

Essays on Labour Markets and Social Protection in Development Economics

Rocco Zizzamia

St Edmund Hall

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy in International Development

Trinity Term 2023

Monday 9th September, 2024

Essays on Labour Markets and Social Protection in Development Economics

Rocco Zizzamia

St Edmund Hall, University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy in International Development

This thesis is made up of three stand-alone papers. While contextually distinct, each paper is organised around an attempt to understand barriers that individuals in developing countries face in making productive use of their labour. Paper 1 is an investigation into the role that rejection during job search has on individuals' search behaviour. Forces that create disincentives to job-search create barriers to labour market entry for workers. Using a lab-in-the-field experiment implemented with young workers in South Africa, I explore whether the psychological experience of rejection may be one of these forces. I find that when exposed to (more frequent) rejection, subjects choose to trade-off expected earnings in order to avoid receiving rejection signals. If this behaviour carries over into the real labour market, this implies that frequent rejection has the effect of depressing search rates and that workers may implicitly trade-off earnings and employment opportunities to minimise exposure to rejection. Paper 2 is a study of a 'big push intervention' in rural Bangladesh. Big push interventions are poverty alleviation programmes which are designed to help ultra-poor individuals escape from poverty by delivering a large bundle of support – including productive assets and regular coaching. Big push interventions appear to work by allowing recipients to move from unproductive to more productive forms of self-employment. However, these programmes are expensive and difficult to scale, with coaching representing a major bottleneck. Using a large cluster randomised control trial, I find that variation in the amount of coaching that beneficiaries receive plays no discernible role in explaining the positive welfare effects of the programme. This is good news for scale – suggesting there is scope to increase the number of people such programmes reach without having to compromise on efficacy on the intensive margin. Paper 3 combines original qualitative data and nationally representative survey data to explore – in the South African context where the most reliable route out of poverty for most is through wage work – those paradoxical cases in which black, urban youths turn down or quit wage work. I use this paradox to explore the disincentives to certain forms of low-skill employment that South African workers face, and what this reveals about the labour market more broadly.

Supervisor: This thesis was supervised by Pramila Krishnan.

Word count: Approximately 58,000 words

Statement of authorship: Paper 1 is is single authored. Paper 2 is joint work with Stefan Dercon, Imran Matin, Ashley Pople, and Narayan Das. Paper 3 is single authored.

Acknowledgement of funding: I gratefully acknowledge funding from the Clarendon Fund, St Edmund Hall, the Department of International Development, and the Center for the Study of African Economies.

Thesis format: I have been granted permission to submit an integrated thesis.

Acknowledgements

For many people the DPhil is a lonely experience. It has been anything but that for me. I am so thankful that I have so many people to be thankful for.

This journey really started for me at SALDRU and UCT. From the moment I first walked into his office, Murray Leibbrandt treated me more like the person and the scholar I aspired to be than the person and scholar I was then. Like so many others who have grown under his leadership, he inspired me to live up to the trust he put in me. SALDRU is the home it is for me because of its people: Thank you especially to Francis Wilson, Vimal Ranchhod, Nicola Branson, Amy Jepthah, Tania Hendricks, Haajirah Esau, Ariane De Lannoy, and Pippa Green. SALDRU was and remains a home for me, and I am proud to be a Saldrupian.

During my DPhil in Oxford, the Oxford Department of International Development and the Center for the Study of African Economies have been where I have worked, learned, debated, made friends, and found support when I needed it. I am so grateful for all those who have played a role in this: especially Dominique Attala and Gary Jones at ODID, and Rose Page at CSAE. I have benefited from the CSAE Seminar series, both as a forum to share my own work and to learn about the work of my peers. Thank you to Simon Quinn and Julien Labonne for organising these. Thank you also to Doug Gollin and Cheryl Doss for convening the Quantitative Methods workshop series, from which I have also benefited over the past years. Special thanks is due to Chris Adam and others at ODID, without whose advocacy and support I would never have even started the DPhil at ODID.

I have had the privilege of working with a number of incredible organisa-

tions over the past years who have made the research in this thesis possible. I learned an enormous amount from my work with J-Pal Africa, unfortunately on a project which was derailed by COVID-19 and which did not make it into this thesis. Wim Louw deserves special thanks. I was energised and inspired by the time I spent with the BRAC Institute for Governance and Development in Dhaka, and in those few weeks learned more about development than in many months studying it in the classroom. I have also been lucky to work with a number of phenomenal research assistants through these organisations. On the Bangladesh paper, I could not have asked for a better partner than Maliha Noshin Khan. Kris Lemon, Raquel Caldiera and Zaakir Essa provided exceptional research assistance on my work with J-Pal, which through it is not included in this thesis, was an invaluable learning experience. Without Mzulungile Cabanga's excellent guidance, interpreting, and research assistance, Paper 3 of this thesis would not have been possible.

