column 2 presents the SUIS index, which captures the frequency with which people use imagery. Columns 3 and 4 present treatment effects for the self-reported vividness and emotion experienced in the PIT scale that captures the quality of imagery used, specifically for positive future business scenarios. Columns 5 and 6 report the quality of imagery used for negative future business scenarios. False discovery rate q-values for the three imagery sub-hypotheses are calculated following the sharpened two-stage procedure of Benjamini et al. (2006). We pool the two follow-up surveys and collapse the data to household level in the case where more than one household member participated in the training. The sample consists of people who confirmed participation in the training prior to knowing their treatment status. * $p<0.10$, ** $p<0.05$, *** $p<0.01$

Table 2.19: Effect on non-business imagery for entire sample

		Non-business imagery sub-components				
	(1) Imagery (non-business)	(2) SUIS	(3) Positive PIT Vivid	(4) Positive PIT Emotion	(5) Negative PIT Vivid	(6) Negative PIT Emotion
Imagery training	-0.006	-0.127**	0.022	0.013	0.029	0.047
	(0.061)	(0.058)	(0.062)	(0.059)	(0.056)	(0.056)
Strata Controls	✓	✓	✓	✓	✓	✓
Survey Wave FE	✓	✓	✓	✓	✓	✓
Wave-Subpop FE	✓	✓	✓	✓	✓	✓
Mean DV in Placebo	0	0.01	-0.03	-0.01	0.05	-0.04
R-squared	0.017	0.020	0.018	0.018	0.023	0.040
Total N	1,796	1,795	1,788	1,788	1,790	1,790
N in Placebo	656	656	653	653	654	654
N in Imagery	1,140	1,139	1,135	1,135	1,136	1,136

Notes: The table presents standardised mean treatment effects on non-business imagery of receiving the imagery training relative to the placebo training for the sample, both with and without a business at baseline. Standard errors clustered at the household level are presented in parenthesis. Columns 1 present the treatment effect for the overall non-business imagery index. Column 2 presents the SUIS index, which captures the frequency with which people use imagery. Columns 3 and 4 present treatment effects for the self-reported vividness and emotion experienced in the PIT scale that captures the quality of imagery used, specifically for positive future business scenarios. Columns 5 and 6 report the quality of imagery used for negative future business scenarios. False discovery rate q-values for the three imagery sub-hypotheses are calculated following the sharpened two-stage procedure of Benjamini et al. (2006). We pool the two follow-up surveys and collapse the data to household level in the case where more than one household member participated in the training. The sample consists of people who confirmed participation in the training prior to knowing their treatment status. * $p<0.10$, ** $p<0.05$, *** $p<0.01$

Table 2.20: Imagery vs No-intervention: Effect on Downstream Outcomes

			Pre-COVID		COVID					COVID	
	(1) Imagery non-business	(2) Imagery business	(3) Earnings	(4) Business survival	(5) Earnings	(6) Business survival	(7) Safety nets	(8) Business behaviour	(9) Investment	(10) Kessler	(11) Psych resilience
Imagery training	-0.013	0.052	-0.007	0.031	0.040	0.033	0.047	-0.003	0.061	0.011	-0.049
	(0.093)	(0.098)	(0.066)	(0.029)	(0.057)	(0.028)	(0.061)	(0.063)	(0.060)	(0.064)	(0.066)
Strata Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Survey Wave FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Wave-Subpop FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Mean DV (No-int)	0	0	0	0.69	0	0.65	0	0	0	0	0
R-squared	0.042	0.056	0.056	0.187	0.031	0.098	0.030	0.020	0.054	0.037	0.016
Total N	873	753	988	1,012	1,654	1,697	1,680	1,693	1,723	1,678	1,679
N in No-int	278	241	323	333	539	552	546	551	564	545	545
N in Imagery	595	513	665	679	1,115	1,145	1,134	1,142	1,159	1,133	1,134

Notes: The table presents standardised mean treatment effects of receiving the imagery training relative to the no-intervention group. Standard errors clustered at the household level are presented in parenthesis. False discovery rate q-values over the family of economic outcomes and the family of psychological outcomes are calculated following the sharpened two-stage procedure of Benjamini et al. (2006). We pool the two follow-up surveys and collapse the data to household level in the case where more than one household member participated in the training. The sample consists of people who confirmed participation in the training prior to knowing their treatment status. * p<0.10, ** p<0.05, *** p<0.01

Table 2.21: Placebo Training vs No-intervention: Effect on Downstream Outcomes

			Pre-COVID		COVID					COVID	
	(1) Imagery non-business	(2) Imagery business	(3) Earnings	(4) Business survival	(5) Earnings	(6) Business survival	(7) Safety nets	(8) Business behaviour	(9) Investment	(10) Kessler	(11) Psych resilience
Placebo training	-0.045	-0.131	-0.222**	0.001	-0.144**	-0.022	0.014	-0.074	-0.048	-0.012	-0.052
	(0.099)	(0.111)	(0.096)	(0.032)	(0.071)	(0.031)	(0.068)	(0.067)	(0.066)	(0.070)	(0.072)
Strata Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Survey Wave FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Wave-Subpop FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Mean DV (No-int)	0	0	0	0.69	0	0.65	0	0	0	0	0
R-squared	0.044	0.066	0.087	0.210	0.072	0.117	0.079	0.033	0.065	0.070	0.036
Total N	619	525	703	725	1,181	1,211	1,196	1,208	1,231	1,192	1,194
N in No-int	278	241	323	333	539	552	546	551	564	545	545
N in Placebo	341	285	380	392	642	659	650	657	667	647	649

Notes: The table presents standardised mean treatment effects of receiving the placebo training relative to the no-intervention group. Standard errors clustered at the household level are presented in parenthesis. False discovery rate q-values over the family of economic outcomes and the family of psychological outcomes are calculated following the sharpened two-stage procedure of Benjamini et al. (2006). We pool the two follow-up surveys and collapse the data to household level in the case where more than one household member participated in the training. The sample consists of people who confirmed participation in the training prior to knowing their treatment status. * p<0.10, ** p<0.05, *** p<0.01

Table 2.22: Imagery vs No int: Effect on Downstream Outcomes for "Territorial"

			Pre-COVID		COVID					COVID	
	(1) Imagery non-business	(2) Imagery business	(3) Earnings	(4) Business survival	(5) Earnings	(6) Business survival	(7) Safety nets	(8) Business behaviour	(9) Investment	(10) Kessler	(11) Psych resilience
Imagery training	-0.143	-0.029	0.052	0.053	-0.107	0.022	0.075	0.023	0.012	0.010	-0.067
	(0.123)	(0.146)	(0.112)	(0.050)	(0.115)	(0.049)	(0.109)	(0.125)	(0.108)	(0.115)	(0.126)
Strata Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Survey Wave FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Wave-Subpop FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Mean DV (No-int)	0.05	0.08	0	0.740	0	0.710	0	0	0	0	0
R-squared	0.060	0.063	0.066	0.143	0.044	0.073	0.039	0.026	0.054	0.030	0.026
Total N	459	354	269	274	449	461	458	460	471	457	457
N in No-int	177	137	106	108	173	177	177	177	182	176	176
N in Imagery	282	217	163	166	276	284	281	283	289	281	281

Notes: The table presents standardised mean treatment effects of receiving the imagery training relative to the no-intervention group for the "Territorial" sub-population only. Standard errors clustered at the household level are presented in parenthesis. We pool the two follow-up surveys and collapse the data to household level in the case where more than one household member participated in the training. The sample consists of people who confirmed participation in the training prior to knowing their treatment status. * p<0.10, ** p<0.05, *** p<0.01

Table 2.23: Placebo vs No int: Effect on Downstream Outcomes for "Territorial"

			Pre-COVID		COVID					COVID	
	(1) Imagery non-business	(2) Imagery business	(3) Earnings	(4) Business survival	(5) Earnings	(6) Business survival	(7) Safety nets	(8) Business behaviour	(9) Investment	(10) Kessler	(11) Psych resilience
Placebo training	-0.372***	-0.263*	-0.198	-0.010	-0.175	-0.074	0.083	-0.094	-0.180	-0.004	-0.162
	(0.139)	(0.155)	(0.177)	(0.056)	(0.133)	(0.056)	(0.116)	(0.119)	(0.115)	(0.123)	(0.136)
Strata Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Survey Wave FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Wave-Subpop FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Mean DV (No-int)	0.05	0.08	0	0.740	0	0.710	0	0	0	0	0
R-squared	0.062	0.100	0.112	0.234	0.103	0.133	0.097	0.055	0.094	0.066	0.035
Total N	355	257	205	210	347	356	354	355	363	353	353
N in No-int	177	137	106	108	173	177	177	177	182	176	176
N in Placebo 178	120	99	102	174	179	177	178	181	177	177	

Notes: The table presents standardised mean treatment effects of receiving the placebo training relative to the no-intervention group for the "Territorial" sub-population only. Standard errors clustered at the household level are presented in parenthesis. We pool the two follow-up surveys and collapse the data to household level in the case where more than one household member participated in the training. The sample consists of people who confirmed participation in the training prior to knowing their treatment status. * p<0.10, ** p<0.05, *** p<0.01

Table 2.24: Effect on downstream outcomes, by trauma

	Pre-COVID		COVID					COVID	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Earnings	Business survival	Earnings	Business survival	Safety nets	Business behaviour	Investment	Kessler	Psych resilience
Imagery training	0.199*	0.012	0.109	0.064**	0.043	0.071	0.101	0.008	-0.026
	(0.105)	(0.032)	(0.075)	(0.031)	(0.070)	(0.070)	(0.066)	(0.066)	(0.076)
High trauma	-0.154	-0.012	-0.150	-0.021	-0.194*	0.025	-0.026	-0.504***	-0.032
	(0.199)	(0.053)	(0.121)	(0.052)	(0.110)	(0.112)	(0.107)	(0.116)	(0.127)
High trauma*Imagery	0.160	0.076	0.085	0.032	-0.020	-0.072	0.018	0.093	-0.039
	(0.229)	(0.066)	(0.147)	(0.064)	(0.139)	(0.140)	(0.134)	(0.144)	(0.161)
Imag*Trauma=0 (p-value)	0.485	0.249	0.564	0.621	0.886	0.610	0.894	0.521	0.806
Imag*Trauma=0 (q-value)	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Imag*Trauma+Imag=0 (p-value)	0.076	0.127	0.124	0.091	0.848	0.993	0.307	0.433	0.647
Imag*Trauma+Imag=0 (q-value)	0.402	0.402	0.402	0.402	1.000	1.000	0.443	0.564	0.859
Strata Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓
Survey Wave FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Wave-Subpop FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Mean DV (No-intervention)	-0.18	0.73	-0.090	0.65	0.05	-0.09	-0.04	0.13	-0.01
R-squared	0.079	0.184	0.046	0.109	0.051	0.025	0.068	0.079	0.030
Total N	907	929	1,530	1,569	1,552	1,564	1,588	1,549	1,551
N in Placebo	380	392	642	659	650	657	667	647	649
N in Imagery	665	679	1115	1145	1134	1142	1159	1133	1134

Notes: The table presents heterogeneous treatment effects on downstream economic and psychological outcomes by trauma, when comparing the imagery to the placebo training arm. Standard errors clustered at the household level are presented in parenthesis. We average across the two follow-up surveys and collapse the data to household level (the unit of randomisation) in the case where more than one household member participated in the training. We control for survey wave and subdivision-wave interaction fixed effects. The sample consists of people who confirmed participation in the training prior to knowing their treatment status. * p<0.10, ** p<0.05, *** p<0.01

Table 2.25: Attrition, conditional on acceptance

	(1) Imagery Treatment		(2) Placebo Treatment		(3) Pure Control		T-test Difference			F-test for joint
Variable	N	Mean/SE	N	Mean/SE	N	Mean/SE	(1)-(2)	(1)-(3)	(2)-(3)	orthogonality
Female	795	0.574	461	0.579	415	0.571	-0.006	0.003	0.008	0.032
		(0.018)		(0.023)		(0.024)				
Age (18-28)	795	0.591	461	0.573	415	0.629	0.019	-0.038	-0.056*	1.491
		(0.017)		(0.023)		(0.024)				
Age (29-45)	795	0.245	461	0.249	415	0.272	-0.004	-0.027	-0.023	0.548
		(0.015)		(0.020)		(0.022)				
Age (46-59)	795	0.096	461	0.119	415	0.094	-0.024	0.002	0.025	1.079
		(0.010)		(0.015)		(0.014)				
Years Education	657	13.157	387	13.035	327	13.366	0.122	-0.209	-0.331	0.877
		(0.129)		(0.182)		(0.173)				
Only Business Owner	795	0.245	461	0.245	415	0.236	0.000	0.009	0.009	0.070
		(0.015)		(0.020)		(0.021)				
Only Have a Business Idea	795	0.477	461	0.477	415	0.494	-0.000	-0.017	-0.017	0.182
		(0.018)		(0.023)		(0.025)				
Have Business and Business Idea	795	0.278	461	0.275	415	0.270	0.002	0.008	0.006	0.045
		(0.016)		(0.021)		(0.022)				
Income Strat: Sales $\leq$ 1 month salary	795	0.665	461	0.677	415	0.641	-0.011	0.024	0.036	0.655
		(0.017)		(0.022)		(0.024)				
Income Strat: Sales $>$ 1 month salary	795	0.238	461	0.221	415	0.248	0.016	-0.010	-0.027	0.455
		(0.015)		(0.019)		(0.021)				
Refused to Provide Income	795	0.097	461	0.102	415	0.111	-0.005	-0.014	-0.009	0.292
		(0.010)		(0.014)		(0.015)				
Assets Owned	657	11.661	387	11.894	327	11.835	-0.233	-0.175	0.058	0.508
		(0.150)		(0.201)		(0.209)				
Household size	657	3.452	386	3.609	326	3.650	-0.157	-0.197*	-0.041	2.143
		(0.059)		(0.082)		(0.095)				
Save Monthly Y/N	654	0.549	387	0.535	326	0.580	0.014	-0.031	-0.045	0.748
		(0.019)		(0.025)		(0.027)				
Had Access to Credit	654	0.153	386	0.150	326	0.138	0.003	0.015	0.012	0.195
		(0.014)		(0.018)		(0.019)				
No. of Traumatic Events	795	2.247	461	2.342	415	2.308	-0.095	-0.061	0.034	0.249
		(0.083)		(0.110)		(0.125)				
Reported Experiencing Trauma (Y/N)	657	0.826	387	0.837	326	0.844	-0.011	-0.017	-0.006	0.254
		(0.015)		(0.019)		(0.020)				
Impact of Event Score	649	17.645	386	17.142	324	17.785	0.503	-0.140	-0.643	0.091
		(0.874)		(1.120)		(1.164)				
Kessler Score	657	13.098	387	13.187	326	13.138	-0.090	-0.040	0.050	0.061
		(0.160)		(0.202)		(0.220)				
Wave 1 Participant	795	0.307	461	0.384	415	0.304	-0.077***	0.003	0.080**	4.655***
		(0.016)		(0.023)		(0.023)				

Notes: The table presents a balance test on the sample who accepted training for all three treatment arms. Columns 1, 2, and 3 show the sample size, mean, and standard errors for the imagery treatment, placebo treatment, and no-intervention groups, respectively. The following 3 columns show the differences between treatment arms and p-values for t-tests measuring differences between treatment arms. The final column presents an F-test across all three treatment arms. * p<0.10, ** p<0.05, *** p<0.01

Table 2.26: Attrition in the first follow-up survey

	(1) Imagery Treatment		(2) Placebo Treatment		(3) Pure Control		T-test Difference			F-test for joint
Variable	N	Mean/SE	N	Mean/SE	N	Mean/SE	(1)-(2)	(1)-(3)	(2)-(3)	orthogonality
Female	681	0.567	392	0.602	304	0.569	-0.035	-0.002	0.033	0.686
		(0.019)		(0.025)		(0.028)				
Age (18-28)	681	0.608	392	0.569	304	0.605	0.039	0.003	-0.036	0.853
		(0.019)		(0.025)		(0.028)				
Age (29-45)	681	0.223	392	0.258	304	0.276	-0.034	-0.053*	-0.019	1.852
		(0.016)		(0.022)		(0.026)				
Age (46-59)	681	0.103	392	0.133	304	0.115	-0.030	-0.012	0.018	1.100
		(0.012)		(0.017)		(0.018)				
Years Education	581	12.917	338	13.015	257	13.126	-0.098	-0.209	-0.112	0.354
		(0.137)		(0.198)		(0.198)				
Only Business Owner	681	0.253	392	0.255	304	0.240	-0.003	0.012	0.015	0.117
		(0.017)		(0.022)		(0.025)				
Only Have a Business Idea	681	0.493	392	0.485	304	0.500	0.009	-0.007	-0.015	0.083
		(0.019)		(0.025)		(0.029)				
Have Business and Business Idea	681	0.254	392	0.258	304	0.260	-0.004	-0.006	-0.002	0.021
		(0.017)		(0.022)		(0.025)				
Income Strat: Sales $\leq$ 1 month salary	681	0.687	392	0.696	304	0.674	-0.009	0.013	0.022	0.194
		(0.018)		(0.023)		(0.027)				
Income Strat: Sales $>$ 1 month salary	681	0.229	392	0.202	304	0.220	0.028	0.009	-0.019	0.552
		(0.016)		(0.020)		(0.024)				
Refused to Provide Income	681	0.084	392	0.102	304	0.105	-0.018	-0.022	-0.003	0.800
		(0.011)		(0.015)		(0.018)				
Assets Owned	581	11.482	338	11.695	257	11.705	-0.213	-0.223	-0.010	0.474
		(0.159)		(0.211)		(0.234)				
Household Size	580	3.536	337	3.641	256	3.749	-0.105	-0.213*	-0.108	1.641
		(0.065)		(0.086)		(0.110)				
Save Monthly Y/N	578	0.559	338	0.536	256	0.559	0.023	0.000	-0.023	0.262
		(0.021)		(0.027)		(0.031)				
Had Access to Credit	578	0.152	337	0.139	256	0.137	0.013	0.016	0.003	0.234
		(0.015)		(0.019)		(0.022)				
No. of Traumatic Events	681	2.360	392	2.583	304	2.550	-0.223	-0.190	0.033	1.261
		(0.091)		(0.126)		(0.148)				
Reported Experiencing Trauma (Y/N)	580	0.833	338	0.861	256	0.855	-0.028	-0.023	0.005	0.766
		(0.016)		(0.019)		(0.022)				
Impact of Event Score	573	18.348	337	18.539	254	18.385	-0.190	-0.036	0.154	0.008
		(0.935)		(1.252)		(1.339)				
Kessler Score	580	13.073	338	13.376	256	13.121	-0.302	-0.048	0.254	0.608
		(0.172)		(0.221)		(0.250)				
Wave 1 Participant	681	0.300	392	0.362	304	0.296	-0.063**	0.004	0.066*	2.645*
		(0.018)		(0.024)		(0.026)				

Notes: The table presents a balance test at the first follow up survey for all three treatment arms. Columns 1, 2, and 3 show the sample size, mean, and standard errors for the imagery treatment, placebo treatment, and no-intervention groups, respectively. The following 3 columns show the differences between treatment arms and p-values for t-tests measuring differences between treatment arms. The final column presents an F-test across all three treatment arms. * p<0.10, ** p<0.05, *** p<0.01

Table 2.27: Treatment effects on imagery, by trauma (non-business scenarios)

	Combined imagery index		Non-business imagery sub-components					Perceived likelihood	
	(1) Imagery (non-business)	(2) Imagery (business)	(3) SUIS	(4) Positive PIT Vivid	(5) Positive PIT Emotion	(6) Negative PIT Vivid	(7) Negative PIT Emotion	(8) Positive Likelihood	(9) Negative Likelihood
Imagery training	-0.024	0.201*	-0.016	-0.043	-0.170*	0.013	0.032	0.107	-0.061
	(0.092)	(0.113)	(0.096)	(0.090)	(0.087)	(0.084)	(0.087)	(0.086)	(0.087)
High trauma	0.013	0.208	0.094	-0.317*	-0.377**	0.456***	0.115	0.380***	-0.295*
	(0.158)	(0.181)	(0.159)	(0.166)	(0.160)	(0.144)	(0.133)	(0.132)	(0.169)
High trauma*Imagery	0.209	0.080	-0.011	0.436**	0.615***	-0.108	0.083	0.011	0.279
	(0.209)	(0.236)	(0.202)	(0.212)	(0.205)	(0.181)	(0.186)	(0.182)	(0.214)
Test: Imag*Trauma+Imag=0 (p-value)	0.330	0.180	0.880	0.040	0.020	0.560	0.490	0.470	0.270
Strata Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓
Survey Wave FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Wave-Subpop FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Mean DV (Placebo+High Trauma)	-0.03	-0.14	-0.07	0.02	0.05	-0.10	0.04	-0.17	0.02
R-squared	0.046	0.057	0.035	0.053	0.051	0.067	0.050	0.075	0.039
Total N	820	697	813	813	813	815	815	815	819
N in Placebo	341	285	339	339	339	339	339	339	341
N in Imagery	595	513	590	590	590	592	592	592	594

Notes: The table presents standardised mean treatment effects on imagery of receiving the imagery training relative to the placebo training for the subsample of participants with a business at baseline. Standard errors clustered at the household level are presented in parenthesis. Columns 1 and 2 present treatment effects for non-business and business-specific imagery respectively. Columns 3 to 7 reports results for the five sub-components of the imagery index for non-business imagery. Column 3 presents the SUIS index, which captures the frequency with which people use imagery. Columns 4 and 5 present treatment effects for the self-reported vividness and emotion experienced when imagining positive future non-business scenarios in the PIT scale that captures the quality of imagery used. Columns (5) and (6) report the quality of imagery used for negative future non-business scenarios. False discovery rate q-values for the five imagery sub-hypotheses are calculated following the sharpened two-stage procedure of Benjamini et al. (2006). Columns 8 and 9 report the perceived likelihood of the positive and negative future non-business scenarios imagined through the PIT occurring in the participant's life, which were not pre-specified outcomes of interest. We pool the two follow-up surveys and collapse the data to household level in the case where more than one household member participated in the training. The sample consists of people who confirmed participation in the training prior to knowing their treatment status.
* p<0.10, ** p<0.05, *** p<0.01

Table 2.28: Entire sample: Treatment effects on imagery (business scenarios)

	Combined imagery index		Business-specific sub-components				
	(1) Imagery (non-business)	(2) Imagery (business)	(3) SUIS	(4) Positive PIT Vivid	(5) Positive PIT Emotion	(6) Negative PIT Vivid	(7) Negative PIT Emotion
Imagery training	0.020	0.201**	0.093	0.075	0.061	0.189**	0.157*
	(0.078)	(0.092)	(0.089)	(0.092)	(0.085)	(0.082)	(0.080)
FDR q-values			0.399	0.399	0.399	0.119	0.119
Strata Controls	✓	✓	✓	✓	✓	✓	✓
Survey Wave FE	✓	✓	✓	✓	✓	✓	✓
Wave-Subpop FE	✓	✓	✓	✓	✓	✓	✓
Mean DV in Placebo	-0.04	-0.11	-0.05	-0.07	-0.07	-0.11	-0.01
R-squared	0.033	0.045	0.031	0.045	0.025	0.026	0.036
Total N	1,005	851	850	845	845	842	842
N in Placebo	367	304	304	302	302	301	300
N in Imagery	638	547	546	543	543	541	542

Notes: The table presents standardised mean treatment effects of receiving the imagery training relative to the placebo training for the subsample of participants with a business at baseline. Standard errors clustered at the household level are presented in parenthesis. Columns (1) and (2) present treatment effects for non-business and business-specific imagery respectively. Column (3) presents the SUIS index, which captures the frequency with which people use imagery. Columns (4) and (5) present treatment effects for the self-reported vividness and emotion experienced in the PIT scale that captures the quality of imagery used, specifically for positive future business scenarios. Columns (5) and (6) report the quality of imagery used for negative future business scenarios. False discovery rate q-values for the three imagery sub-hypotheses are calculated following the sharpened two-stage procedure of Benjamini et al. (2006). We pool the two follow-up surveys and collapse the data to household level in the case where more than one household member participated in the training.
* p<0.10, ** p<0.05, *** p<0.01

Table 2.29: Entire Sample: Treatment Effects on Downstream Outcomes

	Pre-COVID		COVID					COVID	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Earnings	Business survival	Earnings	Business survival	Safety nets	Business behaviour	Investment	Kessler	Psych resilience
Imagery training	0.208**	0.011	0.164***	0.033	0.048	0.082	0.036	0.004	-0.019
	(0.081)	(0.026)	(0.059)	(0.024)	(0.054)	(0.054)	(0.051)	(0.055)	(0.060)
FDR q-values	0.037	0.418	0.037	0.267	0.418	0.267	0.418	1.000	1.000
Strata Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓
Survey Wave FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Wave-Subpop FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
R-squared	0.073	0.181	0.045	0.106	0.043	0.026	0.058	0.052	0.021
Mean DV (Placebo)	-0.23	0.71	-0.17	0.64	-0.02	-0.08	-0.05	0	-0.03
Total N	1,157	1,187	1,942	1,993	1,971	1,988	2,020	1,967	1,970
N in Placebo	423	436	716	733	724	731	744	721	723
N in Imagery	734	751	1226	1,260	1,247	1,257	1,276	1,246	1,247

Notes: The table presents standardised mean treatment effects of receiving the imagery training relative to the placebo training. Standard errors clustered at the household level are presented in parenthesis. False discovery rate q-values over the family of economic outcomes and the family of psychological outcomes are calculated following the sharpened two-stage procedure of Benjamini et al. (2006). We pool the two follow-up surveys and collapse the data to household level in the case where more than one household member participated in the training.
* p<0.10, ** p<0.05, *** p<0.01

CHAPTER 3

Cash Transfers and Crowding Out of Remittances: Evidence from South Africa

Abstract With social protection programmes rapidly expanding worldwide, the extent to which public cash transfers interact with private transfer behaviour remains an important welfare concern. This paper evaluates whether receiving a state old-age pension crowds out remittances in a middle-income country context where private transfers are pervasive. I estimate the crowding out effect using a fuzzy regression discontinuity design in post-Apartheid South Africa, whereby I exploit the discontinuity in the probability of receiving the pension created by the age eligibility rules. I find that a 1 rand increase in pension income crowds out 0.22 rand in remittance income on average. I also find evidence to suggest that the nature of the relationship between the sender and receiver matters, with stronger effects for households with migrants. Failing to account for the endogenous response in remittances may lead to an overstatement of the impact of public cash transfers on poverty alleviation for direct recipients.

Acknowledgements *I am grateful for insightful comments and feedback from Abi Adams-Prassl, Stefan Dercon, Simon Quinn, and the applied microeconomics student group at the University of Oxford.*

3.1 Introduction

Public social protection systems have rapidly expanded in recent years. By mid-2021, a total of 782 cash transfer programmes had been planned or implemented in more than 182 countries, largely in response to the Covid-19 pandemic (Gentilini et al., 2021). While there is strong evidence that cash transfers are effective in smoothing consumption, there is also growing evidence that public cash transfers could interact with other household behaviours that can either reinforce or diminish their effectiveness (Ardington et al., 2009; Parker and Todd, 2017). The interaction between public and private transfers is especially important in determining the effect of cash transfers on overall household welfare (Cox, 1987). The extent to which public cash transfers crowd out private support is arguably most pertinent in contexts where private transfers are pervasive and social protection is expanding, as in many low- and middle-income countries (Cox and Jakubson, 1995; Jensen, 2003; Juarez, 2009; Amuedo-Dorantes and Juarez, 2015).

This paper evaluates the extent to which receiving a state cash transfer in the form of the old-age pension crowds out remittances in post-Apartheid South Africa. Remittances are an important source of private support in alleviating poverty.[1] These transfers are typically made from working-age populations to dependents, such as the elderly or children. Similarly, pensions serve an important role in targeting the elderly when their income is at its lowest in the life cycle. Yet the interaction between public pensions and remittances is theoretically ambiguous: the public cash transfer could either displace or supplement the private support. Identifying the dominant effect and the resulting welfare implications becomes an empirical question. For instance, in crowding out remittances, the pension could serve as a means to reduce the burden of supporting the elderly from working-age populations who are likely to have more productive use for the capital.

In this paper, I estimate the crowding out effect of receiving the public pension on remittances through a fuzzy regression discontinuity design, where I ex-

[1] In 2020, $540 billion in remittances were transferred globally, far exceeding both foreign direct investment and official development assistance in low- and middle-income countries (Ratha et al., 2021).

ploit the discontinuity in the probability of receiving the pension created by the age eligibility rules. I use data from the nationally-representative National Income Dynamics Study (NIDS) and focus on the year 2012, as the first survey year after a policy change to the age eligibility rules. Although I cannot speak to the overall welfare implications of the crowding out, I find that the younger generation typically send a large share of remittances to the receiving households in the sample. Furthermore, receiving the state pension crowds out 266.4 rand (approximately $33) in monthly household remittances on average for household near the pension age eligibility threshold. The crowding out effect is substantial in size: I estimate that a 1 rand increase in income received from the pension crowds out 0.22 rand in remittance income on average ($p < 0.05$), given that nearly all households receive the maximum monthly pension amount of 1,200 rand (approximately $147). Moreover, remittances remains lower even for pension-eligible households located far from the age eligibility threshold of 60 years. The results are robust to a wide range of alternative model specifications and placebo tests.

However, the crowding out effect is only observed in the year 2012, after a decade during which overall social spending had doubled in South Africa. There is no evidence of crowding out in 2017, using the latest survey wave of NIDS. In fact, the number of households sending remittances and the average level of remittances in society has increased over time, including for pension-eligible households. I present descriptive evidence to suggest that the divergence in results in 2017 may be partially driven by a change in the profiles of both those who send and receive remittances. Sending households are wealthier, whereas households receiving remittances appear qualitatively poorer in 2017 relative to 2012.

This paper contributes to a growing literature on the effect of cash transfers on household behaviour, with important implications for welfare. Cash transfers often intended for other purposes have been shown to affect a wide range of household decisions, such as education, health, labour supply and migration (Parker and Todd, 2017). The unintended effects of the South African old-age pension in particular has been subject to extensive research, as it is one of the most far-reaching and generous non-contributory pensions in low- and middle-income countries (Ferguson, 2015). Dubbed the “social" pension, the South African pension has been

shown to improve the health and education outcomes of co-resident children, change household composition and affect labour supply (Case and Deaton, 1998; Duflo, 2003; Bertrand et al., 2003; Edmonds et al., 2004; Edmonds, 2006; Ranchhod, 2006; Posel and Casale, 2006). Ardington et al. (2009) find that the South African pension induces labour migration among prime-aged adults, but could not speak to the question of remittances. In this paper, I present evidence that the old-age pension could also interact negatively with remittance behaviour, thereby diminishing income available to recipient households, but also potentially releasing capital for more productive means.

This paper also presents strong empirical evidence in support of the crowding out hypothesis within a middle-income country context. In high-income countries where social protection systems are well-established and private transfers are small, several empirical studies conclude that the threat of displacement posed by public transfers to private transfers may be overstated (Altonji et al., 1992, 1997; Künemund and Rein, 1999). In contrast, more recent empirical work has shown that crowding out may be a relevant concern in middle-income countries where credit markets do not function well, informal support is widespread and social protection may be expanding. The magnitude of the estimated crowding out effect ranges from 17% to 87% (Cox and Jakubson, 1995; Cox et al., 1998; Albarran and Attanasio, 2003; Cox et al., 2004; Juarez, 2009; Lai and Orsuwan, 2009; Gerardi and Tsai, 2014; Amuedo-Dorantes and Juarez, 2015).[2] Closest to my work, Jensen (2003) shows that the expansion of the South African pension to all races in 1993 displaced remittances for a sub-sample of rural migrant households. The literature on crowding out draws largely on panel data and instrumental variable estimation methods to address the endogeneity of income and self-selection into transfers, such as the triple difference strategy in Jensen (2003). To address these estimation issues, this paper is the first to employ a regression discontinuity design and focuses on modern-day South Africa where social spending had doubled within a decade (National Treasury, 2012). I find a crowding out effect of 0.22 rand

[2]The countries studied include the Philippines, Peru, Mexico and Taiwan. The variation in results may be partly explained by the type of public cash transfer received (Cox and Jakubson, 1995), non-linearities in the crowding out effect (Cox et al., 2004) or due to heterogeneity in the extent of crowding out by income status (Kazianga, 2006; Maitra and Ray, 2003).

for every 1 rand of pension income received on average, which is remarkably similar in size to the 0.25-0.30 rand reduction found by Jensen (2003) twenty years earlier.

The paper is structured as follows. In the next two sections, I motivate the context by providing an overview of the setting, data used and descriptive evidence. Section 3.4 outlines the empirical strategy and presents evidence in support of the identifying assumptions underlying the regression discontinuity design. In Section 3.5, I present results on the local average treatment effect. Section 3.6 demonstrates the robustness of my results to a wide range of alternative model specifications and placebo tests. I present results by year in Section 3.7 and Section 3.8 offers conclusions.

3.2 Context

Modern-day South Africa is characterised by widespread informal support networks and a generous social protection system that has recently expanded. Its Apartheid history of harsh restrictions on settlement and employment induced an oscillatory migration pattern, with individuals temporarily migrating to places of employment (largely domestically) while retaining membership with their households of origin.[3] Even with the removal of Apartheid restrictions in 1994, this circular migrant labour system has remained ingrained, in part due to the HIV epidemic and high rates of unemployment (Bhorat, 2004; Posel and Casale, 2006). Remittances have played an integral role in supporting low-income households as a result. Yet in the first fifteen years after Apartheid, there has been a notable drop in the proportion of households with migrant workers receiving remittances (Posel, 2009).

The fall in remittances coincides with the expansion in social protection. While the old-age pension predominantly benefited white South Africans under the Apartheid regime[4], social protection was rapidly expanded to all races shortly af-

[3]In particular, a harsh closed compound mining system permitted only workers to live onsite and black South Africans were denied residency rights in urban areas without a work permit under Apartheid (Posel, 2010).

[4]During Apartheid, discrimination against black South Africans amounted to lower benefits and an inefficient delivery system, thus resulting in a lower reach (Woolard and Leibbrandt, 2010).

ter the end of Apartheid in the early 1990s and again in the 2000s, corresponding with increasing inequality and endemic unemployment (Ferguson, 2015). Overall social spending doubled in real terms between 2002 and 2012 (National Treasury, 2012). By 2012, approximately 30% of the population received a state grant, of which three million aged 60 years and older received the old-age pension (Statistics South Africa, 2013a).[5]

Several features of the old-age pension make it attractive for empirical analysis. Firstly, eligibility for the non-contributory state pension depends primarily on age through a clearly defined targeting rule. Although the pension is in theory means tested and based on the recipient's income, the test has been found not to be binding on most black households (Case and Deaton, 1998; Bertrand et al., 2003; Duflo, 2003; Jensen, 2003). As a result, most households receive the maximum pension payment amount in practice.[6] Secondly, take up of the state pension is high and its value is substantial relative to the earning capacity of many elderly people. In 2010, the South African old age pension amounted to 1.75 times the median monthly per capita income (Woolard and Leibbrandt, 2010).

3.3 Data

The National Income Dynamics Study (NIDS) is the first nationally representative longitudinal survey of over 28,000 individuals and their households in South Africa. Collected in two-year increments since 2008, the study tracks these individuals over five waves. Broadly, the data describe household composition, income, labour market participation, migration, health and education. Moreover, for the purpose of this paper, the panel surveys capture a rich picture of all income sources, including government transfers and remittances.

Using NIDS, I aggregate the analysis to the household level and construct the key variables of interest as follows. The age of the oldest household member forms the running variable in the regression discontinuity design. A dummy variable for

[5]The grants have an even wider coverage at the household level, with 44% of all households nationwide receiving at least one grant (Statistics South Africa, 2013a).

[6]The maximum pension amount was R940, R1080, R1200, R1250 and R1600 in 2008, 2010, 2012, 2014 and 2017 respectively.

pension receipt indicates whether the oldest household member reported receiving a state pension - the dependent variable. I construct an instrument using a second dummy variable indicating pension eligibility, defined as the presence of a household member older than the age eligibility threshold. The primary outcome of interest is real household remittances. In the five waves of NIDS, all adults above the age of fifteen are surveyed about their 'contributions' given and received from people who are not normally residents of the households in the past 30 days and 12 months, including cash and in-kind transfers.[7] The NIDS team aggregate the individual-level data to derive and impute household-level remittances received, combining both cash and in-kind transfers (Brophy et al., 2018).

In Table 3.1, I present descriptive evidence for all five waves of NIDS to show that there is high take up of the South African old-age pension and its monetary value is substantial. Across all five waves, approximately a fifth of households in the population has a pension-eligible household member. There is high take up by those households that are age eligible: between 66% and 73% take up the state pension, depending on the year. Moreover, the pension comprises a significant proportion of total household income of recipient households, ranging from approximately 40% to 50%. The pension has increased in nominal value over time, but there is little variation in the monthly amount received within each year, as shown by the pension amount at the 10th, 50th and 90th percentile of its distribution. The data affirms that, conditional on receiving the pension, the means test is only loosely applied and age eligible households typically receive the maximum amount.

Table 3.2 provides descriptive statistics for control households (column 1) in Wave 3 of NIDS in the year 2012, the first survey conducted after a policy change to age eligibility and thus used in my main analysis.[8] Control households are defined as those households where the oldest household member falls just below the age

[7]The survey question is framed as follows: "In the last 12 months, did you receive money, food or any other kind of contribution from people who do not usually sleep under this roof for four nights a week? If you receive maintenance for you or your child, please include it here."

[8]Across all five waves of NIDS, women are eligible for the state pension at 60 years or older, as defined by the South African Social Service Agency. In contrast, the eligibility threshold for men was gradually lowered from 65 to 60 years between 2008 and 2010 (Manuel, 2008).