Those people who I have learned most from are those with whom I have collaborated with as co-authors and co-PIs on academic projects: For all you have taught me, and for the privilege of being partners, thank you to Ihsaan Bassier, Josh Budlender, Ronak Jain, Nilmini Herath, Patrizio Piraino, Imran Matin, Narayan Das, Ashley Pople, Simone Schotte, Murray Leibbrandt and Vimal Ranchhod. Many others have inspired me, encouraged me, been generous with their time and ideas, and fanned the flames of my curiosity. Thank you to Craig McIntosh, Anne Mager, Kai Barron, Rob Garlick, David Lam, Christian Meyer, Vimal Ranchhod, Hannah Dawson, Jonny Steinberg, William Beinhart, Jann Lay, Andy McKay, and Sam Jones.

I am very grateful to Muhammad Meki, Sandra Sequeira and Christopher Adam, who served as my examiners, and whose comments have greatly improved the quality of this thesis.

I am fortunate to have friends who in so many ways have provided support, companionship and levity over the years, the value of which is it impossible to exaggerate. Thank you to Tariq Blecher, Mischa Blecher, Emily Gammon, Zahraa Mohamed, Dave Viljoen, Dan Hartford, Josh Hartford, Josh Budlender, Ashley Pople, Simone Schotte, Kayla Brown, Gabriel Espi-Sanchez, Nick Budlender, Taylor Winchell, Josh Wen, Kyle Pratt, Gus Greenstein, Theresa Lieb, Biruk Terrefe, Tim Lefroy, Ilona Quahe, Max Nathanson, Luciano Rodriguez-Carrington, Maggie Neil, Alberto Lo Bue, Ollie Ballinger, Zac Cesaro, Giacomo Bertuzzi, Julia Ruiz Pozuelo, Max Roser, Imi Dobie, Vatsal Khandelwal, Chiara Beghe, Nisreen Khambati, Verena Wiedemann, Laura Sochas, Gabriela Smarelli, Rahpael Bradenbrink, Sanghamitra Mukherjee, Hannah Zillessen, Luke Milsom, and Giulio Schinaia. Lena Reim was my closest companion for most of my time in the PhD and I could always count on her support.

A huge part of my life over the last years, and a pressure valve and counterbalance to my academic work has been running and climbing. It is not a stretch to acknowledge that my companions in these pursuits have contributed in a major way to the output in other areas of my life. These are some of my deepest friendships. Thank you Francesco Pelizza, Gioele Passoni, Jessie Waldman, Laura Coggins, Dan Radford, Zoe Duby, Marc Fehr, Jimbo Smith, Ben Heargreaves, Jeanne Bouyat, Bella Pickett, James Potticary, Warren Handley, Lachlan Authur, Matt Seddon, Sacha Salphati, Mike Woodward, Keith Forbes, Cliff Hakimi, Mike Clark, Jabez Magomere, Matt Berrington, Alex Tridgell, Sarah Short and Ethan Wight.

To my family, Nina, Craig, Gail and Antonio: Your love and support has been constant and unwavering, a cherished comfort and a source of great happiness for me. I am who I am, and am able to do what I do, because of

you. Thank you.

I do not have the words to describe the depth of my gratitude to Stefan Dercon and Pramila Krishnan. They have been far more than supervisors and mentors to me over the past years. I looked up to Stefan before ever meeting him. The privilege of working with Stefan, and observing him work, has filled him out as a person who engages with people and with ideas with energy, curiosity, frankness and a sense of humour. He has treated me with a generosity and a trust which I aspire to be worthy of.

Pramila's support for me has extended well beyond the expectations of a supervisor – she has in so many ways and on so many occasions shown her care for me as a person first and foremost, and as her student only secondarily. And as a supervisor I could not have asked for a better match. I have always felt guided, never pressured or coerced, always respected, and encouraged and reassured when that was what I needed. Our long discussions about my research ideas were among the most stimulating moments in the last years. I hope that one day I will be for others what Pramila has been for me.

Finally, I am grateful for the participation of thousands of people whose experiences, stories, and difficulties are the threads from which this thesis is woven. There are both positive stories – of people working with amazing organisations to transform their economic lives – but there are also sadder stories – of young people playing in a labour market with dice loaded against them. Each has given their time to contribute to this research, and I am immensely thankful.