Table 3.1: Summary statistics

	Wave 1	Wave 2	Wave 3	Wave 4	Wave 5
Pension-eligible households (%)	20.25	21.62	21.02	20.86	19.65
Pension-eligible households receiving a pension (%)	68.55	65.72	69.77	72.64	68.06
Pension as a proportion of household income (%), conditional on pension receipt	47.77	50.10	46.52	40.55	39.07
Pension amount at the 10th percentile (Rand)	870	1000	1200	1300	1500
Median pension amount received (Rand)	870	1080	1200	1350	1590
Pension amount at the 90th percentile (Rand)	940	1080	1200	1410	1600

Notes: Each column presents summary statistics on pension take up and nominal value for Waves 1 to 5 of NIDS. I use post-stratification weights created by Brophy et al. (2018). Wave 1 was conducted in 2008; Wave 2 in 2010-11; Wave 3 in 2012; Wave 4 in 2014-15; and Wave 5 in 2017. Note that the survey waves sometime overlap across multiple financial years, during which the state pension was increased in value. For instance, the state pension amounted to 870 rand at the start of 2008 and was increased to 940 rand after the budget speech in April 2008. Similarly, the maximum pension amount after April 2014 was 1350 rand, but this amount was increased to 1410 rand in April 2015.

eligibility threshold of 60 years and within the bandwidth used in the main specification (ten years below the threshold). Approximately 11% of the control group report receiving remittances, whereas 10% of control households report sending remittances. Conditional on receiving remittances, the average monthly remittances received is R1529 (approximately USD $190) and the average amount sent every month is R1427 (approximately USD $177).

I also present summary statistics for households receiving remittances (column 2) and households sending remittances (column 3) across the entire sample. Households sending remittances are different from households receiving remittances in nearly all respects. Compared to recipient households, households sending remittances tend to be wealthier. Most sending households (94%) have at least one household member who is employed, whereas only half of receiving households report an employed member. Sending households also tend to be more educated, smaller in size with fewer co-resident children, and more likely to be male-headed and located in urban areas. With race as a strong correlate of socioeconomic status in South Africa, sending households are also marginally less likely to be classified as black. Households receiving remittances are more likely to report a migrant.

Table 3.2: Descriptive statistics

	Control mean	Recipients mean	Senders mean	Norm. Diff.	p-value
Household characteristics					
Household size	4.02	3.77	2.55	0.50	0.000
Co-resident children (<=15 years)	1.07	1.18	0.54	0.51	0.000
Household classified as black	0.72	0.87	0.80	0.16	0.007
Urban household	0.68	0.60	0.81	-0.49	0.000
Education of household head	0.49	0.64	0.74	-0.21	0.005
Female household head	0.53	0.61	0.35	0.53	0.000
Age of the oldest household member	54.34	43.74	42.39	0.14	0.213
Income & employment					
Household monthly income	10978.02	7058.99	14665.36	-0.39	0.000
Income, net of pension & remittances	10761.42	5294.63	14450.48	-0.47	0.000
One member is employed	0.76	0.52	0.93	-1.17	0.000
Household has a migrant	0.09	0.14	0.05	0.38	0.000
Remittances sent & received					
Household receives remittances	0.11	1.00	0.11	2.82	
Amount of remittance received	1431.71	1529.34	1102.13	0.27	0.091
Household sends remittances	0.10	0.11	1.00	-4.18	
Amount of remittance sent	196.70	96.17	1427.11	-0.55	0.000
Observations	1718	1133	724		

Notes: Column (1) presents the control mean for households whereby the age of the oldest household member falls between 50 and 59 years, i.e. the bandwidth used in the main specification. Columns (2) and (3) presents the mean for households receiving and sending remittances respectively, across the entire sample. I use post-stratification weights created by Brophy et al. (2018). Norm. Diff. reports the normalised difference between the receiving and sending household means, following (Imbens and Rubin, 2015). The last column reports the p-value on the difference in means between receiving and sending households.

NIDS not only tracks the amount and frequency with which remittances are sent or received, but also captures the relationship between the receiver and sender. As shown in Table 3.9, the primary sender of remittances to control households[9] is most likely to be a family member belonging to a younger generation (43% of the group). The second most likely source of remittances are family members of the same generation (29%), followed by older family members (19%) and non-family households (7%). These figures are consistent with the notion that remittances received by more elderly family members are more likely to be sent by younger generations. Although the data does not allow for further analysis into the relationship between the sender and receiver, they do nevertheless suggest that if a pension crowds out remittances, there may be potential benefits arising from the reallocation of productive capital back to younger populations.

3.4 Empirical strategy

3.4.1 Fuzzy regression discontinuity design

I estimate the effect of receiving a cash transfer on remittances using a fuzzy regression discontinuity design. With imperfect compliance, the take up of the state pension is likely to be endogenous and thus bias the estimates of interest. To circumvent this challenge, my identification strategy exploits the abrupt change in the probability of treatment arising when the oldest household member passes over a fixed age eligibility threshold for the state pension, along otherwise smooth distributions of household characteristics. I employ an instrumental variables strategy that closely resembles that proposed by Hahn et al. (2001) and use the age eligibility threshold as an instrument for pension status.

Using households that fall within a pre-specified neighbourhood of the age eligibility threshold, I begin by estimating the following first-stage equation for household i:

$$pension_i = \gamma_0 + \gamma_1 eligible_i + \gamma_2 age_i + \delta X_i + \eta_i + \epsilon_i \tag{3.1}$$

[9]Control households are again defined as having the oldest household member aged between 50 and 59 years. This is the size of the bandwidth used in the main specification.

where $pension_i$ is an indicator for whether the oldest household member receives a state pension. $eligible_i$ is an indicator for whether the oldest household member is above the age eligibility threshold for the state pension. I refer to households located above the age eligibility threshold as 'treated households' and households below the age eligibility threshold as 'control households'. age_i is the normalised age of the oldest household member, constructed by subtracting the pension age eligibility threshold from age. X_i is a vector of predetermined household characteristics to increase precision in the estimates, when included. I control for a geographic fixed effect η_i: the district using the 2011 census. ϵ_i is a mean zero error term.

Using two-stage least squares (2SLS), I can then estimate the following second-stage equation:

$$Y_i = \beta_0 + \beta_1 pension_i + \beta_2 age_i + \delta X_i + \eta_i + \epsilon_i \tag{3.2}$$

where the outcome variable of interest or Y_i is household real remittances. β_1 can be interpreted as the local average treatment effect of the pension on remittances as an elderly household member crosses the age eligibility threshold, provided the assumption of monotonicity holds. The running variable (age of the oldest household member) takes a discrete form with only a moderate number of distinct values that are shared by many observations – or mass points – within the bandwidths used. As suggested by Kolesár and Rothe (2018), I present heteroskedasticity-robust standard errors for inference.[10]

I first present the results without controlling for pre-determined household characteristics and then show the results controlling for covariates X_i to increase precision in the estimates. The pre-determined covariates include: household size; total number of co-resident children below the age of 15 years; modal race of the

[10]Following Lee and Card (2008), it has been common practice to use confidence intervals based on standard errors that are clustered by the running variable when it is discrete. However, Kolesár and Rothe (2018) find that these confidence intervals do not guard against model misspecification as intended and have poor coverage properties. Clustering by the running variables tend to result in smaller confidence intervals as a result. As a more conservative approach, they instead propose to use heteroskedasticity-robust standard errors for inference and smaller bandwidths to reduce the bias in the estimation of the treatment effect. I show that my results remain robust to even smaller bandwidths in Section 3.6, where I also discuss possible alternatives for future work.

household; education and gender of household head; and urban status. To account for sampling bias in NIDS, I use post-stratification weights created by Brophy et al. (2018).

To implement the local linear regression approach, I use a data-driven bandwidth that minimises the approximate mean squared error (MSE) or the sum of the squared bias and variance of the regression discontinuity estimator, following Imbens and Kalyanaraman (2012). I employ the same MSE-optimal bandwidth in the regression functions on either side of the age eligibility threshold and in both the first and second stage (Lee and Lemieux, 2010). For the main specification, I weight all observations within the bandwidth uniformly and check that the results are robust to alternative weighting approaches in Section 3.6.

3.4.2 Validating identifying assumptions

In this section, I present evidence to show the plausibility of the assumptions underlying the RD design. I show that there is a sharp discontinuity in the probability of receiving a pension at the age eligibility threshold, while the running variable and predetermined covariates remain largely continuous at the threshold. In Section 3.6, I present further robustness checks for the main results. Specifically, to address any concerns about systematic manipulation around the age eligibility threshold, I confirm that my results are not being driven entirely by observations located immediately on either side of the threshold. I further show that the use of placebo age eligibility thresholds cannot explain my results. Lastly, I check robustness of my results to model specification by changing the size of the bandwidth around the age eligibility threshold; including higher order polynomial terms (quadratic terms in particular); using triangular kernel weights; and introducing more mass points by defining the running variable in terms of months, rather than years.[11] To control for any potential endogenous behavioural change in response to the policy reform, I also present results where age, household composition and other household characteristics are determined by the previous survey

[11] I define the age variable in years for the main specification, as this definition produces less measurement error and results in a sharper discontinuity in the probability of pension receipt at the age eligibility threshold.

wave.

Discontinuity in the first stage

Across all five waves of NIDS, women are eligible for the state pension at 60 years or older, as defined by the South African Social Service Agency. In contrast, the eligibility threshold for men was gradually lowered from 65 to 60 years between 2008 and 2010 (Manuel, 2008). Due to the gradual policy change, the discontinuity in the probability of receiving the pension is most evident in 2012, the third wave of NIDS, when the age eligibility threshold is 60 years for both genders. As a result, I use the third wave of NIDS for my main specification and explore the crowding out effect over time in Section 3.7.

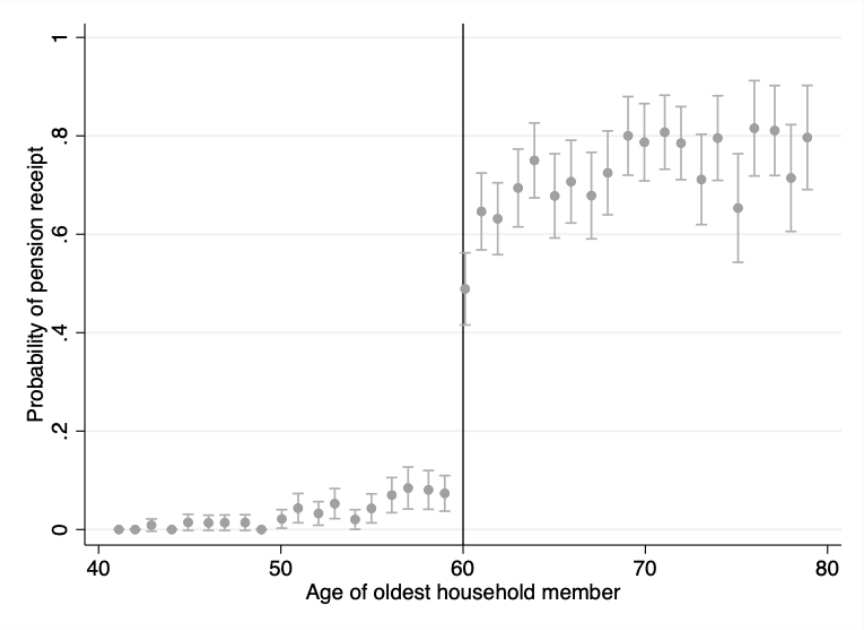
There are strong reasons to believe that using pension age eligibility as an instrument is both relevant and valid. Figure 3.1 illustrates the abrupt jump in the probability of receiving a pension at age 60 for the oldest household members, using Wave 3 of NIDS. There is relatively high take up of the state pension after the age eligibility threshold: whereas only 7.4% of the oldest household members aged 59 years received the pension, 48.9% of individuals aged 60 years received the pension.[12] The probability of receiving a pension increases by less than one at the age eligibility threshold due to imperfect compliance, with not everyone above the threshold receiving the pension and not everyone below the threshold remaining untreated.[13] The imperfect compliance motivates the use of a 'fuzzy' regression discontinuity design.

Moreover, the age eligibility threshold is unlikely to induce other relevant behavioural changes that affect remittances in the local neighbourhood around the threshold, even if recipients essentially age into the programme. Direct beneficiaries are ineligible to receive a pension in conjunction with other state transfers. Receipt of the old-age pension is not dependent on retirement and employment

[12]The proportion of individuals receiving the pension increases to 67.8% by age 65 and 78.7% by age 70.

[13]Imperfect compliance may be due to misreporting of age or incomplete take up of the pension by wealthier households.

Figure 3.1: Probability of pension receipt by age of oldest household member



Notes: Figure 3.1 presents the unconditional predicted probability of receiving a state pension by age of the oldest household member using Wave 3 of NIDS. There is a sharp discontinuity at the pension age eligibility threhsold of 60 years.

is allowed.[14] Appendix Figure 3.9 confirms that, although there is a gradual decline in the employment status of the oldest household members as they grow older, there is no discontinuity in employment status at the pension age eligibility threshold. Figure 3.9 also illustrates that there is no systematic change in migration patterns around the threshold, which is consistent with Jensen (2003) and in contrast to findings by Ardington et al. (2009). I further test that other predetermined household characteristics are continuous at the threshold in Section 3.4.2.

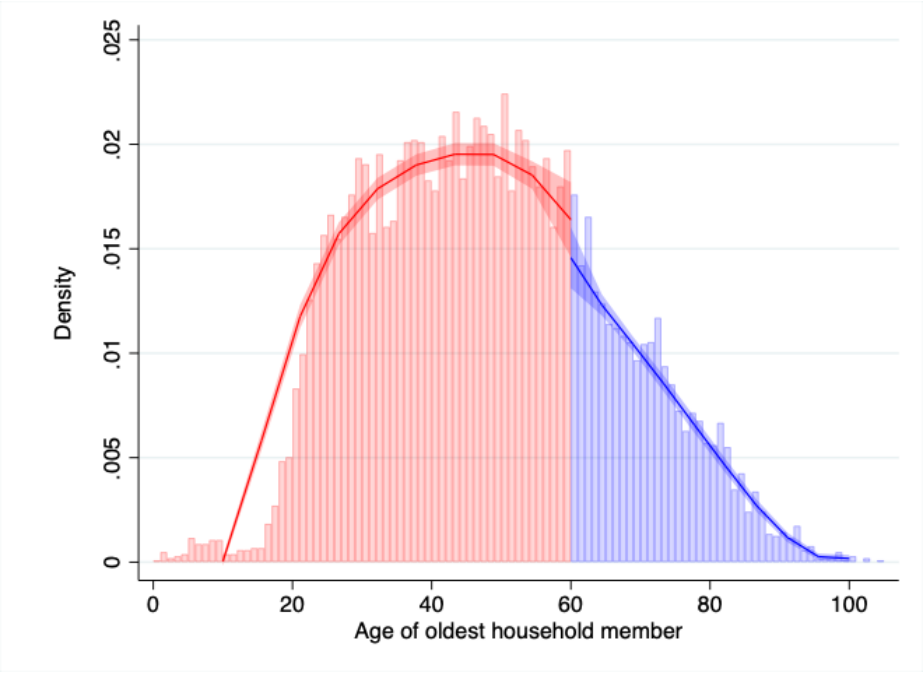
Continuity in the age variable

The discontinuous change in the probability of receiving a pension can only be used to identify the local average treatment effect if households are not able to sort perfectly around the age eligibility threshold. The registration process requires formal documentation and finding loopholes in the bureaucracy is expensive for low-

[14]Recipients remain eligible for the pension as long as the recipient's income does not exceed the means test threshold (OECD, 2017). However, noted in Section 3.3, the means test is unlikely to be a binding constraint.

income households, so it is unlikely that the system allows for systematic manipulation. To affirm this, Figure 3.2 presents a graphical representation of the empirical distribution of the age of the oldest household member and the non-parametric McCrary (2008) density test with overlapping 95% confidence intervals for Wave 3 of NIDS (Cattaneo et al., 2019). The graph suggests that there may be a small drop in the density at the age eligibility threshold. However, when I use a formal test, I fail to reject the null hypothesis of no difference in the density of treated and control observations at the cutoff, with a statistic of -1.2258 ($p = 0.2203$). Thus, the observed drop in the age variable at the age eligibility threshold is not sufficiently large to be statistically significantly different from zero. There is no concerning discontinuity in the running variable at the threshold.

Figure 3.2: Density of age of oldest household member



Notes: Figure 3.2 provides a graphical representation of the empirical distribution of the age of the oldest household member and the non-parametric McCrary (2008) density test with overlapping 95% confidence intervals for Wave 3 of NIDS. The pension age eligibility threshold is 60 years.

Continuity in predetermined household covariates

To test whether control households just below the age eligibility threshold are a valid counterfactual to treated households, I check that predetermined household

characteristics are continuous at the age eligibility threshold. In the absence of the treatment, covariates determined before treatment assignment should be similar across the two groups.

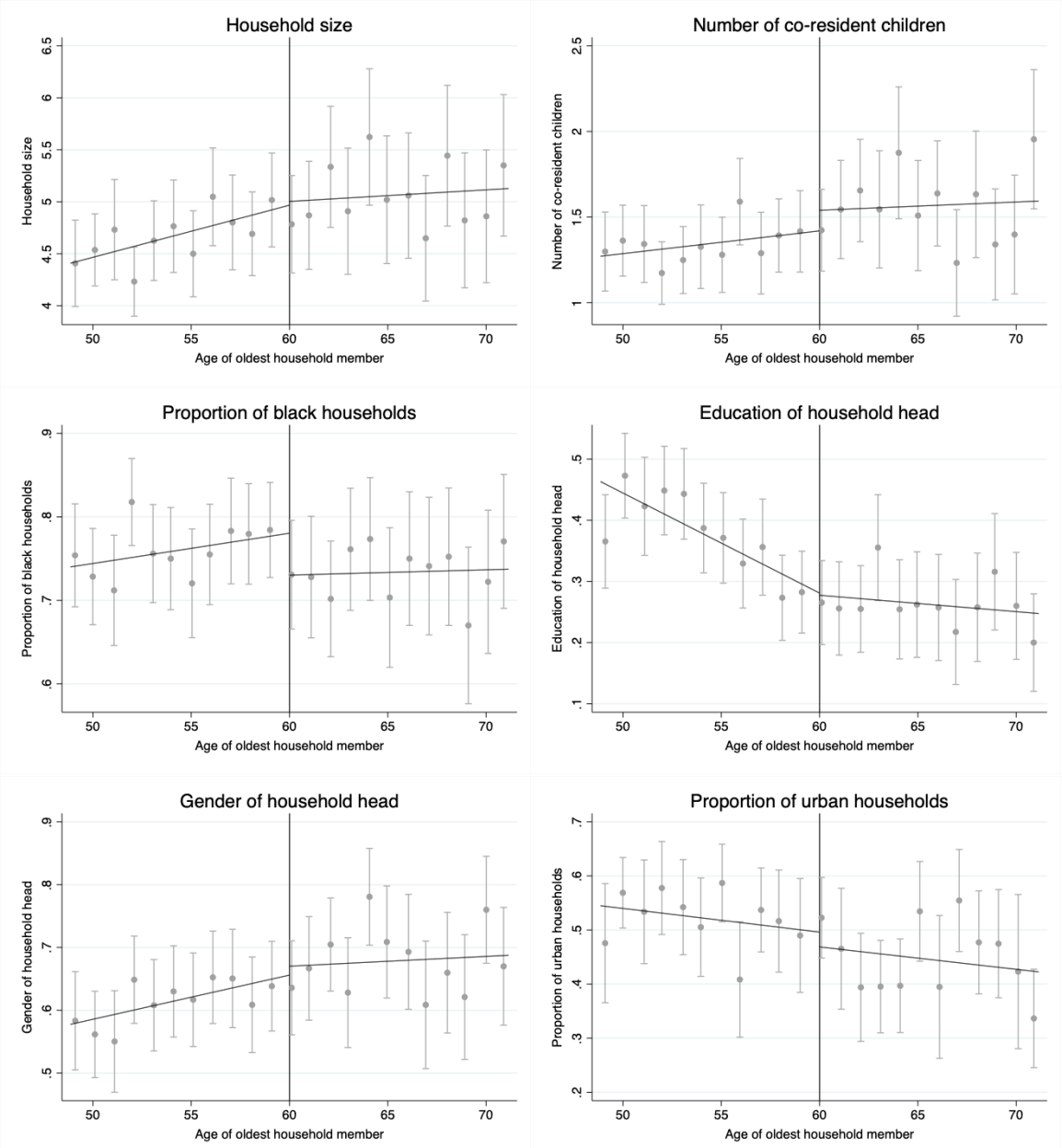
Figure 3.3 illustrates the continuity in observable household characteristics at the age eligibility threshold for Wave 3 (main specification), by plotting the unconditional expectation functions against the age variable on either side of the age eligibility threshold. There is no noticeable discontinuity in household characteristics at the age eligibility threshold, with the exception of the number of co-resident children and race variable. For the latter two variables, there is an observed change at the pension age eligibility threshold, moving in the direction of potentially endogenous effects. Thus, I formally test whether there is a discontinuity by estimating a local linear regression on either side of the threshold for each covariate. I report the results in Appendix Table 3.10. For all variables (including the race and co-resident children variables), I fail to reject the null hypothesis that there is no discontinuous effect at the threshold. Therefore, while there is some suggestive evidence that households may be responding to the pension, these effects are not sufficiently large to conclude that these characteristics are discontinuous at the age eligibility threshold.