Rocco Zizzamia

Oxford, April 2023

Table of Contents

1 Introduction **1**

References 12

2 Ignorance is bliss? Rejection and discouragement in on-the-job search **16**

2.1 Introduction 17

2.2 Context 25

2.2.1 Understanding inadequate job search 25

2.2.2 Rejection avoidance as information avoidance 28

2.2.3 The South African labour market 31

2.3 Empirical approach and data 33

2.3.1 Sample 33

2.3.2 The structure of the experimental sessions 37

2.3.3 Game protocol 40

2.3.4 Top-up rounds 46

2.3.5 The real effort task 47

2.3.6 Survey 49

2.4 Treatments 50

2.5 Results 54

2.5.1 Descriptive comparisons of search behaviour across treatments 54

2.5.1.1 Treatment 1: *Stay* 54

2.5.1.2 Treatment 2: *Search* 56

2.5.1.3 Treatment 3: *SearchNoRej* 59

2.5.1.4 Treatment 4: *SearchMoreRej* 62

2.5.1.5 Comparing player strategies across games . . 64
2.5.2 Estimating the effect of rejection on job-search 66
2.5.3 Correlates of search behaviour and discouragement . . 74
2.6 A simple decision-making framework accounting for preferences for information on rejection 75
2.7 Discussion 82
2.7.1 Discussion of main results 82
2.7.2 Alternative explanations for observed behaviour 85
2.7.2.1 The salience of rejection 85
2.7.2.2 Redirecting effort in response to (perceived) signals about comparative advantage 86
2.7.2.3 Relevance of self-esteem beliefs in a quasi-lottery 88
2.7.2.4 Contamination through peer effects 90
2.7.2.5 Errors in probabilistic reasoning 91
2.8 Conclusion 92

Appendices **94**

2.A Robustness to ordering effects 94
2.B Comprehension tests 96
2.C Full list of treatment conditions 97
2.D Pre-analysis plan 101
2.E Correlates of search behaviour and discouragement 101
2.F Balance 106
2.G Self-image in the South African sample 107
References 112

3 A push which is less big but just as good? Experimental

evidence on the role of coaching within bundled ultra-poor graduation programmes **120**

3.1 Introduction . . . 121

3.2 The graduation model and the role of coaching . . . 129

3.2.1 The graduation model . . . 129

3.2.2 The effects of coaching: Possible mechanisms and hypotheses . . . 137

3.3 Experimental design . . . 140

3.3.1 Study Sites and Participants . . . 140

3.3.2 The intervention . . . 143

3.3.3 Experimental design . . . 149

3.3.4 Data . . . 153

3.4 Empirical approach . . . 155

3.5 Main results . . . 156

3.5.1 Effects on primary outcomes . . . 156

3.5.2 Putting effect sizes in context . . . 161

3.5.3 Effects on asset holdings . . . 166

3.5.4 Cost-effectiveness . . . 169

3.6 Effects of variation in coaching intensity on intermediate outcomes . . . 169

3.6.1 Hard skills . . . 170

3.6.2 Coaching as a commitment device . . . 176

3.7 Discussion and conclusion . . . 181

Appendices **184**

3.A Details on the asset package . . . 185

3.B Robustness checks . . . 185

3.B.1 Double-post selection LASSO controls . . . 185

3.B.2 Accounting for differences in Village Action Committee formation within Treatment Arm 3 . . . 186
3.B.3 Log transformation of outcome variables . . . 186
3.B.4 Winsorization at 95% . . . 187
3.C Cost benefit analysis . . . 188
3.D Evidence of treatment effects on other outcomes . . . 193
3.D.1 Does the UPGP crowd in or crowd out additional governmental/non-governmental support? . . . 193
3.D.2 Social inclusion . . . 194
3.D.3 Psycho-social outcomes . . . 198
3.D.4 Financial inclusion . . . 202
3.E Balance . . . 205
3.F Attrition . . . 207
3.G Pre-analysis plan . . . 208
3.H Construction of variables for comparison with Bandiera et al. (2017) and Banerjee et al. (2015) . . . 209
References . . . 214

4 Is employment a panacea for poverty? A mixed-methods investigation of employment decisions in South Africa 219
4.1 Introduction . . . 220
4.2 Contested terrain: the meaning of work in South Africa . . . 223
4.3 Methods and data . . . 227
4.3.1 Quantitative data . . . 227
4.3.2 Qualitative data and methods . . . 228
4.4 Poverty dynamics and employment volatility - a quantitative perspective . . . 233
4.5 A model for analysing worker employment decisions . . . 239

4.6 Is employment a panacea to poverty? Qualitative evidence . . 243
4.6.1 Aggregate patterns: labour market events and wellbeing 243
4.6.2 Low, variable wages and expensive commutes 245
4.6.3 Burden of dependents 248
4.6.4 Perception of exclusion 251
4.6.5 The role of weak outside options in job-quit decisions . 256
4.7 Conclusion 258

Appendices **262**

4.A Site selection, field research methods and post-collection analysis techniques 262
References 270