Moreover, I find similar patterns for the race variable when I plot the unconditional expectation functions against the age variable for the first wave of NIDS, the year in which the gradual policy change to lower the eligibility age of men from 65 years was first initiated (Appendix Figure 3.10). The drop after age 60 years is thus unlikely to be driven by either the pension or differential attrition in survey response over time, but rather reflect broader structural reasons, such as high mortality rates.[15]

Finally, in Section 3.5, I control for and exclude these covariates in my main specification and find that my results do not differ significantly.

[15]The overall life expectancy in South Africa was only 59.6 years in 2013, due to the high incidence of HIV/AIDS and tuberculosis (Statistics South Africa, 2013b). Black South Africans are disproportionately affected by these two epidemics.

Figure 3.3: Continuity test of predetermined household characteristics at the age eligibility threshold



Notes: Each panel shows the unconditional mean and expectation function of the predetermined household characteristic against the age variable on either side of the age eligibility threshold. 95% confidence intervals are shown. The covariates determined before treatment assignment are defined as: household size; total number of co-resident children below the age of 15; a dummy variable for a household with predominantly black household members; education of the household head; a dummy variable for a female household head; and a dummy variable for an urban household. For all specifications, I define the bandwidth size to be the same as my preferred main specification, i.e. 10.356 years on either side of the threshold.

3.5 Results

3.5.1 Main specification

The primary question of interest is whether receiving the state pension crowds out remittances. I evaluate the local average treatment effect of the state pension on household real remittances using a local linear regression approach. Table 3.3 presents the first stage and RD treatment effects using MSE-optimal bandwidths, controlling for predetermined household covariates (columns 1 and 2) and excluding these covariates (columns 3 and 4). In both specifications, there is a strong first stage. Controlling for predetermined covariates, crossing the age eligibility threshold predicts a 55.8% take up of the state pension ($p = 0.000$). Excluding covariates, pension eligibility predicts 48.8% of pension take up ($p = 0.000$).

I find that receiving the state pension crowds out 266.4 rand (approximately USD $33) of monthly household remittances on average when the oldest household member becomes age eligible, controlling for predetermined covariates. The effect is statistically significant at the 5% level ($p = 0.013$). Excluding these covariates does not substantively change the point estimate (-221.1 rand), albeit estimated with less precision ($p = 0.038$). Control households receive 214 rand of monthly remittances on average[16] and nearly all households received the maximum monthly pension amount of 1,200 rand in 2012. Using my main result, this suggests that a 1 rand increase in income received from the pension crowds out 0.22 rand in remittance income on average.

Figure 3.4 graphically illustrates the mean predicted value of remittances by age of the oldest household member, controlling for household characteristics. The mean predicted value of household remittances increases as the oldest household member grows older between 50 and 60 years. At the pension age eligibility threshold (60 years), the average level of household remittances received drops substantially and remains persistently lower, even for pension-eligible households

[16]Control households also include non-compliers who would not necessarily take up the pension once the oldest household member becomes age eligible. In contrast, the RD estimate is the local average treatment effect for compliers that are age eligible and take up the state pension within the local neighbourhood of the age eligibility threshold.

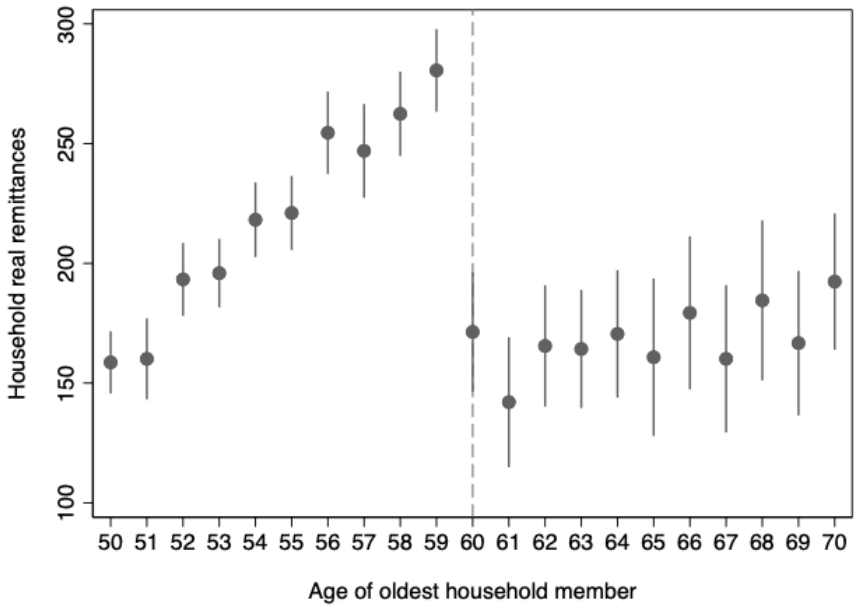
Table 3.3: RD effect of the state pension on household remittances

	(1) Pension received	(2) Remittances	(3) Pension received	(4) Remittances
Pension eligibility (>=60 years)	0.558***		0.488***	
	(0.0259)		(0.0240)	
Pension received		-266.4**		-221.1**
		(107.1)		(106.7)
Covariates	✓	✓		
Fixed effect	✓	✓		
N	2,964	2,964	3,702	3,702
MSE-optimal bandwidth	10.36	10.36	11.14	11.14
F-stat for weak IV test	464.32	464.32	412.00	412.00
AR confidence intervals	[-474.17,-58.68]	[-474.17,-58.68]	[-428.11,-14.13]	[-428.11,-14.13]

Notes: Columns (1) and (2) report the first and second stage respectively for an instrumental variables approach using local linear regressions and MSE-optimal bandwidths, controlling for predetermined household characteristics to increase the precision of estimates. Covariates include: household size; total number of co-resident children; modal race of the household; education and gender of household head; and urban status. I control for district as the geographic fixed effect. Columns (3) and (4) report the first and second stage respectively excluding these covariates. I select a data-driven bandwidth that minimises the approximate mean squared error of the RD estimator, following Imbens and Kalyanaraman (2012). Robust standard errors are reported in parenthesis. I also report the F-statistic for the weak IV test proposed by Olea and Pflueger (2013) and Anderson-Rubin confidence intervals. I use post-stratification weights created by Brophy et al. (2018). The remittances variable is winsorised at the 1st and 99th percentile to account for outliers. Monthly household remittances are deflated using the monthly consumer price index, with December 2016 as the base year. * $p<0.10$, ** $p<0.05$, *** $p<0.01$

located further away from the threshold.[17]

Figure 3.4: Mean predicted value of remittances, by age of the oldest household member



Notes: This figure plots the mean predicted value of remittances by age of the oldest household member, controlling for household characteristics.

3.6 Robustness to model specification and placebo tests

My results are robust to a wide range of alternative model specifications and placebo tests. I first show that the results are robust to bandwidths of different sizes and other model specifications, including the inclusion of higher order polynomial terms and triangular kernel weights. To control for any potential endogenous behavioural change in response to the policy reform, I also present results where age, household composition and other household characteristics are determined by the previous survey wave. I then present evidence to suggest that systematic manipulation is likely to be limited around the age eligibility threshold, by checking the sensitivity of my estimates to the exclusion of observations in the immediate vicin-

[17]Given that the relationship between age and household remittances appears non-linear, I show that the crowding out effect remains robust to the inclusion of the square of age in Section 3.6.

ity of the threshold. Lastly, I check that placebo age eligibility thresholds cannot explain the results, offering suggestive evidence in support of the continuity assumptions underlying the RD design.

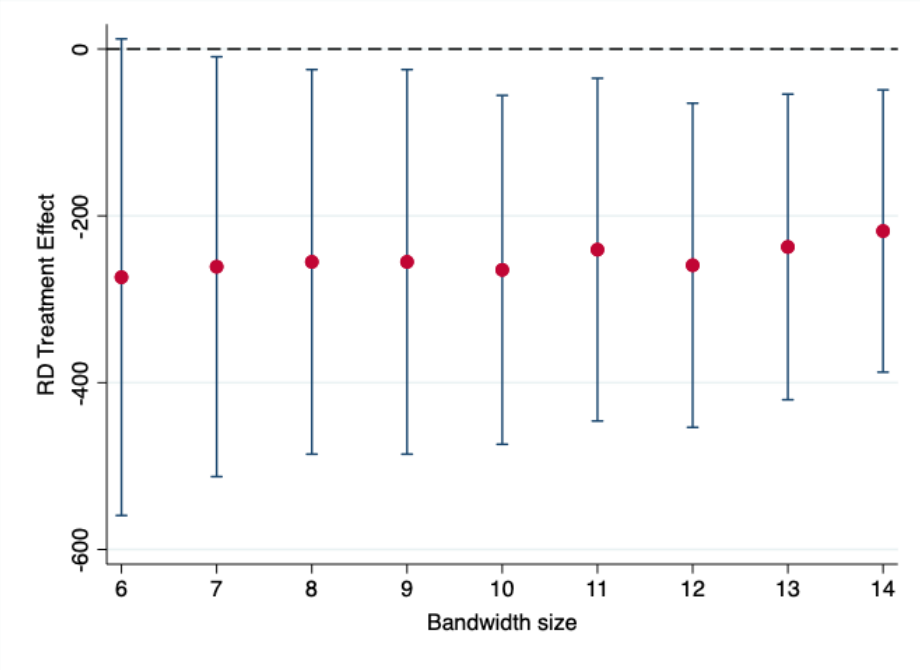
The choice of bandwidth determines the number of observations in the local neighbourhood around the age eligibility threshold. While a smaller bandwidth reduces the bias in the estimation of the treatment effect, it also increases the variance in the RD estimator. I test the sensitivity of results to bandwidth choice, while holding constant other estimation and inference procedures.[18] Figure 3.5 provides a graphical representation of the crowding out effect with 95% confidence intervals for bandwidths that vary in size from 6 to 14 years on either side of the age eligibility threshold. Further details are provided in Appendix Table 3.13. As expected, the variance in the RD estimator decreases with larger bandwidths. However, the magnitude of the coefficients across all specifications remain remarkably consistent with the main point estimate obtained using the MSE-optimal bandwidth of 10.36. Moreover, the treatment effect for all bandwidth choices are statistically significant at the 5% level, with the exception of the smallest bandwidth choice of 6 where the effect is statistically significant at the 10% level ($p = 0.059$).

Table 3.4 shows that my results are robust to other model specifications. My preferred model specification (shown in column 1) uses linear local regressions, age defined in years as the running variable, rectangular kernel weights and the MSE-optimal bandwidth for both estimation and inference. However, my results are robust to the inclusion of quadratic terms (column 2); defining age in terms of months (column 3); using triangular kernel weights (column 4) and using the bias correction proposed by Calonico et al. (2014) (column 5). While local linear regressions has been shown to be an effective low-order polynomial approximation (Hahn et al., 2001), it is typically best practice to show robustness to quadratic polynomials.[19] I introduce more mass points by defining the age variable in terms

[18] Kolesár and Rothe (2018) make a case for using heteroskedasticity-robust standard errors for inference and making the bandwidth smaller to reduce the bias of the treatment effect estimate. They also propose the use of "honest confidence intervals" to address the bias explicitly and serves as a useful robustness check in future work.

[19] The use of higher order polynomials are discouraged, leading to noisy estimates, poor coverage of confidence intervals and sensitivity to the degree of the polynomial (Gelman and Imbens, 2019). Note that the F-stat remains large (314) in the quadratic case.

Figure 3.5: RD effect of the pension on remittances for different bandwidth choices



Notes: The markers indicate the local average treatment effect of receiving the state pension on real household remittances obtained from an instrumental variables approach using local linear regressions and controlling for predetermined household characteristics. 95% confidence intervals are shown, using robust standard errors. I manually vary the size of the bandwidth to test the sensitivity of the point estimates to bandwidth choice. The data-driven bandwidth that minimises the approximate mean squared error of the RD estimator is 10.36 years on either side of the age eligibility threshold.

of months, rather than years, albeit introducing more noise.[20] The triangular kernel function places greater weight on observations closer to the age eligibility threshold. The bias correction addresses the concern that MSE-optimal bandwidths tend to be large in size and thus introduce bias in the distributional approximation of the RD estimators.[21] For all model specifications, the point estimates remain relatively stable, if not larger in magnitude, and statistically significant at the 5% level.

To control for any potential endogenous behavioural change in response to the policy reform, Table 3.5 shows results for Wave 3 where age and thus eligibility of the oldest household member is determined by household composition in Wave 2. Columns (1) and (2) present results controlling for household characteristics as determined in Wave 2, whereas columns (3) and (4) present results without controls. Using these controls, I still find a strong crowding out effect of R126 and R130 rand respectively, although I can no longer reject the null hypothesis that these results are statistically significantly different from zero. The noisier estimates may partially reflect that household characteristics change rapidly in a South African context, for reasons not necessarily related to the pension. As a result, the NIDS surveys follow the individual over time, and not the household.

I investigate the sensitivity of the results to observations closest to the age eligibility threshold. The “donut hole" test assumes that those households closest to the threshold are most likely to engage in systematic manipulation if it were to occur. The test omits these units to assess the sensitivity of results to these omitted units and the extrapolation that occurs at the threshold using a local linear regression approach. In Table 3.6 and Appendix Figure 3.11, I omit households with the oldest member aged within one and two years of either side of the age eligibility threshold and proceed with the usual estimation and inference procedures using

[20] I define the age variable in years for the main specification, as this definition produces less measurement error and results in a sharper discontinuity in the probability of pension receipt at the age eligibility threshold. In contrast, the variables used in constructing age in months suffered from many missing observations, with artificial pile ups at the year mark as a result. Mismeasurement error is likely to be correlated with the income status of respondents. Nevertheless, using age defined in months serves as a useful robustness check.

[21] The bias correction first uses an estimate of the leading bias to recenter the usual t-statistic in order to account for the “large" bandwidth choice. The t-statistic is then re-scaled with a standard error formula that accounts for the additional variability introduced by the estimated bias (Calonico et al., 2014).

Table 3.4: Robustness to alternative model specifications

	(1) Main specification	(2) Quadratic Function	(3) Age in months	(4) Triangular kernel weights	(5) Bias correction
	Panel A: First stage				
	Pension received	Pension received	Pension received	Pension received	Pension received
Pension eligibility	0.558***	0.526***	0.518***	0.480***	0.467***
	(0.0259)	(0.0297)	(0.0328)	(0.0438)	(0.0508)
	Panel B: Second stage				
	Remitt	Remitt	Remitt	Remitt	Remitt
Pension received	-266.4**	-326.2**	-250.7**	-357.45**	-379.71**
	(107.1)	(132.0)	(126.2)	(177.7)	(165.25)
Covariates	✓	✓	✓	✓	✓
Fixed effect	✓	✓	✓	✓	✓
Observations	2,964	4,691	2,191	2,209	2,209
MSE-optimal bandwidth	10.36	17.48	7.81	13.91	13.91
Order of polynomial	1	2	1	1	1
Definition of age	Year	Year	Months	Year	Year
Kernel weights	Rectangular	Rectangular	Rectangular	Triangular	Triangular
Bias-correction					✓

Notes: Panel A and B report the first and second stage respectively for an instrumental variables model using local linear regressions and controlling for predetermined household characteristics to increase the precision of estimates. Column (1) presents the main specification. Column (2) presents results including the quadratic term for the running variable. Column (3) defines age in terms of months, rather than years. Column (4) employs triangular kernel weights in the local linear regression, thereby placing greater weight on households closer to the age eligibility threshold. Column (5) uses triangular kernel weights and the bias correction proposed by Calonico et al. (2014). Covariates include: household size; total number of co-resident children; modal race of the household; education and gender of household head; and urban status. I control for district as the geographic fixed effect. * $p<0.10$, ** $p<0.05$, *** $p<0.01$

Table 3.5: Robustness to controls from the previous wave

	(1) Pension received	(2) Remittances	(3) Pension received	(4) Remittances
Pension eligibility (>=60 years)	0.402***		0.368***	
	(0.0252)		(0.0243)	
Pension received		-126.2		-130.0
		(145.6)		(146.2)
Covariates	✓	✓		
Fixed effect	✓	✓		
N	2,966	2,966	3,498	3,498
MSE-optimal bandwidth	9.20	9.20	10.04	10.04
F-stat for weak IV test	254.86	254.86	229.19	229.19

Notes: Columns (1) and (2) report the first and second stage respectively for an instrumental variables approach using local linear regressions and MSE-optimal bandwidths, controlling for pre-determined household characteristics from Wave 2. The age and thus eligibility of the oldest household member is also determined by the household composition of Wave 2. Covariates include: household size; total number of co-resident children; modal race of the household; education and gender of household head; and urban status. I control for district as the geographic fixed effect. Columns (3) and (4) report the first and second stage respectively, but excluding these covariates. * $p<0.10$, ** $p<0.05$, *** $p<0.01$

the remaining sample. My conclusions are robust to the donut hole test: excluding both the first and second window leads to similar point estimates as in the main specification, but estimated with less precision. Both results are weakly statistically significant at the 10% level.