5 Conclusion **275**

References 280

List of Figures

2.1 Session structure 38
2.2 Anatomy of a round 44
2.3 Distribution of effort in the real effort task 49
2.4 Evolution of search rates in the *Stay* treatment condition . . . 56
2.5 Evolution of search rates in the *Stay* vs *Search* treatment conditions 58
2.6 Evolution of search rates in the *Search* vs *SearchNoRej* treatment conditions 61
2.7 Evolution of search rates in the *Search* vs *SearchMoreRej* treatment conditions 63
2.8 Distribution of search behaviour across treatment conditions . 66
2.9 Comparing job search behaviour between individuals with different parameter values 80
2.B.1 Main comprehension check 96
2.B.2 Supplementary comprehension check 97
2.G.1 Grit and self-efficacy: Distribution of scores 109

3.3.1 Selection of study areas and allocation of branches to treatment and control 142
3.3.2 Experimental design 151
3.6.1 Proportion of recipients retaining their asset by a given year . 178
3.6.2 Relationship between BRAC asset retention and livelihood outcomes 180

4.3.1 Mapping welfare trends over time 231
4.A.1 Sample Selection 263

4.A.2Illustration of the analysis of LHI events – Sicelo's life history. 268

List of Tables

2.3.1 Lab sample compared to National Income Dynamics Study samples . . . 36
2.4.1 Primary treatments . . . 51
2.5.1 Treatment effects on search rates, comparing *Search* to *Stay* (inverted) conditions . . . 68
2.5.2 Treatment effects on search rates, comparing *Search* to *Search-NoRej* conditions . . . 70
2.5.3 Treatment effects on search rates, comparing *Stay* (inverted) to *SearchNoRej* conditions . . . 72
2.5.4 Treatment effects on search rates, comparing *Search* to *Search-MoreRej* conditions . . . 73
2.A.1 Treatment effects of rejection on search rates: Search vs Stay (inverted), with and without sequencing fixed effects . . . 95
2.C.1 All treatments . . . 100
2.E.1 Correlates of search behaviour and discouragement . . . 105
2.F.1 Balance between sample that completed the top-up session and those that did not . . . 106
2.G.1 Labour market expectations (confidence expressed on a scale of 10): 1 month vs 4 months after starting the internship . . . 110

3.3.1 Baseline sample statistics . . . 143
3.5.1 Overall treatment effects of the BRAC UPGP with weekly, fortnightly, or fortnightly + group coaching . . . 160
3.5.2 Comparison with Bandiera et al. (2017) and Banerjee et al. (2015) . . . 165
3.5.3 Treatment effects on assets . . . 168

3.6.1 Likelihood of being confident in ability to perform various tasks 173
3.6.2 Animal husbandry best practices 175
3.A.1 Details on the asset transfer package 185
3.B.1 Overall treatment effects of the BRAC UPGP with weekly (T1), fortnightly (T2), or fortnightly + group (T3) coaching . 185
3.B.2 Overall treatment effects of the BRAC UPGP with weekly (T1), fortnightly (T2), fortnightly + group coaching with VAC (T3a) vs without VAC (T3b) 186
3.B.3 Overall treatment effects on log-transformed outcome variables 187
3.B.4 Overall treatment effects with Winsorization at the 95% level . 188
3.C.1 Roughly estimated cost-benefit analysis across three treatment arms 192
3.D.1 Number of instances household received help from government / NGO in last year 194
3.D.2 Treatment effects on social inclusion 197
3.D.3 Treatment effects on psycho-social outcomes 200
3.D.4 Treatment effects on financial inclusion 204
3.E.1 Baseline balance across all control and treatment groups . . . 206
3.F.1 Differential attrition 207

4.3.1 Sample means - Qualitative sample compared to NIDS for African adults in Cape Town (2017). 232
4.4.1 Poverty transition matrices for South Africa, 2008–2017 234
4.4.2 Number of periods employed 235
4.4.3 Changes in indicators of wellbeing co-occurring with job loss . 237
4.6.1 Incidence of labour market events and associated wellbeing effects 244

1 | Introduction

Across the developing world, economic mobility is largely determined by the terms on which individuals engage in the labour market. In middle-income countries with large formal sectors, the economic lives of the poor are largely shaped by unemployment and their exclusion from the formal labour market. In most lower-income countries where the informal sector is far larger, upward economic mobility typically means graduating from survivalist to more economically productive forms of self-employment.

In this thesis I seek to understand some of the barriers that stand in the way of individuals engaging more productively in the labour market. I focus on two countries facing very different labour market challenges – South Africa and Bangladesh. Two of the papers that make up this thesis focus on the challenges that young South Africans face in entering into employment, specifically on barriers to effective job search and on several overlooked disincentives to formal work. The remaining paper focuses on a 'graduation programme' in Bangladesh, in which very poor rural women receive large productive asset transfers and regular training with the aim of facilitating a transition away from survivalist to more productive economic livelihoods.