Lastly, I examine whether there are treatment effects at placebo thresholds at which treatment assignment should not change. In the absence of treatment, the regression functions for households above and below the age eligibility threshold should be continuous at the threshold for the RD design to remain valid. Therefore, the presence of empirical discontinuities in the estimable regression functions for control and treated households away from the age eligibility threshold should cast doubt on the RD design (Cattaneo et al., 2019). To indirectly test this assumption, I perform the same estimation and inference procedures as for the main specification (including holding constant bandwidth size), but I now replace the true age eligibility threshold with four placebo thresholds at age 55, 57, 63 and 65. As shown in Figure 3.6, there are no large jumps in the probability of receiving a pension at these placebo thresholds. Moreover, I observe no significant changes in household

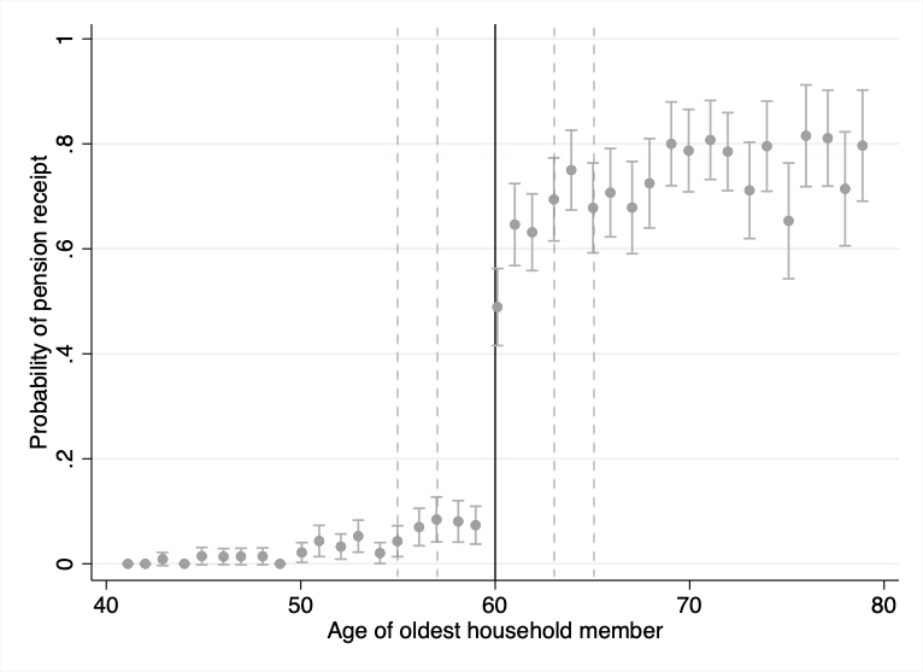
Table 3.6: Donut hole test

	No donut hole		Donut radius of 1		Donut radius of 2	
	(1) Pension received	(2) Remittances	(3) Pension received	(4) Remittances	(5) Pension received	(6) Remittances
Pension eligibility	0.558***		0.621***		0.621***	
	(0.0259)		(0.0290)		(0.0353)	
Pension received		-266.4**		-207.6*		-271.4*
		(107.1)		(114.1)		(147.2)
Covariates	✓	✓	✓	✓	✓	✓
Fixed effect	✓	✓	✓	✓	✓	✓
Observations	2,964	2,964	2,802	2,802	2,496	2,496
Size of donut	0	0	1	1	2	2
Bandwidth size	10.36	10.36	10.36	10.36	10.36	10.36
F-stat for weak IV test	464.32	464.32	458.24	458.24	309.51	309.51
AR confidence intervals	[-474,-59]	[-474,-59]	[-429,14]	[-429,14]	[-557,14]	[-557,14]

Notes: Each column reports the second-stage RD treatment effect of receiving the state pension on real household remittances, using local linear regressions and controlling for predetermined household characteristics to increase the precision of estimates. I present results for the main specification (no donut hole), omitting the first window and omitting the second window around the age eligibility threshold. The same bandwidth size – 10.36 years on either side of the threshold – is used for all specifications. Covariates include: household size; total number of co-resident children; modal race of the household; education and gender of household head; and urban status. I control for district as the geographic fixed effect. Robust standard errors are reported in parenthesis. I also report the F-statistic for the weak IV test proposed by Olea and Pflueger (2013) and Anderson-Rubin confidence intervals. I use post-stratification weights created by Brophy et al. (2018). The remittances variable is winsorised at the 1st and 99th percentile to account for outliers. Monthly household remittances are deflated using the monthly consumer price index, with December 2016 as the base year. * $p<0.10$, ** $p<0.05$, *** $p<0.01$

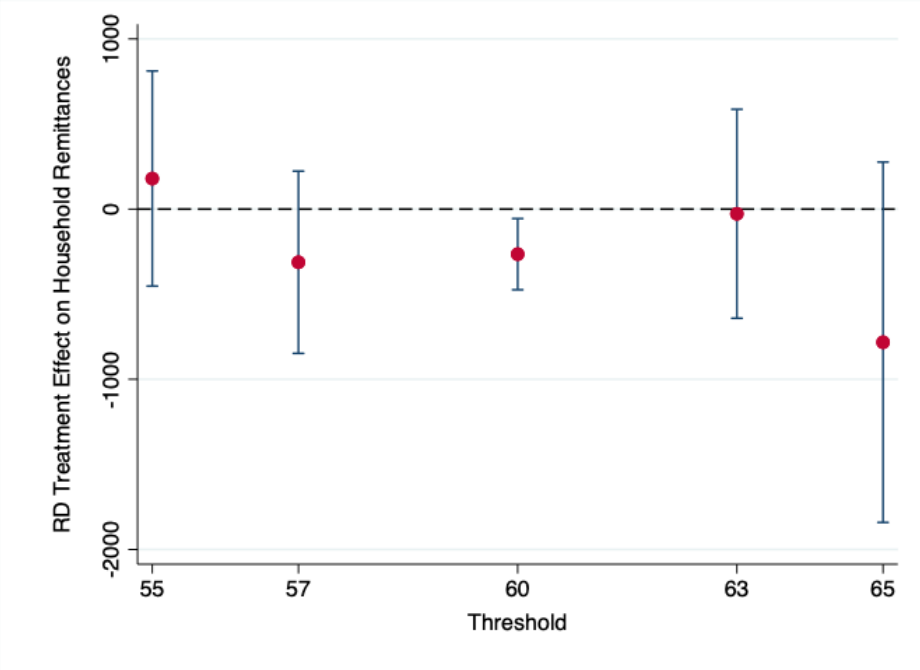
remittances at the placebo thresholds, as shown in Figure 3.7. The average level of real household remittances only decreases significantly at the true age eligibility threshold of 60 years. Further details are provided in Appendix Table 3.14.

Figure 3.6: Probability of receiving a pension, with placebo thresholds



Notes: The figure shows the probability of receiving a pension by age of the oldest household member. In addition to the true age eligibility threshold (60 years), I indicate four placebo thresholds to show that there are no large discontinuities away from the true threshold. The four placebo thresholds include age 55, 57, 63 and 65.

Figure 3.7: RD effect of the pension on remittances for placebo thresholds



Notes: The markers indicate the local average treatment effect of receiving the state pension on real household remittances obtained from an instrumental variables approach using local linear regressions and controlling for predetermined household characteristics. 95% confidence intervals are shown, using robust standard errors. In addition to the true age eligibility threshold, I use four placebo thresholds to test whether there are any discontinuities away from the true threshold. I use the same bandwidth size - 10.36 years on either side of the threshold - for all specifications.

3.7 Crowding out over time

In this section, I explore the crowding out effect over time using three waves of NIDS.[22] Table 3.7 presents the first stage and RD treatment effects for survey Wave 1 (2008), Wave 3 (2012) and Wave 5 (2017), using MSE-optimal bandwidths and controlling for household covariates. Wave 1 was conducted just prior to and during the start of the policy change, whereby the age eligibility threshold for men was gradually reduced from 65 to 60 years between 2008 and 2010. I use Wave 3 in my main specification above, as the first survey conducted immediately after the policy change. Wave 5 was conducted approximately seven years after the policy change. I use the official age eligibility thresholds for each year.[23]

I find that the crowding out effect is only observed in the year 2012. Columns 1 and 2 in Table 3.7 show a strong crowding out effect of 266.4 rand ($p < 0.05$) for the main specification using Wave 3, the period shortly after the policy change. For the year in which the policy change was first introduced (Wave 1), I observe a crowding out effect of 130.1 rand (column 4). However, the effect for this year is imprecisely estimated and I cannot reject the hypothesis that the effect is different from zero ($p = 0.404$).[24] By Wave 5 in 2017, seven years after the policy change, I find that receiving a state pension has no effect ($p = 0.805$) on household remittances in the local neighbourhood of the age eligibility threshold of 60 years (columns 5 and 6).

Figure 3.8 plots the average level of real household remittances by age of the oldest household member within ten years of either side of the age eligibility threshold for the three waves of NIDS. The discontinuous fall in average remittances received at the age eligibility threshold of 60 years is clearly illustrated for Wave 3. The drop is less precisely shown in Wave 1 due to the slow roll out of the

[22] I examine the regression discontinuity effect within each survey wave, effectively treating each wave as a cross-section. I am unable to exploit the panel dimension of NIDS to explore the effects over time, as control households eventually age into receiving the pension.

[23] The official age eligibility threshold was 60 years for females in 2008 and for both genders in Waves 3 and 5; and 65 years before April 2008 and 63 years after April 2008 for men. By the end of the 2010 budget year, the age eligibility for men stabilised at age 60.

[24] I find a similarly sized crowding out effect of 144.51 rand in Wave 4 (2014), although this effect is also imprecisely estimated with $p = 0.238$. I find no crowding out effect in Wave 2, the survey wave conducted during the policy change.

Table 3.7: RD effect of the state pension on household remittances, over time

	Main specification Wave 3 (2012)		Pre-policy change Wave 1 (2008)		Post-policy change Wave 5 (2017)	
	(1) Pension received	(2) Remittances	(3) Pension received	(4) Remittances	(5) Pension received	(6) Remittances
Pension eligibility	0.558***		0.436***		0.517***	
	(0.0259)		(0.0385)		(0.0223)	
Pension received		-266.4**		-130.1		28.28
		(107.1)		(155.8)		(114.8)
Covariates	✓	✓	✓	✓	✓	✓
Fixed effect	✓	✓	✓	✓	✓	✓
N	2,964	2,964	1,665	1,665	3,648	3,648
MSE-optimal bandwidth	10.36	10.36	7.26	7.26	9.04	9.04
F-stat for weak IV test	464.32	464.32	128.14	128.14	538.75	538.75
AR confidence intervals	[-474.1,-58.7]	[-474.1,-58.7]	[-432.2,172.1]	[-432.2,172.1]	[-194.4,251.0]	[-194.4,251.0]

Notes: Columns (1) and (2) present the first and second stage for an instrumental variables approach using local linear regressions and MSE-optimal bandwidths for Wave 3 (2012), controlling for predetermined household characteristics to increase the precision of estimates. This is the main specification used in the core results. In Wave 3, data were collected shortly after the end of the policy change in age eligibility for men. Columns (3) and (4) present the same results for Wave 1 (2008), whereby data was collected before and at the start of the gradual policy change. Columns (5) and (6) present the same results for Wave 5 (2017), more than half a decade after the end of the policy change. Covariates include: household size; total number of co-resident children; modal race of the household; education and gender of household head; and urban status. I control for district as the geographic fixed effect. I select a data-driven bandwidth that minimises the approximate mean squared error of the RD estimator, following Imbens and Kalyanaraman (2012). Robust standard errors are reported in parenthesis. I also report the F-statistic for the weak IV test proposed by Olea and Pflueger (2013) and Anderson-Rubin confidence intervals. I use post-stratification weights created by Brophy et al. (2018). The remittances variable is winsorised at the 1st and 99th percentile to account for outliers. Monthly household remittances are deflated using the monthly consumer price index, with December 2016 as the base year. * $p<0.10$, ** $p<0.05$, *** $p<0.01$

policy reform. In contrast, there is no clear jump in Wave 5 at the age eligibility threshold. In fact, the average level of real household remittances has increased for control households below the threshold from R213 in 2012 to R252 in 2017. The proportion of control households receiving remittances also rose from 11% in 2012 to almost 17% in 2017. Similarly, the proportion of households sending remittances in the sample as a whole doubled from 10% to 20% between 2012 and 2017.

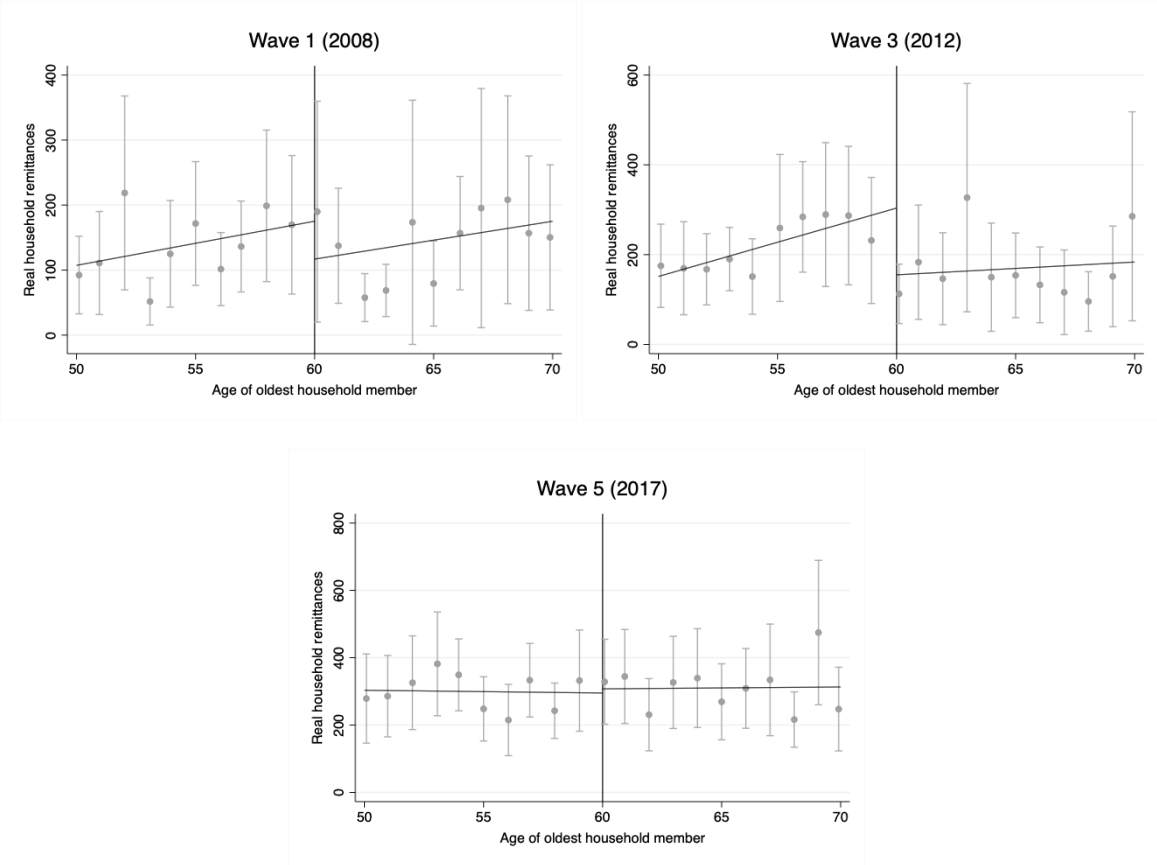
These results suggest that the expansion of public social protection does not appear to be unravelling private support networks over time. In fact, remittances appear to be increasing in size and coverage, despite no change in pension take up.[25] While it is not possible to pin down the exact reason for these trends, it is evident that the profile of the sending households has changed over time. Table 3.8 reports the mean on household characteristics, income and remittances for control households, whereby the oldest household member falls between ages 50 and 59 (the main bandwidth used). Households sending remittances in 2017 compared to 2012 are indeed wealthier in terms of real monthly income. Sending households are also more likely to be classified as black in 2017, potentially reflecting the growing black middle class in post-Apartheid South Africa. In contrast, households receiving remittances appear qualitatively poorer in 2017 relative to 2012, but there are no statistically significant differences across most variables, except for a greater presence of female-headed households in 2017. Nevertheless, these findings provide suggestive evidence to explain why average remittance levels may have been higher in 2017 and crowding out was not detected.

3.8 Conclusion

In this paper, I find evidence of significant crowding out of remittances in response to a public cash transfer. Every 1 rand of pension income received leads to 0.22 rand decrease in remittance income on average for age-eligible households that take up the pension. The result is robust to a wide range of alternative model specifications.

[25] In 2012, 54% of households took up the pension at the age eligibility threshold of 60 years, whereas 55% of households took up the pension in 2017. Similarly, 73% of households took up the pension within the ten-year window after the age eligibility threshold in both waves.

Figure 3.8: Real household remittances, by age of the oldest household member



Notes: The three figures illustrate the average level of real household remittances by the age of the oldest household member for three waves of NIDS, controlling for household characteristics and district fixed effect.

Table 3.8: Descriptive statistics for households that receive or send remittances in Wave 3 vs Wave 5

	Wave 3 recipient mean	Wave 5 recipient mean	*p*-value	Wave 3 sender mean	Wave 5 sender mean	*p*-value
Household characteristics						
Household size	4.87	4.79	0.731	2.55	2.43	0.174
Co-resident children (<=15 years)	1.38	1.72	0.247	0.54	0.54	0.965
Household classified as black	0.87	0.89	0.910	0.80	0.90	0.000
Urban household	0.52	0.49	0.455	0.81	0.74	0.000
Education of household head	0.46	0.48	0.694	0.74	0.83	0.170
Female household head	0.63	0.72	0.058	0.35	0.33	0.965
Age of household head	51.16	51.63	0.951	40.65	39.67	0.292
Income & employment						
Real monthly income	13,702	8,601.20	0.1899	14,665.36	14,937.96	0.039
Income, net of pension & remittances	11,723.18	6,989.58	0.2184	14,450.48	14,614.58	0.036
One member is employed	0.69	0.63	0.625	0.93	0.92	0.321
Household has a migrant	0.21	0.14	0.126	0.05	0.04	0.692
Remittances sent & received						
Amount of remittance received (conditional)	1,908.93	1,489.53	0.269	162.89	202.62	0.126
Amount of remittance sent	15.78	174.17	0.005	1,532.39	1591.87	0.748
Observations	249	466		725	2,080	

Notes: Columns (1) and (2) report the control mean for households receiving remittances in Wave 3 and 5 respectively. A control household is defined as having the oldest household member aged between 50 and 59 years, i.e. the bandwidth used in the main specification. The third column reports the *p*-value on the difference in means between receiving households in 2012 and 2015 respectively. Columns (4) and (5) report the control mean for households sending remittances, within the main bandwidth. The last column reports the *p*-value on the difference in means.

This paper contributes to a growing literature that find substantial crowding out effects in low- and middle-income countries. In most of these contexts, as in my own, social protection had rapidly expanded in size or coverage in the years prior to the analysis. Yet I show that the crowding effect is only observed in the year 2012 and not in later years. In fact, remittances grow in both size and coverage over the years across the board. Although I cannot pin down the precise reason for this increase, I find consistent descriptive evidence that sending households are becoming wealthier over time and receiving households appear qualitatively poorer over time. The growing social protection system does not appear to be unravelling private support networks, but further research is required to verify this hypothesis.

These results have important implications for the efficacy of public cash transfers, as the pension is in effect shared across many household members. The welfare effect of the pension for direct recipients is likely to be partly neutralised by the endogenous response in remittances. However, if younger populations are indeed the primary source of remittances, the crowding out thereof in response to social protection may relieve the burden for working-age populations to support the elderly and instead free up capital for more productive uses. The broader welfare effects of the crowding out remains an open question for further investigation.

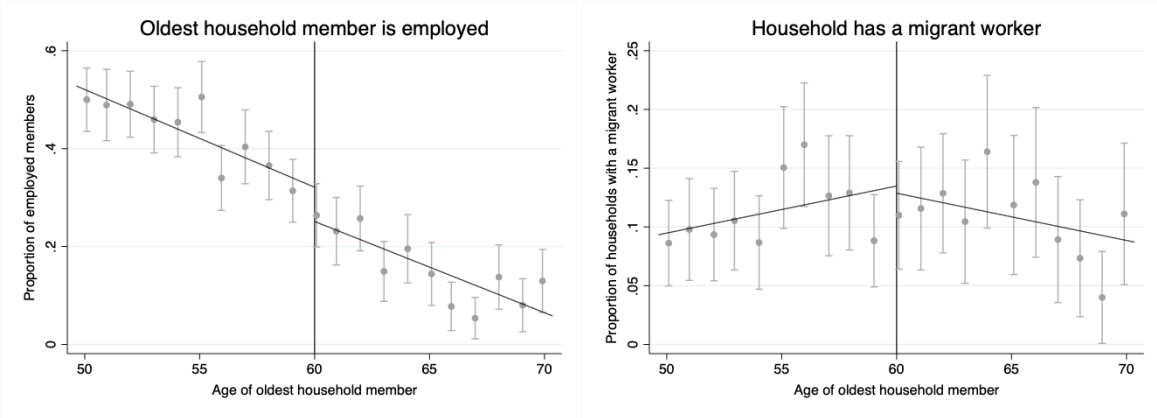
Appendices of Chapter 3

Table 3.9: Relationship of the sender of remittances for control households

Relationship of the sender	Percentage of the control group
Generation below (family)	43%
Same generation (family)	29%
Generation above (family)	19%
Non-family	7%
Other family members	2%

Notes: The table shows the relationship of the primary sender of household remittances, as reported by recipients in control households, whereby the age of the oldest household member falls between 50 and 59 years, i.e. the bandwidth used in the main specification

Figure 3.9: Employment and migration patterns around the age eligibility threshold for Wave 3



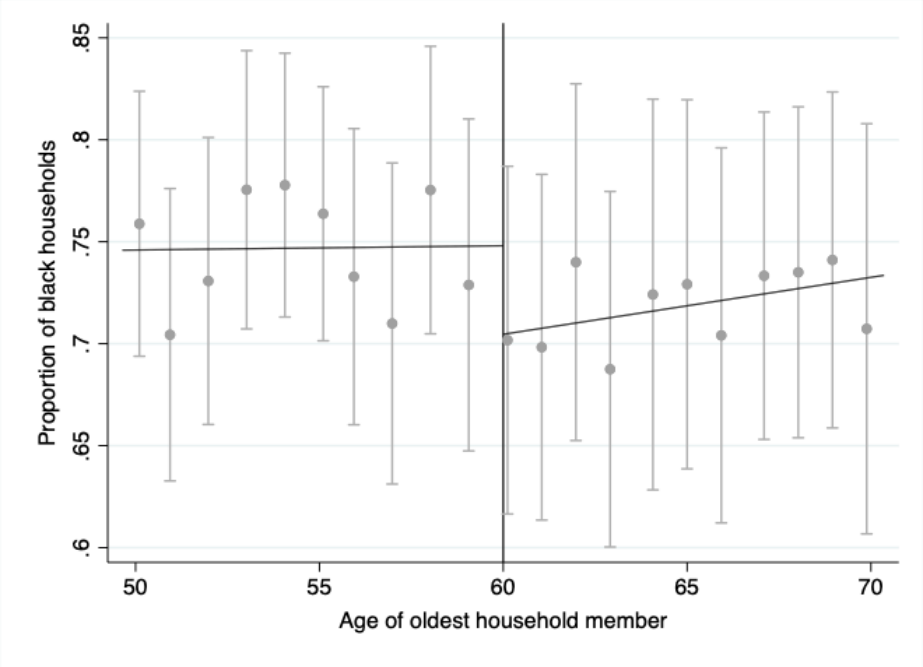
Notes: Each panel shows the unconditional mean and expectation function of the the specified variable against the age variable on either side of the pension age eligibility threshold of 60 years. The left panel uses a dummy variable for whether the oldest household member is employed, whereas the right panel uses a dummy for whether the household reports a migrant worker as a household member. 95% confidence intervals are shown.

Table 3.10: RD effect on predetermined covariates at the age eligibility threshold

	(1) Household size	(2) Number of Children	(3) Proportion of Black Households	(4) Education of Household Head	(5) Female Household Head	(6) Proportion of Urban Households
Pension received	0.071	0.166	-0.049	0.012	0.011	-0.023
	(0.321)	(0.169)	(0.043)	(0.049)	(0.052)	(0.068)
Observations	2,976	3,402	3,402	2,964	2,964	3,309

Notes: To show that predetermined covariates do not differ systematic at the age eligibility threshold of 60 years, each column reports the sharp RD effect of predetermined household characteristics at the threshold, using a local linear regression approach. The predetermined household characteristics include: household size (column 1); total number of co-resident children below the age of 15 (column 2); a dummy variable for a household with predominantly black household members (column 3); education of the household head (column 4); a dummy variable for a female household head (column 5); and a dummy variable for an urban household. For all specifications, I define the bandwidth size to be the same as my preferred main specification, i.e. 10.356 years on either side of the threshold. Robust standard errors are reported in parenthesis.
* p<0.10, ** p<0.05, *** p<0.01

Figure 3.10: Continuity test of the race variable for Wave 1



Notes: The figure shows the unconditional mean and expectation function of the the race variable against the age variable on either side of the pension age eligibility threshold of 60 years. The race variable takes the form of a dummy variable indicating whether a household has predominantly black household members. 95% confidence intervals are shown.

Table 3.11: RD effect of the pension on remittances, by gender of oldest household member

	(1) Pension received	(2) Household Remittances	(3) Pension received	(4) Household Remittances
Pension eligibility (>=60 years)	0.522***		0.439***	
	(0.0292)		(0.0271)	
Pension eligibility × Female	0.0653**		0.0886***	
	(0.0260)		(0.0252)	
Pension received		-162.3		-122.0
		(127.1)		(132.0)
Pension received × Female		-179.6**		-169.6**
		(88.10)		(85.36)
RD effect for Female		-341.9***		-291.6***
Covariates	✓	✓		
Fixed effect	✓	✓		
N	2,964	2,964	3,702	3,702
MSE-optimal bandwidth	10.36	10.36	11.14	11.14

Notes: Columns (1) and (2) report the first and second stage respectively for a fully-interacted 2SLS model with gender of the oldest household member, using local linear regressions and controlling for predetermined household characteristics to increase the precision of estimates. Columns (3) and (4) report the first and second stage respectively excluding covariates. The base category include households where the oldest household member is male. 45% of households have a female oldest household member, whereas 55% of households have a male oldest member. Covariates include: household size; total number of co-resident children; modal race of the household; education and gender of household head; and urban status. I control for district as the geographic fixed effect. I use the same data-driven bandwidths selected for the main specification in Table 3.3. Robust standard errors are reported in parenthesis. * p<0.10, ** p<0.05, *** p<0.01

Table 3.12: RD effect of the pension on remittances, by household wealth

	(1) Pension received	(2) Household Remittances	(3) Pension received	(4) Household Remittances
Pension eligibility (>=60 years)	0.510***		0.503***	
	(0.0308)		(0.0294)	
Pension eligibility × Low assets	0.132***		0.124***	
	(0.0277)		(0.0269)	
Pension received		-359.7**		-275.5**
		(143.0)		(133.6)
Pension received × Low assets		116.2		74.25
		(89.54)		(86.20)
RD effect for Low assets		-243.4**		-201.2*
Covariates	✓	✓		
Fixed effect	✓	✓		
N	2,365	2,365	2,694	2,694
MSE-optimal bandwidth	10.36	10.36	11.14	11.14

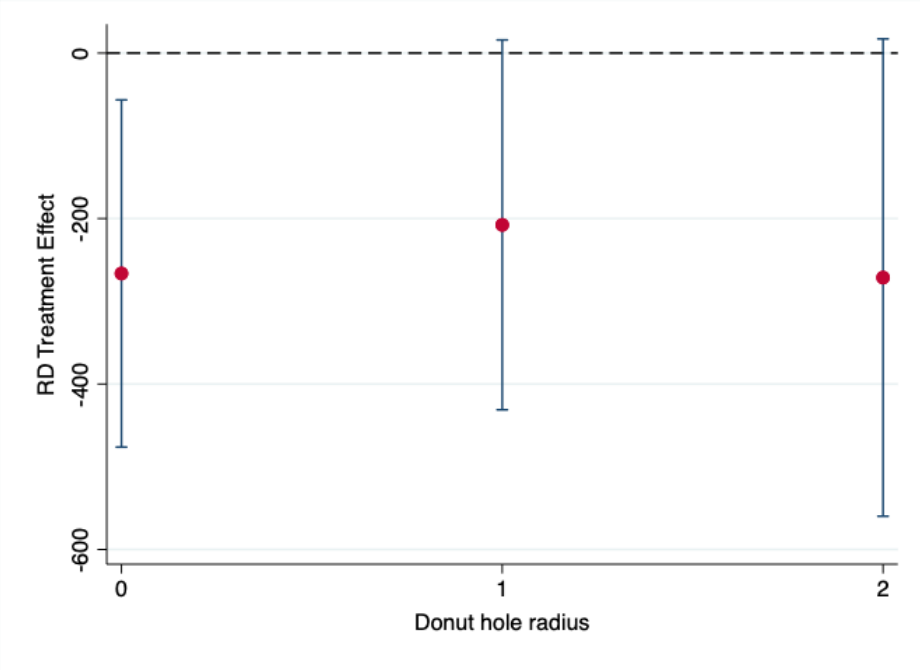
Notes: Columns (1) and (2) report the first and second stage respectively for a fully-interacted 2SLS model with household assets, using local linear regressions and controlling for predetermined household characteristics to increase the precision of estimates. Columns (3) and (4) report the first and second stage respectively excluding covariates. The base category include households with above-median assets. I use the household asset index created by Daniels and Augustine (2016). The index adds up all assets owned by the household, including business assets, vehicles, financial assets, superannuation and durable household assets. I use the household asset index constructed using Wave 4 data, as the only available and comprehensive asset index measure within NIDS, while recognising the limitations of using an asset measure from a future wave. Wave 2 of NIDS also collects asset data, but this measure omits household positions and thus alters the composition of household assets most markedly for low-income households (Daniels and Augustine, 2016). Covariates include: household size; total number of co-resident children; modal race of the household; education and gender of household head; and urban status. I control for district as the geographic fixed effect. I use the same data-driven bandwidths selected for the main specification in Table 3.3. Robust standard errors are reported in parenthesis. * $p<0.10$, ** $p<0.05$, *** $p<0.01$

Table 3.13: RD effect of the pension on household remittances, by bandwidth size

	(1) Remitt	(2) Remitt	(3) Remitt	(4) Remitt	(5) Remitt	(6) Remitt	(7) Remitt	(8) Remitt	(9) Remitt
Pension received	-276.8*	-263.5**	-257.0**	-249.1**	-266.4**	-242.3**	-260.7***	-238.5**	-219.3**
	(146.4)	(129.0)	(118.1)	(111.9)	(107.1)	(105.1)	(99.35)	(93.68)	(86.54)
Covariates	✓	✓	✓	✓	✓	✓	✓	✓	✓
Fixed effect	✓	✓	✓	✓	✓	✓	✓	✓	✓
N	1,867	2,135	2,417	2,661	2,964	3,220	3,501	3,761	4,019
Bandwidth size	6	7	8	9	10.36	11	12	13	14
F-stat for weak IV test	242.29	295.67	350.75	403.56	464.32	511.30	566.56	628.32	688.89
AR confidence intervals	[573,-4]	[-514,-23]	[-486,-37]	[-466,-32]	[-474,-59]	[-446,-38]	[-453,-68]	[-422,-55]	[-389,-50]

Notes: Each column reports the second-stage RD treatment effect of receiving the state pension on real household remittances, using local linear regressions and controlling for predetermined household characteristics to increase the precision of estimates. I manually vary the size of the bandwidth to test the sensitivity of the point estimates to bandwidth choice. The data-driven bandwidth that minimises the approximate mean squared error of the RD estimator is 10.36 years on either side of the age eligibility threshold. Covariates include: household size; total number of co-resident children; modal race of the household; education and gender of household head; and urban status. I control for district as the geographic fixed effect. Robust standard errors are reported in parenthesis. I also report the F-statistic for the weak IV test proposed by Olea and Pflueger (2013) and Anderson-Rubin confidence intervals. I use post-stratification weights created by Brophy et al. (2018). The remittances variable is winsorised at the 1st and 99th percentile to account for outliers. Monthly household remittances are deflated using the monthly consumer price index, with December 2016 as the base year. * p<0.10, ** p<0.05, *** p<0.01

Figure 3.11: Donut hole test



Notes: The markers indicate the local average treatment effect of receiving the state pension on real household remittances obtained from an instrumental variables approach using local linear regressions and controlling for predetermined household characteristics. 95% confidence intervals are shown, using robust standard errors. I present results for the main specification (no donut hole), omitting the first window and omitting the second window around the age eligibility threshold. The same bandwidth size – 10.36 years on either side of the threshold – is used for all specifications.

Table 3.14: RD effect of the pension on household remittances, using placebo thresholds

	(1) Remittances	(2) Remittances	(3) Remittances	(4) Remittances	(5) Remittances
Pension received	179.5	-312.3	-266.4**	-27.48	-782.5
	(322.4)	(273.2)	(107.1)	(313.2)	(540.2)
Covariates	✓	✓	✓	✓	✓
Fixed effect	✓	✓	✓	✓	✓
Observations	2,964	2,964	2,964	2,964	2,964
Threshold	55	57	60	63	65
F-stat for weak IV test	96.54	136.20	464.32	77.32	13.18
AR confidence intervals	[-446,831]	[-864,218]	[-474,59]	[-635,605]	[-2301,223]

Notes: Each column reports the second-stage RD treatment effect of receiving the state pension on real household remittances, using local linear regressions and controlling for predetermined household characteristics to increase the precision of estimates. In addition to the true age eligibility threshold (column 3), I use four placebo thresholds to test whether there are any discontinuities away from the true threshold: age 55 (column 1), age 57 (column 2), age 63 (column 4) and age 65 (column 5). I use the same bandwidth size - 10.36 years on either side of the threshold - for all specifications. Covariates include: household size; total number of co-resident children; modal race of the household; education and gender of household head; and urban status. I control for district as the geographic fixed effect. Robust standard errors are reported in parenthesis. I also report the F-statistic for the weak IV test proposed by Olea and Pflueger (2013) and Anderson-Rubin confidence intervals. I use post-stratification weights created by Brophy et al. (2018). The remittances variable is winsorised at the 1st and 99th percentile to account for outliers. Monthly household remittances are deflated using the monthly consumer price index, with December 2016 as the base year. * $p<0.10$, ** $p<0.05$, *** $p<0.01$

Conclusion

This thesis speaks to the broad theme of the effectiveness of policy responses to shocks. Chapter 1 on anticipatory cash transfers presented quasi-experimental evidence that a timelier response is more impactful in mitigating the worst effects of a climate disaster. Although we were unable to directly compare an anticipatory approach to a traditional response, we captured the welfare benefits of the anticipatory cash transfers three months after the intervention, which is before a traditional humanitarian intervention would have even been implemented. Moreover, we use the variation in the number of days between receiving a cash transfer and peak local flooding to show that speed of delivery matters: an earlier cash transfer had stronger effects on food consumption even three months later, with dissipating welfare effects for each day that transfers are delayed.

Several questions on anticipatory cash transfers remain open for further investigation. First, what is the optimal timing for delivering cash transfers in anticipation of climate shocks? A direct comparison between anticipatory cash transfers to a traditional response could serve to quantify the relative effectiveness of an anticipatory cash transfer. In our study, we find evidence that the cash transfer provided important liquidity to households at a critical juncture in their decision making, which enabled households to act in ways that improved their welfare three months after the intervention. While we provide suggestive evidence on the mechanisms at play, further theoretical and empirical work is needed to understand the ways in which anticipatory cash transfers support household behavioural change, which may differ relative to a traditional response.

Second, while we present strong evidence in favour of timing, we cannot speak to the question of how best to target households in anticipation of a climate

shock. There is a trade-off between targeting efficiency and timing when budgets are limited. While ex-post interventions might be better at targeting those worst affected by the shock (assuming that more information is available), they are less likely to be impactful relative to a timelier, but less well targeted approach. This tension is likely to become less prominent over time, as advances in data quality and machine learning improve targeting efficiency. For example, as a first proof of concept, Aiken et al. (2021) use traditional survey-based measures of consumption and wealth to train algorithms on non-traditional big data to predict poverty. These poverty predictions were used to target households through the Novissi cash transfer programme during the Covid-19 pandemic, in partnership with the Government of Togo.[26] Further research is needed to show whether this approach could be extended to identify at-risk households in real-time and at scale in the context of shocks, such as in climate disaster response. For instance, forecast-based triggers should be designed to measure the predicted impact of the shock on households, rather than just the magnitude of the climate shock itself.

The final question relates to the external validity and scale up potential of using anticipatory cash transfers in response to other shocks. For example, the optimal timing for delivering cash transfers is not evident for slow onset climate disasters, such as drought, given the large window between forecasting the shock and its impact materialising. While anticipatory cash transfers are likely to be effective across a multitude of shocks, further research is needed on the design of such programmes.

Chapter 2 speaks to the broader question of whether policy could address behavioural constraints that may arise due to exposure to shocks. Specifically, the chapter provides a strong proof of concept that mental imagery – the ability to conger images, rather than words – might matter for decision making. In our randomised control trial, we designed an imagery training to be identical in format and entrepreneurship content to a placebo entrepreneurship training. The sole difference between the two training lies in the use of imagery exercises in the former, and written and verbal group-based activities in the latter. When we compare the

[26] Aiken et al. (2021) first created poverty maps to identify the poorest geographies in Togo using satellite prediction, before identifying the poorest people within the target geographies using cell usage patterns.

imagery and placebo training, we find an improvement in the frequency and quality of imagery used for imagery participants, especially when applied to business decision making. Moreover, imagery participants report higher earnings relative to the placebo group.

However, we do not find promising results in favour of business trainings in general. When compared to the no-intervention group, we cannot reject the hypothesis that the imagery training had no effect on any of our outcomes of interest, even if our results point to the direction of positive effects. Moreover, we find that the placebo training creates significantly negative effects on earnings relative to the no-intervention group. There is some evidence of differential attrition in survey response that could partly explain these results, but we cannot rule out the hypothesis that the business training could be detrimental. The broader literature on business training finds modestly positive effects at most (McKenzie, 2021; McKenzie, Woodruff et al. 2021). However, further research is needed to understand why we observe these effects and the broader implications for policy.

For mental imagery to be used more widely in public policy, further work is also needed on the mechanisms through which imagery affects economic outcomes. Our current working hypothesis is that traumatic experiences accumulate in people's memory and interfere with the recollection of positive past experiences. In turn, such a biased recollection affects current and future-oriented decision making, making people more risk averse and prone to stick with the status quo. We propose that mental imagery can rebalance people's memory database and thus improve economic decisions. Imagery allows people to simulate positive experiences in a safe space, and such experiences get stored in memory "as-if" they were real.

In Chapter 3, I find significant evidence that social protection systems do not operate in a vacuum and show that public cash transfers can interact negatively with private transfers in the context of post-Apartheid South Africa. Every 1 rand of pension income received leads to 0.22 rand decrease in remittance income on average for age-eligible households that take up the pension. Age-eligible households located further away from the age eligibility threshold continue to experience a reduction in remittances within the cross-sectional data of interest. My

findings suggest that the welfare effect of the public cash transfers for direct recipients may be neutralised by the endogenous response in household behaviour. Instead, the benefits of the pension also appear to accrue to other family members. This is likely to be especially true in middle-income country contexts where risk-sharing networks are widespread and social protection has been expanding during the Covid-19 pandemic.