South Africa is an upper-middle income country characterised by a dual labour market with high levels of inequality (Leibbrandt et al., 2010). It has a strong formal sector, and a robust social protection regime (Kingdon and Knight, 2004; Satumba et al., 2017). At the same time, South Africa is facing a protracted unemployment crisis, with especially high rates of youth unemployment.

Part of the reason that the youth struggle to find jobs is that the ed-

ucation system fails to adequately signal ability to employers (Lam et al., 2011), with the result that signals from the labour market itself are especially important. However, to the extent that young workers remain 'locked out' of employment opportunities, these routes to climbing the job ladder may remain closed off. For young South Africans, finding a 'good' formal sector job is therefore often as difficult as it is important.

At the same time, in a labour market dominated by the formal sector, large firms wield substantial monopsony power (Bassier, 2021; Bassier, 2023), exacerbated by spatial mismatch (Banerjee and Sequeira, 2020). These features, combined with the fact that supply outstrips demand in low-wage sectors, conspire to erode the terms on which employment is offered to low-wage workers.

In Paper 1, I focus on a disincentive to job search which may emerge in a slack labour market such as South Africa's: The psychological experience of repeated rejection. Specifically, I ask whether workers adjust their search behaviour downward to reduce their exposure to rejection, even at the cost of forfeiting potential earnings and employment opportunities. Inadequate search might contribute to the difficulties that young workers face in entering the labour market.

This question arose from two observations in prior research I undertook over the first two years of my DPhil. In this research, I conducted a study on the impacts of a large public internship programme (which was unfortunately derailed by the COVID-19 pandemic and associated lockdowns). In this project I ran a panel study in which I followed interns from soon after they entered the programme, to five months later. I was interested in what these workers – who were in most cases working their first job – were learning about

the job and about their own traits as workers. To do so, I measured their beliefs and expectations going into the job, and asked how they updated these expectations following several months of exposure to the work environment.

I found that this sample of young workers was remarkably optimistic about their employment prospects for after the internship, and that they sustained these high levels of optimism from when they begun working to five months later. This optimism stands in stark contrast with actual employment rates for those who have completed the internship in the past: Data from the internship provider suggests that over 60% of former interns are not employed. Optimism is therefore not only persistent, but also appears unwarranted.

This lack of learning about employment prospects poses a puzzle, but also raises some concerns: If workers are over-optimistic about their employment prospects – especially if they are over-optimistic about retention in their current workplace – this may lead to complacency that depresses job search. Search may also be costly – monetarily and psychologically – and present biased workers may delay search (Field et al., 2023). At the same time, workers may derive positive utility from optimistic beliefs in the present (Brunnermeier and Parker, 2005), and may choose to actively avoid exposure to information – i.e. such as frequent rejection in job search – that would force a downward revision in these beliefs (Golman et al., 2017).

In Paper 1 I focus on this latter hypothesis: Studying the impact of rejection on job search in the real world is difficult since various factors affect search behavior, including expected pecuniary returns, search costs, and outside options. These factors may also interact with experiences of rejection, making it difficult to identify the role of rejection itself. As a

result, the effect of rejection on job search behavior in slack labor markets with high rejection rates is not well understood.

Therefore, in order to isolate a rejection effect and net out potential confounders, I study this question using a lab-in-the-field experiment with a theoretically relevant population – young workers in South Africa.

The experiment consists of a series of incentivised multi-round games in which subjects are asked to choose between performing a task which has a high expected exposure to receiving a rejection signal, and a task which has a low expected exposure to receiving a rejection signal. In both tasks, the round-by-round offer rate is calibrated to be identical. The crucial difference between the two tasks is that in the 'high rejection exposure' task – framed as 'searching for a new job' – outcomes are revealed on a round-by-round basis, while in the 'low rejection exposure' task – framed as 'focusing on your current job' – outcomes are withheld from the subject until the game is over. The probability of receiving a payoff at the end of the game is identical for both activities and is determined by chance and by performance on a real-effort task. Importantly, this fact is known by all subjects.

Across versions of the game I vary the payoffs for the two activities ('search' or 'not search'), and the amount of rejection that subjects receive if they choose to 'search'. This allows me to measure the extent to which subjects respond to rejection signals by reducing their exposure to activities with frequent rejection exposure in favour of activities with infrequent rejection exposure, even where this comes at the cost of a lower expected payoff. We can interpret this as a willingness to pay to avoid receiving information. Using the multi-round feature of the games I am also able to track how rejection avoidant behaviour evolves with past exposure to rejection.

Subjects in the game are willing to accept lower earnings to avoid rejection signals, leading to under-searching in treatment conditions where there is a higher monetary return to searching than not searching. Exposure to rejection signals decreases search rates by 18 to 25 percent compared to how much we would expect rejection indifferent players to search. Importantly, search rates drop as players who were initially induced to search by the monetary premium stop searching, with search rates eventually becoming indistinguishable from a condition where there is a monetary incentive *not* to search.