However, this paper does not take a stance on the broader welfare implications of the crowding out effect. To speak to these implications, further research is required on two open questions. First, what are the effects of crowding out on remitters? Does their consumption go up in response to their sending reduced remittances? I am limited by my data to speak to this question directly. Second, given that we observe small to no crowding out in high-income countries and substantial crowding out in middle-income countries, what happens as countries adopt more established and widespread social protection programmes? Are private support networks unravelled over time? I do not find evidence of crowding out around the age eligibility threshold in later survey years of NIDS. Further work on the trajectory of private support and the dynamic nature of the displacement effect over time would be worthy of further investigation.

In sum, more effective policy responses are needed to address intensifying and more frequent income shocks due to climate change and conflict. Policy innovations that use data to forecast the shocks and act early, as well as those that combine insights across multiple disciplines, show promise. This thesis contributes to a growing evidence base on how to design more resilient systems and strengthen households' capacity to navigate these shocks on their own terms.

References

Abay, K., Berhane, G., Hoddinott, J., and Tafere, K. (2020). COVID-19 and Food Security in Ethiopia: Do Social Protection Programs Protect?

Adams, T., Sridhar, P., and Kaur, J. (2021). Collective Anticipatory Humanitarian Action for Monsoon Flooding - Bangladesh.

Adhvaryu, A., Nyshadham, A., Molina, T., and Tamayo, J. (2018). Helping children catch up: Early life shocks and the progresa experiment. *NBER Working Paper*, 24848.

Aguayo, V. M., Sharma, A., and Subedi, G. R. (2015). Delivering essential nutrition services for children after the Nepal earthquake. *The Lancet Global Health*, 3(11):665–e666.

Aiken, E., Bellue, S., Karlan, D., Udry, C., and Blumenstock, J. (2021). Machine learning and mobile phone data can improve the targeting of humanitarian assistance. *NBER Working Paper Series*, 29070.

Aker, J. C. (2013). Cash or Coupons? Testing the Impacts of Cash versus Vouchers in the Democratic Republic of Congo. *CGD Working Paper*, 320.

Aker, J. C., Boumnijel, R., McClelland, A., and Tierney, N. (2016). Payment mechanisms and antipoverty programs: Evidence from a mobile money cash transfer experiment in Niger. *Economic Development and Cultural Change*, 65(1):1–37.

Alan, S. and Ertac, S. (2018). Fostering patience in the classroom: Results from randomized educational intervention. *Journal of Political Economy*, 126(5):1865 – 1911.

Alaref, J., Brodmann, S., and Premand, P. (2020). The medium-term impact of entrepreneurship education on labor market outcomes: Experimental evidence from university graduates in tunisia. *Labour Economics*, 62.

Albarran, P. and Attanasio, O. P. (2003). Limited commitment and crowding out of private transfers: Evidence from a randomised experiment. *The Economic Journal*, 113:77–85.

Alderman, H., Hoddinott, J., and Kinsey, B. (2006). Long term consequences of early childhood malnutrition. *Oxford Economic Papers*, 58(3):450–474.

Alibhai, S., Buehren, N., Frese, M., Goldstein, M., Papineni, S., and Wolf, K. (2019). Full esteem ahead? mindset-oriented business training in ethiopia. *World Bank Policy Research Working Paper*, 8892.

Altonji, J., Hayashi, F., and Kotlikoff, L. (1992). Is the extended family altruistically linked? direct tests using micro data. *The American Economic Review*, 82.

Altonji, J., Hayashi, F., and Kotlikoff, L. (1997). Parental altruism and inter vivos transfers: Theory and evidence. *Journal of Political Economy*, 105.

Amuedo-Dorantes, C. and Juarez, L. (2015). Old-age government transfers and the crowding out of private gifts: The 70 and above program for the rural elderly in mexico. *Southern Economic Journal*, 81.

Anderson, M. L. (2008). Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects. *Journal of the American Statistical Association*, 103(484):1481–1495.

Andrade, J., May, J., Deeprose, C., Baugh, S.-J., and Ganis, G. (2014). Assessing vividness of mental imagery: The plymouth sensory imagery questionnaire. *British Journal of Psychology*, 105(4):547–563.

Andrews, C., Hsiao, A., and Ralston, L. (2018). Social Safety Nets Promote Poverty Reduction, Increase Resilience, and Expand Opportunities. In Beegle, K., Coudouel, A., and Monsalve, E., editors, *Realizing the Full Potential of Social Safety Nets in Africa*, chapter 2, pages 87–138. World Bank, Washington, DC.

Anthony, W. P., Bennett III, R. H., Maddox, E. N., and Wheatley, W. J. (1993). Picturing the future: Using mental imagery to enrich strategic environmental assessment. *Academy of Management Perspectives*, 7(2):43–56.

Ardington, C., Case, A., and Hosegood, V. (2009). Labor supply response to large social transfers: Longitudinal evidence from south africa. *American Economic Journal: Applied Economics*, 1:22–48.

Arndt, C., Davies, R., Gabriel, S., Harris, L., Makrelov, K., Robinson, S., Levy, S., Simbanegavi, W., van Seventer, D., and Anderson, L. (2020). Covid-19 lockdowns, income distribution, and food security: An analysis for South Africa. *Global Food Security*, 26:100410.

Arntz, A., Rauner, M., and Van den Hout, M. (1995). "if i feel anxious, there must be danger": Ex-consequentia reasoning in inferring danger in anxiety disorders. *Behaviour Research and Therapy*, 33(8):917 – 925.

Asfaw, S., Carraro, A., Davis, B., Handa, S., and Seidenfeld, D. (2017). Cash transfer programmes, weather shocks and household welfare: evidence from a randomised experiment in Zambia. *Journal of Development Effectiveness*, 9(4):419–442.

Asukai, N., Kato, H., Kawamura, N., Kim, Y., Yamamoto, K., Kishimoto, J., Miyake, Y., and Nishizono-Maher, A. (2002). Reliabiligy and validity of the japanese-language version of the impact of event scale-revised (ies-rj): four studies of different traumatic events. *The Journal of nervous and mental disease*, 190(3):175–182.

Banks, S. and Dinges, D. F. (2007). Behavioral and physiological consequences of sleep restriction. *Journal of Clinical Sleep Medicine*, 3(5):519–528.

Barro, R. J. (1974). Are government bonds net wealth? *Journal of Political Economy*, 82.

Battistin, F. (2016). Lebanon Cash Consortium (LCC) impact evaluation of the multipurpose cash assistance programme.

Beck, J. G., Grant, D. M., Read, J. P., Clapp, J. D., Coffey, S. F., Miller, L. M., and Palyo, S. A. (2008). The impact of event scale-revised: psychometric properties in a sample of motor vehicle accident survivors. *Journal of anxiety disorders*, 22(2):187–198.

Becker, G. S. (1974). A theory of social interactions. *Journal of Political Economy*, 82.

Becker, G. S. and Mulligan, C. B. (1997). The Endogenous Determination of Time Preference*. *The Quarterly Journal of Economics*, 112(3):729–758.

Benjamini, Y., Krieger, A. M., and Yekutieli, D. (2006). Adaptive linear step-up procedures that control the false discovery rate. *Biometrika*, 93(3):491–507.

Benjet, C., Bromet, E., Karam, E. G., Kessler, R. C., McLaughlin, K. A., Ruscio, A. M., Shahly, V., Stein, D. J., Petukhova, M., Hill, E., Alonso, J., Atwoli, L., Bunting, B., Bruffaerts, R., de Almeida, J. M. C., de Girolamo, G., Florescu, S., Gureje, O., Huang, Y., Lepine, J. P., Kawakami, N., Kovess-Masfety, V., Medina-Mora, M. E., Navarro-Mateu, F., Piazza, M., Posada-Villa, J., Scott, K. M., Shalev, A., Slade, T., ten Have, M., Torres, Y., Viana, M. C., Zarkov, Z., and Koenen, K. C. (2016). The epidemiology of traumatic event exposure worldwide: Results from the world mental health survey consortium. *Psychological medicine*, 46:327–343.

Bertrand, M., Mullainathan, S., and Miller, D. (2003). Public policy and extended families: Evidence from pensions in south africa. *The World Bank Economic Review*, 17:27–50.

Bessone, P., Rao, G., Schilbach, F., Schofield, H., and Toma, M. (2020). The Economic Consequences of Increasing Sleep Among the Urban Poor. *NBER Working Paper*, 26746.

Bhorat, H. (2004). Labour market challenges in the post-apartheid south africa. *South African Journal of Economics*, 72.

Bordalo, P., Coffman, K., Gennaioli, N., and Shleifer, A. (2016). Stereotypes. *The Quarterly Journal of Economics*, 131(4):1753–1794.

Bordalo, P., Gennaioli, N., and Shleifer, A. (2018). Diagnostic expectations and credit cycles. *The Journal of Finance*, 73(1):199–227.

Bottan, N., Hoffmann, B., and Vera-Cossio, D. A. (2021). Stepping up during a crisis: The unintended effects of a noncontributory pension program during the Covid-19 pandemic. *Journal of Development Economics*, 150:102635.

Bremner, J. D. (2006). Traumatic stress: effects on the brain. *Dialogues in clinical neuroscience*, 8(4):445–461.

Brophy, T., Branson, N., Daniels, R. C., Leibbrandt, M., Mlatsheni, C., and Woolard, I. (2018). National Income Dynamics Study Panel User Manual.

Brunnermeier, M. K. and Parker, J. A. (2005). Optimal expectations. *American Economic Review*, 95(4):1092–1118.

Buckley, T. C., Blanchard, E. B., and Neill, W. (2000). Information processing and ptsd: A review of the empirical literature. *Clinical Psychology Review*, 20(8):1041–1065.

Böhm-Bawerk, E. V. (1889). *The Positive Theory of Capital*.

Callen, M., Isaqzadeh, M., Long, J. D., and Sprenger, C. (2014). Violence and risk preference: Experimental evidence from afghanistan. *American Economic Review*, 104(1):123–48.

Calonico, S., Cattaneo, M. D., and Titiunik, R. (2014). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica*, 82:2295–2326.

Campos, F., Frese, M., Goldstein, M., Iacovone, L., Johnson, H., McKenzie, D., and Mensmann, M. (2017). Teaching personal initiative beats traditional training in boosting small business in west africa. *Science*, 357:1287–1290.

Carroll, J. S. (1978). The effect of imagining an event on expectations for the event: An interpretation in terms of the availability heuristic. *Journal of Experimental Social Psychology*, 14:88–96.

Carsell, K. M., Pingel, N. D., and Ford, D. T. (2004). Quantifying the Benefit of a Flood Warning System. *Natural Hazards Review*, 5(3):131–140.

Case, A. and Deaton, A. (1998). Large cash transfers to the elderly in south africa. *The Economic Journal*, 108.

Castro, L. S., Poyares, D., Leger, D., Bittencourt, L., and Tufik, S. (2013). Objective prevalence of insomnia in the São Paulo, Brazil epidemiologic sleep study. *Annals of Neurology*, 74(4):537–546.

Cattaneo, M. D., Idrobo, N., and Titiunik, R. (2019). *A Practical Introduction to Regression Discontinuity Designs*. Cambridge University Press.

Chen, Y. and Zhou, L. A. (2007). The long-term health and economic consequences of the 1959-1961 famine in China. *Journal of Health Economics*, 26(4):659–681.

Chioda, L., Contreras-Loya, D., Gertler, P., and Carney, D. (2021). Making entrepreneurs: Returns to training youth in hard versus soft business skills. *NBER Working Paper Series*, 28845.

Clavijo, I. (2020). Long-term Impacts of the timing of Conditional Cash Transfers on Households' Economic Mobility.

Conway, Singer, and Tagini, A. (2004). Conway, m. a., singer, j. a., tagini, a. (2004). the self andautobiographical memory: Correspondence and coherence.social cognition, 22, 495–537. *Social Cognition*.

Cox, D. (1987). Motives for Private Income Transfers. *Journal of Political Economy*, 95(3):508–546.

Cox, D., Eser, Z., and Jimenez, E. (1998). Motives for private transfers over the life cycle: An analytical framework and evidence for Peru. *Journal of Development Economics*, 55:57–80.

Cox, D., Hansen, B. E., and Jimenez, E. (2004). How responsive are private transfers to income? evidence from a laissez-faire economy. *Journal of Public Economics*, 88:2193–2219.

Cox, D. and Jakubson, G. (1995). The connection between public transfers and private interfamily transfers. *Journal of Public Economics*, 57(1):129–167.

Cunha, J. M. (2014). Testing paternalism: Cash versus in-kind transfers. *American Economic Journal: Applied Economics*, 6(2):195–230.

Cunha, J. M., De Giorgi, G., and Jayachandran, S. (2019). The Price Effects of Cash Versus In-Kind Transfers. *Review of Economic Studies*, 86(1):240–281.

Daniels, R. C. and Augustine, T. (2016). The measurement and distribution of household wealth in south africa using the national income dynamics study (nids) wave 4. *Southern Africa Labour and Development Research Unit Working Paper*, 183.

de Janvry, A., Finan, F., Sadoulet, E., and Vakis, R. (2006). Can conditional cash transfer programs serve as safety nets in keeping children at school and from working when exposed to shocks? *Journal of Development Economics*, 79(2):349–373.

De Mel, S., McKenzie, D., and Woodruff, C. (2012). Enterprise Recovery Following Natural Disasters. *Economic Journal*, 122(559):64–91.

Deaton, A. (2008). Income, health, and well-being around the world: Evidence from the Gallup World Poll. *Journal of Economic Perspectives*, 22(2):53–72.

Del Carpio, X. V. and Macours, K. (2010). *Leveling the intra-household playing field: Compensation and specialization in child labor allocation*, volume 31. World Bank.

Dercon, S. (2002). Income risk, coping strategies, and safety nets. *The World Bank Research Observer*, 17:141–166.

Dercon, S. (2004). Growth and shocks: Evidence from rural Ethiopia. *Journal of Development Economics*, 74(2):309–329.

Dercon, S. and Porter, C. (2014). Live aid revisited: Long-term impacts of the 1984 Ethiopian famine on children. *Journal of the European Economic Association*, 12(4):927–948.

Dinkelman, T. (2017). Long-run Health Repercussions of Drought Shocks: Evidence from South African Homelands. *Economic Journal*, 127(604):1906–1939.

Doocy, S. and Tappis, H. (2017). Cash-based approaches in humanitarian emergencies: a systematic review. *Campbell Systematic Reviews*, 13(1):1–200.

Douglas Bremner, J., Randall, P., Scott, T. M., Capelli, S., Delaney, R., McCarthy, G., and Charney, D. S. (1995). Deficits in short-term memory in adult survivors of childhood abuse. *Psychiatry Research*, 59(1):97–107.

Driskell, J., Copper, C., and Moran, A. (1994). Does mental practice enhance performance? *Journal of Applied Psychology*, 79:481–492.

Duflo, E. (2003). Grandmothers and granddaughters: Old-age pensions and intrahousehold allocation in south africa. *The World Bank Economic Review*, 17.

Edmonds, E. (2006). Child labor and schooling responses to anticipated income in south africa. *Journal of Development Economics*, 81:386–414.

Edmonds, E., Mammen, K., and Miller, D. (2004). Rearranging the family? income support and elderly living arrangements in a low-income country. *The Journal of Human Resources*, 40.

Eil, D. and Rao, J. M. (2011). The good news-bad news effect: Asymmetric processing of objective information about yourself. *American Economic Journal: Microeconomics*, 3(2):114–38.

Ferguson, J. (2015). *Give a man a fish: Reflections on the new politics of distribution*. Duke University Press.

Forgas, J. (1995). Mood and judgment: The affect infusion model (aim). *Psychological bulletin*, 117:39–66.

Gabaix, X. and Laibson, D. (2017). Myopia and Discounting. Technical report, National Bureau of Economic Research, Cambridge, MA.

Gelman, A. and Imbens, G. (2019). Why high-order polynomials should not be used in regression discontinuity designs. *Journal of Business and Economic Statistics*, 37:447–456.

Gennaioli, N. and Shleifer, A. (2010). What Comes to Mind. *Quarterly Journal of Economics*, 125(4):1399–1433.

Gentilini, U., Almenfi, M., Orton, I., and Dale, P. (2020). Social Protection and Jobs Responses to COVID-19: A Real-Time Review of Country Measures (September 18, 2020).

Gentilini, U., Almenfi, M., Orton, I., and Dale, P. (2021). Social Protection and Jobs Responses to COVID-19: A Real-Time Review of Country Measures.

Gerardi, K. and Tsai, Y. (2014). The effect of social entitlement programmes on private transfers: New evidence of crowding out. *Economica*, 81:721–746.

Gregory, W. L., Cialdini, R. B., and Carpenter, K. M. (1982). Self-relevant scenarios as mediators of likelihood estimates and compliance: Does imagining make it so?

Gros, C., Bailey, M., Schwager, S., Hassan, A., Zingg, R., Uddin, M. M., Shahjahan, M., Islam, H., Lux, S., Jaime, C., and Coughlan de Perez, E. (2019). Household-level effects of providing forecast-based cash in anticipation of extreme weather events: Quasi-experimental evidence from humanitarian interventions in the 2017 floods in Bangladesh. *International Journal of Disaster Risk Reduction*, 41:101275.

Gros, C., Easton-Calabria, E., Bailey, M., Dagys, K., de Perez, E. C., Sharavnyambuu, M., and Kruczkiewicz, A. (2020). The effectiveness of forecast-based humanitarian assistance in anticipation of extreme winters: Evidence from an intervention for vulnerable herders in mongolia. *Disasters*.

Guiteras, R., Jina, A., and Mushfiq Mobarak, A. (2015). Satellites, self-reports, and submersion: Exposure to Floods in Bangladesh. *American Economic Review*, 105(5):232–236.

Hahn, J., Todd, P., and Klaauw, W. V. D. (2001). Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica*, 69:201–209.

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

Haushofer, J. and Shapiro, J. (2018). The Long-Term Impact of Unconditional Cash Transfers: Experimental Evidence from Kenya.

Hershfield, H. E., John, E. M., and Reiff, J. S. (2018). Using vividness interventions to improve financial decision making. *Policy Insights from the Behavioral and Brain Sciences*, 5:209–215.

Hidrobo, M., Hoddinott, J., Peterman, A., Margolies, A., and Moreira, V. (2014). Cash, food, or vouchers? Evidence from a randomized experiment in northern Ecuador. *Journal of Development Economics*, 107:144–156.

Holbrook, T. L., Hoyt, D. B., Stein, M. B., and Sieber, W. J. (2002). Gender differences in long-term posttraumatic stress disorder outcomes after major trauma: women are at higher risk of adverse outcomes than men. *Journal of Trauma and Acute Care Surgery*, 53(5):882–888.

Holmes, E. and Mathews, A. (2010). Mental imagery in emotion and emotional disorders. *Clinical psychology review*, 30:349–62.

Holmes, E. A., Arntz, A., and Smucker, M. R. (2007). Imagery rescripting in cognitive behaviour therapy: Images, treatment techniques and outcomes. *Journal of Behavior Therapy and Experimental Psychiatry*, 38(4):297–305.