To help interpret these results, I lay out a simple theoretical framework and propose that the rejection avoidant behaviour I observe in the lab may be interpreted as a case of active information avoidance (Golman et al., 2017) in which individuals become increasingly averse to receiving information on rejection which is discordant with their initial beliefs about ability. Intuitively, past signals reduce the 'ignorability' of new signals – repeated rejections are more difficult to brush off than just one. In turn, this affects individual incentives to avoid or seek out new information. Therefore, anticipating the possibility of receiving more rejections, individuals who have already received rejection(s) and who do not want to risk a downward revision to their self-esteem beliefs will become more motivated to change their behaviour to avoid further exposure to rejection.

However, searching for and successfully finding a job is not always sufficient for enabling meaningful improvements in wellbeing: The terms on which employment is offered and how these interact with an individual's economic and social situation are also important considerations in determining whether a job has a positive net effect on infinitival well being. In Paper 3 I ask what those paradoxical cases in which South African workers turn

down or quit wage jobs reveal about the quality of entry-level jobs and about heterogeneity in individual incentives to take up formal jobs.

If workers often choose to turn down or leave wage employment, this is an important potential determinant of both unemployment and employment dynamics. Consequently, this will have profound implications for the effectiveness of strategies which hope to address unemployment through job creation. This study attempts to understand the conditions under which disadvantaged South Africans choose to turn down economic 'inclusion' in the form of wage employment. I propose that these decisions are determined by two distinct but interrelated factors. First, a decision to leave or turn down wage employment will be more likely if the benefits are not significantly greater than the costs of employment. Second, those with stronger outside options may be less tolerant of 'bad' forms of work relative to those with weaker outside options. This model of behaviour is consistent with the standard neoclassical model of labour choice, where agents make a utility maximising decision by weighing the costs and benefits of an employment option against the costs and benefits of an outside option. However, both the costs of employment and the strength of outside options are often not amenable to quantitative measurement. For instance, costs of employment might include psychological burdens while outside options may be determined by the strength of complex and unquantifiable social networks. One consequence of this is that, if workers value their psychological wellbeing over their wage employment, they may make welfare-maximising choices which lead to decreases in observed income or consumption. This may lead to misleading results in a standard quantitative analysis. Furthermore, the relatively high levels of social protection through state grants and the high rates of formal employment mean that the outside options of the unemployed are often

stronger than would be the case in other African contexts, making this an interesting case for studying this question.

I study this question using mixed-methods, combining National Income Dynamics Study (NIDS) panel data with original data from in-depth life history interviews (LHIs) and focus group discussions (FGDs) which were collected in Khayelitsha, a large black township in Cape Town.

Unsurprisingly, getting a job is on average linked to improvements in material wellbeing. However, these averages hide heterogeneity which is of special analytical interest: I find that younger workers, those with few dependants, those with alternative sources of support and those with strong social networks appear to have stronger outside options and are consequently more likely to turn down or quit 'bad' jobs. Drawing on my qualitative data, I provide a detailed investigation into some of the factors which push these workers to quit or turn down wage work, focusing on geographical frictions, the unpredictability of earnings, the psychological experience of working at the bottom of the formal job ladder in a high inequality society, and the role of re-distributive pressures within households.

I find that for workers with relatively strong outside options, wage employment is often deliberately temporary and might be taken up to smooth consumption, so that job-loss, whether deliberate or not, will have a less severe negative impact. I propose that for the unskilled urban poor, certain forms of wage work can be considered as 'survivalist', in a similar sense that certain forms of unprofitable informal self-employment are conventionally considered to be survivalist. That is, like other activities where the benefits are only marginally greater than the costs, workers consider wage employment as a 'second best' option, to be avoided when possible. Understanding

that low-skill employment may be seen by many as an aspirational compromise is imperative if efforts to combat unemployment through job creation are to be effective.

While Papers 1 and 3 deal with job search in South Africa, Paper 2 focuses on a very different context: Rural Bangladesh. Unlike South Africa, the formal sector in Bangladesh is much smaller, especially in rural areas, and the state social protection system is much thinner. Two of the most important elements in individuals' livelihood portfolios in South Africa – income from formal work (whether directly through employment or indirectly through networks) and state cash grants – are almost completely absent in rural Bangladesh. Without recourse to the formal labour market or to state social protection, economic wellbeing in rural Bangladesh is largely determined by the livelihood strategies an individual engages in in the informal economy. The productivity of these livelihood strategies in the informal sector are, in turn, determined in part by access to productive capital where those without productive capital are forced to sell their labour on unfavorable terms to those who do have productive capital (Bandiera et al., 2017).

To the extent that social protection is provided in Bangladesh, this is often done by private actors[1]. BRAC is the primary private actor in this space, and provides a range of services, from microfinance to health services. BRAC's flagship social protection programme is the Ultra-Poor Graduation Programme (UPGP), first launched in 2002 as the Targeting the Ultra-Poor programme. The UPGP is an intervention aimed at reaching the poorest women in rural areas who are not eligible for most other interventions, such as microfinance. Rather than merely providing consumption support, the

[1]For a detailed discussion on the political economy of social protection provision in Bangladesh, see (Hossain, 2017)

UPGP aims to deliver a programme that has the potential to be economically transformative for recipients: by delivering a composite bundle consisting of a large productive asset transfer, consumption support, savings facilities, and regular training/coaching, the UPGP aims to facilitate sustained escapes from poverty by breaking structural barriers to upward mobility.

Bandiera et al. (2017) show that the UPGP in Bangladesh has positive impacts on a range of wellbeing indicators, including increasing annual income, consumption, and asset accumulation, and that these effects were sustained (and even increased) after four to seven years. Importantly, they show that these poverty-reducing effects were driven in particular by the effect that the intervention had on recipients' labour supply: Access to productive assets moved women from low return occupations (maid work and casual agricultural labour) to higher return work which required capital inputs (livestock rearing). Results like these point to a poverty-trap interpretation of the effects of the programme, in which the intervention functions as a 'big push', moving poor households onto a positive economic trajectory in which they are able to sustain an escape from poverty (Balboni et al., 2022). Evaluations in other contexts reach similar conclusions (Blattman et al., 2016; Burchi and Strupat, 2018; Bedoya et al., 2019; Banerjee, Duflo, and Sharma, 2021; Banerjee, Karlan, et al., 2022; Bossuroy et al., 2022).

As promising as they are, a major draw-back of these ultra-poor graduation programmes is that they are expensive and logistically burdensome to implement. Coaching – which is traditionally delivered on a one-on-one basis at regular intervals over a period spanning up to two years – is one of the most expensive elements of many big push interventions. Understanding which elements of the programme are less cost-effective than others will be necessary in identifying where to scale back on the intensive margin in order

to scale up on the extensive margin and reach more people.

Since these interventions are really a composite bundle consisting of an asset transfer, consumption support, savings facilities, and regular training/coaching, it is also unclear what is driving impact and through what channels. For instance, is the transformation in beneficiaries' labour supply driven only by the productive asset transfer or also by the transfer of skills and other benefits through coaching? This may provide insight into what constraints are holding individuals back from using their labour more productively absent the intervention: Is it merely a lack of productive capital, or also of human capital?

We ask three questions in this third paper: First, did the 2016 intervention have lasting effects on the welfare of recipient women and households when we followed up with them six years later? Second, does changing the mode of delivery of the coaching component or reducing the amount of coaching recipients receive affect the magnitude of welfare impacts? Third, we investigate some of the potential mechanisms of action of coaching, attempting to discern whether we can detect effects of coaching within the UPGP programme on some of the intermediate outcomes through which coaching has been hypothesised to operate.

We seek to answer these questions through the evaluation of a field experiment which was carried out with approximately 8000 households in 2016 in rural Bangladesh. Three treatment arms received a version of the UPGP, with variation in the type and amount of coaching delivered, while a control arm did not receive the programme. The first treatment group received the 'full' UPGP programme – including weekly coaching. The second treatment group received the same UPGP programme, but with fortnightly instead of

weekly coaching, while the third treatment group received fortnightly coaching alongside fortnightly group-based coaching.

Comparing treatment arms to the control arm, we are able to evaluate the overall impact of the programme on a range of welfare indicators. By exploiting the randomised variation in the intensity (weekly vs fortnightly) and form (individual vs individual and group-based) of the coaching component, we are also able to study the marginal impact of these coaching variations by comparing outcomes across the three treatment groups.

We find that the UPGP programme has sizeable impacts on asset holdings, consumption, earnings, land holdings, and labour supply in higher-productivity economic activities. These impacts are sustained six years after the programme was introduced, despite the impact of COVID-19 from 2020 to 2022. The magnitude of the effects in our study are comparable to the three-year effects estimated by Banerjee, Duflo, Goldberg, et al. (2015) in their five-country pilot study, and to the four-year effects estimated by Bandiera et al. (2017).

However, when comparing effects *across* treatment conditions, we find almost no evidence that variation in the coaching component of the UPGP has an effect on welfare outcomes. Treatment effect point estimates are similar across the three treatment arms, and – apart from a small but significant difference in consumption between the 'weekly coaching' and the 'fortnightly coaching' treatment groups – we fail to reject the null of no differential impact across treatment groups for the full range of welfare indicators. Overall, our findings strongly suggest that the marginal impact of coaching is small and does not justify the substantial costs required. To the extent that in-person coaching is a bottleneck to bringing graduation programmes to scale,

reducing the quantity of coaching is likely to be an important tool.