Holmes, E. A., Blackwell, S. E., Burnett Heyes, S., Renner, F., and Raes, F. (2016). Mental imagery in depression: Phenomenology, potential mechanisms, and treatment implications. *Annual review of clinical psychology*, 12(1):249–280.

Holmes, E. A., Hales, S. A., Young, K., and Di Simplicio, M. (2019). *Imagery-Based Cognitive Therapy for Bipolar Disorder and Mood Instability*. The Guilford Press, New York.

Holmes, E. A., Lang, T. J., Moulds, M. L., and Steele, A. M. (2008). Prospective and positive mental imagery deficits in dysphoria. *Behaviour Research and Therapy*, 46(8):976–981.

Holmes, E. A., O'Connor, R. C., Perry, V. H., Tracey, I., Wessely, S., Arseneault, L., Ballard, C., Christensen, H., Silver, R. C., Everall, I., Ford, T., John, A., Kabir, T., King, K., Madan, I., Michie, S., Przybylski, A. K., Shafran, R., Sweeney, A., Worthman, C. M., Yardley, L., Cowan, K., Cope, C., Hotopf, M., and Bullmore, E. (2020). Multidisciplinary research priorities for the covid-19 pandemic: a call for action for mental health science. *The Lancet Psychiatry*, 7:547–560.

Hou, X. (2010). Can drought increase total calorie availability? The impact of drought on food consumption and the mitigating effects of a conditional cash transfer program. *Economic Development and Cultural Change*, 58(4):713–737.

Hunt, N. and Evans, D. (2004). Predicting traumatic stress using emotional intelligence. *Behaviour Research and Therapy*, 42(7):791–798.

Imbens, G. and Kalyanaraman, K. (2012). Optimal bandwidth choice for the regression discontinuity estimator. *Review of Economic Studies*, 79:933–959.

Imbens, G. W. and Rubin, D. B. (2015). *Causal inference: For statistics, social, and biomedical sciences an introduction*. Cambridge University Press.

International Finance Corporation (2016). IFC Inclusive Business Case Study: bKash.

IPCC (2021). *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press.

Ivaschenko, O., Doyle, J., Kim, J., Sibley, J., and Majoka, Z. (2020). Does "Manna from Heaven" help? The role of cash transfers in disaster recovery—lessons from Fiji after Tropical Cyclone Winston. *Disasters*, 44(3):455–476.

Jenkins, M., Langlais, P., Delis, D., and Cohen, R. (1998). Learning and memory in rape victims with posttraumatic stress disorder. *The American Journal of Psychiatry*, 155:278–9.

Jensen, N. D., Barrett, C. B., and Mude, A. G. (2017). Cash transfers and index insurance: A comparative impact analysis from northern Kenya. *Journal of Development Economics*, 129:14–28.

Jensen, R. (2003). Do private transfers 'displace' the benefits of public transfers? evidence from south africa. *Journal of Public Economics*, 88(1-2):89–112.

John, A. and Orkin, K. (2021). Can simple psychological interventions increase preventive health investment? *Journal of the European Economic Association, Conditionally accepted*.

Juarez, L. (2009). Crowding out of private support to the elderly: Evidence from a demogrant in Mexico. *Journal of Public Economics*, 93(3-4):454–463.

Kazianga, H. (2006). Motives for household private transfers in burkina faso. *Journal of Development Economics*, 79:73–117.

Kessler, R. C., Aguilar-Gaxiola, S., Alonso, J., Benjet, C., Bromet, E. J., Cardoso, G., Degenhardt, L., de Girolamo, G., Dinolova, R. V., Ferry, F., Florescu, S., Gureje, O., Haro, J. M., Huang, Y., Karam, E. G., Kawakami, N., Lee, S., Lepine, J. P., Levinson, D., Navarro-Mateu, F., Pennell, B. E., Piazza, M., Posada-Villa, J., Scott, K. M., Stein, D. J., Have, M. T., Torres, Y., Viana, M. C., Petukhova, M. V., Sampson, N. A., Zaslavsky, A. M., and Koenen, K. C. (2017). Trauma and PTSD in the WHO World Mental Health Surveys. *European Journal of Psychotraumatology*, 8.

Kleim, B. and Ehlers, A. (2008). Reduced autobiographical memory specificity predicts depression and posttraumatic stress disorder after recent trauma. *Journal of Consulting and Clinical Psychology*, 76(2):231.

Kleim, B., Graham, B., Fihosy, S., Stott, R., and Ehlers, A. (2014). Reduced specificity in episodic future thinking in posttraumatic stress disorder. *Clinical Psychological Science*, 2(2):165–173.

Kling, J. R., Liebman, J. B., and Katz, L. F. (2007). Experimental analysis of neighborhood effects. *Econometrica*, 75(1):83–119.

Knippenberg, E. and Hoddinott, J. (2017). Shocks, social protection, and resilience: Evidence from Ethiopia.

Koehler, D. J. (1991). Explanation, imagination, and confidence in judgment.

Kolesár, M. and Rothe, C. (2018). Inference in regression discontinuity designs with a discrete running variable. *American Economic Review*, 108:2277–2304.

Kreibich, H., Thieken, A. H., Petrow, T., Müller, M., and Merz, B. (2005). Flood loss reduction of private households due to building precautionary measures – Lessons learned from the Elbe flood in August 2002. *Natural Hazards and Earth System Science*, 5(1):117–126.

Künemund, H. and Rein, M. (1999). There is more to receiving than needing: Theoretical arguments and empirical explorations of crowding in and crowding out. *Ageing and Society*, 19:93–121.

Lai, M. S. and Orsuwan, M. (2009). Examining the Impact of Taiwan's Cash Allowance Program on Private Households. *World Development*, 37:1250–1260.

Lee, D. S. (2009). Training, wages, and sample selection: Estimating sharp bounds on treatment effects. *Review of Economic Studies*, 76(3):1071–1102.

Lee, D. S. and Card, D. (2008). Regression discontinuity inference with specification error. *Journal of Econometrics*, 142:655–674.

Lee, D. S. and Lemieux, T. (2010). Regression discontinuity designs in economics. *Journal of Economic Literature*, 48:281–355.

Lehmann, C. and Masterson, D. T. R. (2014). Emergency economies: The impact of cash assistance in Lebanon. An impact evaluation of the 2013–2014 winter cash assistance program for Syrian Refugees in Lebanon. Technical report, International Rescue Committee, New York.

Leroy, J. L., Gadsden, P., Rodríguez-Ramírez, S., and De Cossío, T. G. (2010). Cash and in-kind transfers in poor rural communities in Mexico increase household fruit, vegetable, and micronutrient consumption but also lead to excess energy consumption. *Journal of Nutrition*, 140(3):612–617.

Li, D., Christ, S. E., and Cowan, N. (2014). Domain-general and domain-specific functional networks in working memory. *Neuroimage*, 102:646–656.

Libby, L., Shaeffer, E., Eibach, R., and Slemmer, J. (2007). Picture yourself at the polls visual perspective in mental imagery affects self-perception and behavior. *Psychological Science*, 18:199–203.

Lim, J. and Dinges, D. F. (2010). A Meta-Analysis of the Impact of Short-Term Sleep Deprivation on Cognitive Variables. *Psychological Bulletin*, 136(3):375–389.

Loewenstein, G. (1992). *The Fall and Rise of Psychological Explanations in the Economics of Intertemporal Choice*, pages 3–34. Russell Sage Foundation.

Maccini, S. and Yang, D. (2009). Under the weather: Health, schooling, and economic consequences of early-life rainfall. *American Economic Review*, 99(3):1006–1026.

Mackinnon, J. G. and Webb, M. D. (2017). Wild Bootstrap Inference for Wildly Different Cluster Sizes. *Journal of Applied Econometrics*, 32(2):233–254.

MacLeod, A. K. (1996). Affect, emotional disorder, and future-directed thinking. *Cognition & Emotion*, 10(1):69–86.

Macours, K., Premand, P., and Vakis, R. (2012). Transfers, diversification and household risk strategies: experimental evidence with lessons for climate change adaptation. *Policy Research Working Paper*, 6053.

Maitra, P. and Ray, R. (2003). The effect of transfers on household expenditure patterns and poverty in South Africa. *Journal of Development Economics*, 71:23–49.

Mansur, A., Doyle, J., and Ivaschenko, O. (2017). Social Protection and Humanitarian Assistance Nexus for Disaster Response.

Manuel, T. (2008). *Budget Speech*. Department of National Treasury, South African Ministry of Finance.

Mathews, A., Ridgeway, V., and Holmes, E. A. (2013). Feels like the real thing: Imagery is both more realistic and emotional than verbal thought. *Cognition and Emotion*, 27(2):217–229. PMID: 22780220.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics*, 142:698–714.

McKenzie, D. (2021). Small business training to improve management practices in developing countries: re-assessing the evidence for 'training doesn't work'. *Oxford Review of Economic Policy*, 37:276–301.

McKenzie, D. and Woodruff, C. (2014). What are we learning from business training and entrepreneurship evaluations around the developing world? *The World Bank Research Observer*, 29(1):48–82.

McKenzie, D., Woodruff, C., Bjorvatn, K., Bruhn, M., Cai, J., Gonzalez-Uribe, J., Quinn, S., Sonobe, T., and Valdivia, M. (2021). Training entrepreneurs. *VoxDevLit*, 1.

McNally, R. J., Lasko, N. B., Macklin, M. L., and Pitman, R. K. (1995). Autobiographical memory disturbance in combat-related posttraumatic stress disorder. *Behaviour research and therapy*, 33(6):619–630.

Moore, S. A. and Zoellner, L. A. (2007). Overgeneral autobiographical memory and traumatic events: an evaluative review. *Psychological bulletin*, 133(3):419.

Moya, A. (2018). Violence, psychological trauma, and risk attitudes: Evidence from victims of violence in Colombia. *Journal of Development Economics*, 131:15–27.

Mullainathan, S. (2002). A Memory-Based Model of Bounded Rationality*. *The Quarterly Journal of Economics*, 117(3):735–774.

Munroe, K. J., Giacobbi Jr., P. R., Hall, C., and Weinberg, R. (2000). The four Ws of imagery use: Where, when, why, and what.

National Treasury (2012). Budget Review 2012.

Neck, C. P., Neck, H. M., Manz, C. C., and Godwin, J. (1999). "i think i can; i think i can": A self-leadership perspective toward enhancing entrepreneur thought patterns, self-efficacy, and performance. *Journal of Managerial Psychology*.

Nelis, S., Holmes, E. A., Griffith, J. W., and Raes, F. (2019). Mental imagery during daily life: Psychometric evaluation of the spontaneous use of imagery scale (suis). *Psychologica Belgica*, 54(1):19.

Newby, J. and Moulds, M. (2011). Characteristics of intrusive memories in a community sample of depressed, recovered depressed and never-depressed individuals. *Behaviour research and therapy*, 49:234–43.

OECD (2017). *Pensions at a glance 2017: OECD and G20 Indicators*. OECD Publishing.

Ohtani, T., Iwanami, A., Kasai, K., Yamasue, H., Kato, T., Sasaki, T., and Kato, N. (2004). Post-traumatic stress disorder symptoms in victims of tokyo subway attack: a 5-year follow-up study. *Psychiatry and Clinical Neurosciences*, 58(6):624–629.

Olea, J. L. M. and Pflueger, C. (2013). A robust test for weak instruments. *Journal of Business and Economic Statistics*, 31:358–369.

Oosterbeek, H., Van Praag, M., and Ijsselstein, A. (2010). The impact of entrepreneurship education on entrepreneurship skills and motivation. *European Economic Review*, 54(3):442–454.

Osborne, M. S., Greene, D. J., and Immel, D. T. (2014). Managing performance anxiety and improving mental skills in conservatoire students through performance psychology training: a pilot study. *Psychology of Well-Being*, 4(1):18.

Parker, S. W. and Todd, P. E. (2017). Conditional cash transfers: The case of progresa/oportunidades. *Journal of Economic Literature*, 55:866–915.

Pearson, J. (2019). The human imagination: the cognitive neuroscience of visual mental imagery. *Nature Reviews Neuroscience*, 20(10):624–634.

Pearson, J., Naselaris, T., Holmes, E. A., and Kosslyn, S. M. (2015). Mental imagery: functional mechanisms and clinical applications. *Trends in cognitive sciences*, 19(10):590–602.

Posel, D. (2009). Migration: Analysis of the nids wave 1 dataset. *NIDS Discussion Paper*, no. 1.

Posel, D. (2010). Households and labour migraton in post-apartheid south africa. *Journal for Studies in Economics and Econometrics*, 34:129–141.

Posel, D. and Casale, D. (2006). Internal labour migration and household poverty in post-apartheid south africa.

Premand, P. and Stoeffler, Q. (2020). Do Cash Transfers Foster Resilience? Evidence from Rural Niger. *Policy Research Working Paper*, 9473.

Puri, J., Aladysheva, A., Iversen, V., Ghorpade, Y., and Brück, T. (2017). Can rigorous impact evaluations improve humanitarian assistance? *Journal of Development Effectiveness*, 9(4):519–542.

Ranchhod, V. (2006). The effect of the south african old age pensin on labour supply of the elderly. *South African Journal of Economics*, 74.

Ratha, D., Kim, E. J., Plaza, S., and Seshan, G. (2021). Migration and development brief 34: Resilience: Covid-19 crisis through a migration lens.

Reisberg, D., Pearson, D., and Kosslyn, S. (2003). Intuitions and introspections about imagery: The role of imagery experience in shaping an investigator's theoretical views. *Applied Cognitive Psychology*, 17:147 – 160.

Renner, F., Murphy, F., Ji, J., Manly, T., and Holmes, E. (2019). Mental imagery as a "motivational amplifier" to promote activities. *Behaviour Research and Therapy*, 114.

Rutchick, A. M., Slepian, M. L., Reyes, M. O., Pleskus, L. N., and Hershfield, H. E. (2018). Future self-continuity is associated with improved health and increases exercise behavior. *Journal of Experimental Psychology: Applied*, 24:72–80.

Schacter, D. L., Addis, D. R., and Buckner, R. L. (2007). Remembering the past to imagine the future: the prospective brain. *Nature Reviews Neuroscience*, 8(9):657–661.

Schokman, A., Bin, Y. S., Simonelli, G., Pye, J., Morris, R., Sumathipala, A., Siribaddana, S. H., Hotopf, M., Rijsdijk, F., Jayaweera, K., and Glozier, N. (2018). Agreement between subjective and objective measures of sleep duration in a low-middle income country setting. *Sleep Health*, 4(6):543–550.

Schwab, B. (2019). Comparing the Productive Effects of Cash and Food Transfers in a Crisis Setting: Evidence from a Randomised Experiment in Yemen. *Journal of Development Studies*, 55(sup1):29–54.

Siegler, R. S. (2006). *How Children Develop, Exploring Child Develop Student Media Tool Kit+ Scientific American*. Worth Pub.

Singha, M., Dong, J., Sarmah, S., You, N., Zhou, Y., Zhang, G., Doughty, R., and Xiao, X. (2020). Identifying floods and flood-affected paddy rice fields in Bangladesh based on Sentinel-1 imagery and Google Earth Engine. *ISPRS Journal of Photogrammetry and Remote Sensing*, 166:278–293.

Skoufias, E., Unar, M., and Gonzalez de Cossio, T. (2013). The poverty impacts of cash and in-kind transfers: Experimental evidence from rural Mexico. *Journal of Development Effectiveness*, 5(4):401–429.

Statistics South Africa (2013a). General Household Survey 2013.

Statistics South Africa (2013b). Mid-year population estimates.

Stöber, J. (2000). Prospective cognitions in anxiety and depression: Replication and methodological extension. *Cognition & Emotion*, 14(5):725–729.

Subbiah, A., Bildan, L., and Narasimhan, R. (2008). Background Paper on Assessment of the Economics of Early Warning Systems for Disaster Risk Reduction.

Thieken, A. H., Kreibich, H., Müller, M., and Merz, B. (2007). Coping with floods: Preparedness, response and recovery of flood-affected residents in Germany in 2002. *Hydrological Sciences Journal*, 52(5):1016–1037.

Tranchant, J. P., Gelli, A., Bliznashka, L., Diallo, A. S., Sacko, M., Assima, A., Siegel, E. H., Aurino, E., and Masset, E. (2019). The impact of food assistance on food

insecure populations during conflict: Evidence from a quasi-experiment in Mali. *World Development*, 119:185–202.

Tulving, E. (2002). Episodic memory: From mind to brain. *Annual Review of Psychology*, 53.

Ubfal, D., Arraiz, I., Beuermann, D., Frese, M., Maffioli, A., and Verch, D. (2019). The Impact of Soft-Skills Training for Entrepreneurs in Jamaica. *IZA Discussion Paper Series*, 12325.

Uddin, K., Matin, M. A., and Meyer, F. J. (2019). Operational flood mapping using multi-temporal Sentinel-1 SAR images: A case study from Bangladesh. *Remote Sensing*, 11(13):1581.

United Nations and World Bank (2018). *Pathways for Peace: Inclusive Approaches to Preventing Violent Conflict*. World Bank.

United Nations Resident Coordinator Office (2020). HCTT Monsoon Flood Humanitarian Response Plan: Monitoring Dashboard (5 September 2020). Technical report, United Nations Resident Coordinator Office, Humanitarian Coordination Task Team, Dhaka.

Victora, C. G., Christian, P., Vidaletti, L. P., Gatica-Domínguez, G., Menon, P., and Black, R. E. (2021). Revisiting maternal and child undernutrition in low-income and middle-income countries: Variable progress towards an unfinished agenda. *The Lancet*.

Von Graevenitz, G., Harhoff, D., and Weber, R. (2010). The effects of entrepreneurship education. *Journal of Economic Behavior & Organization*, 76(1):90–112.

Voors, M. J., Nillesen, E. E. M., Verwimp, P., Bulte, E. H., Lensink, R., and Soest, D. P. V. (2012). Violent Conflict and Behavior: A Field Experiment in Burundi. *American Economic Review*, 102(2):941–964.

Webb, P. (2015). How strong is our evidence for effective management of wasting? A review of systematic and other reviews. *Food and Nutrition Bulletin*, 36:S65–S71.

Webb, P., Boyd, E., de Pee, S., Lenters, L., Bloem, M., and Schultink, W. (2014). Nutrition in emergencies: Do we know what works? *Food Policy*, 49(P1):33–40.

Weingärtner, L., Pforr, T., and Wilkinson, E. (2020). The evidence base on anticipatory action.

Weingärtner, L. and Spencer, A. (2019). Analysing gaps in the humanitarian and disaster risk financing landscape.

WFP (2020). Anticipatory Action Bangladesh: Outcome Survey Report.

Woolard, I. and Leibbrandt, M. V. (2010). The evolution and impact of unconditional cash transfers in south africa.

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

World Bank (2019). *Bangladesh Poverty Assessment: Facing Old and New Frontiers in Poverty Reduction*. World Bank, Washington, DC.

Zaleskiewicz, T., Bernady, A., and Traczyk, J. (2020). Entrepreneurial risk taking is related to mental imagery: A fresh look at the old issue of entrepreneurship and risk. *Applied Psychology*, 69(4):1438–1469.

Zeman, A. Z., Dewar, M., and Della Sala, S. (2015). Lives without imagery-congenital aphantasia. *Cortex: A Journal Devoted to the Study of the Nervous System and Behavior*, 73:378–380.