We also provide suggestive evidence ruling out a number of 'classic' proposed mechanisms of the effectiveness of coaching – including that coaching functions as a transfer of human capital through skills development. Rather, we argue that, if coaching works, the most likely explanation for its effectiveness is the hypothesis that coaching increases programme compliance and provides a link between recipients and the implementing organisation.

Paper 1 – on rejection and job search in South Africa – is single authored. Paper 2 – on the role of coaching in graduation programmes – is authored with Ashley Pople, Narayan Das, Imran Matin and Stefan Dercon. Paper 3 is single authored, and is built off my MPhil thesis, submitted in 2018.

References

[1] Geeta Gandhi Kingdon and John Knight. "Unemployment in South Africa: The nature of the beast". In: *World development* 32.3 (2004), pp. 391–408.

[2] Markus K Brunnermeier and Jonathan A Parker. "Optimal expectations". In: *American Economic Review* 95.4 (2005), pp. 1092–1118.

[3] Murray Leibbrandt, Ingrid Woolard, Arden Finn, and Jonathan Argent. *Trends in South African income distribution and poverty since the fall of apartheid*. Tech. rep. 2010.

[4] David Lam, Cally Ardington, and Murray Leibbrandt. "Schooling as a lottery: Racial differences in school advancement in urban South Africa". In: *Journal of development economics* 95.2 (2011), pp. 121–136.

[5] Abhijit Banerjee, Esther Duflo, Nathanael Goldberg, Dean Karlan, Robert Osei, William Parienté, Jeremy Shapiro, Bram Thuysbaert, and Christopher Udry. "A multifaceted program causes lasting progress for the very poor: Evidence from six countries". In: *Science* 348.6236 (2015), p. 1260799. DOI: 10.1126/science.1260799.

[6] Christopher Blattman, Eric P. Green, Julian Jamison, M. Christian Lehmann, and Jeannie Annan. "The Returns to Microenterprise Support among the Ultrapoor: A Field Experiment in Postwar Uganda". In: *American Economic Journal: Applied Economics* 8.2 (Apr. 2016), pp. 35–64. DOI: 10.1257/app.20150023. URL: https://www.aeaweb.org/articles?id=10.1257/app.20150023.

[7] Oriana Bandiera, Robin Burgess, Narayan Das, Selim Gulesci, Imran Rasul, and Munshi Sulaiman. "Labor markets and poverty in village economies". In: *The Quarterly Journal of Economics* 132.2 (2017), pp. 811–870.

[8] Russell Golman, David Hagmann, and George Loewenstein. "Information avoidance". In: *Journal of economic literature* 55.1 (2017), pp. 96–135.

[9] Naomi Hossain. *The aid lab: Understanding Bangladesh's unexpected success*. Oxford University Press, 2017.

[10] Takunda Satumba, Amiena Bayat, and Seeraj Mohamed. "The impact of social grants on poverty reduction in South Africa". In: *Journal of Economics* 8.1 (2017), pp. 33–49.

[11] Francesco Burchi and Christoph Strupat. "Unbundling the impacts of economic empowerment programmes: evidence from Malawi". In: (2018).

[12] Guadalupe Bedoya, Aidan Coville, Johannes Haushofer, Mohammad Isaqzadeh, and Jeremy P Shapiro. *No Household Left Behind: Afghanistan Targeting the Ultra Poor Impact Evaluation*. Working Paper 25981. National Bureau of Economic Research, June 2019. DOI: 10.3386/w25981.

[13] Abhijit Banerjee and Sandra Sequeira. "Spatial mismatches and imperfect information in the job search". In: (2020).

[14] Abhijit Banerjee, Esther Duflo, and Garima Sharma. "Long-Term Effects of the Targeting the Ultra Poor Program". In: *American Economic Review: Insights* 3.4 (Dec. 2021), pp. 471–86. DOI: 10.1257/aeri.20200667. URL: https://www.aeaweb.org/articles?id=10.1257/aeri.20200667.

[15] Ihsaan Bassier. *The impact of centralized bargaining on spillovers and the wage structure in monopsonistic labour markets*. Tech. rep. WIDER Working Paper, 2021.

[16] Clare Balboni, Oriana Bandiera, Robin Burgess, Maitreesh Ghatak, and Anton Heil. "Why do people stay poor?" In: *The Quarterly Journal of Economics* 137.2 (2022), pp. 785–844.

[17] Abhijit Banerjee, Dean Karlan, Robert Osei, Hannah Trachtman, and Christopher Udry. "Unpacking a multi-faceted program to build sustainable income for the very poor". In: *Journal of Development Economics* 155 (2022), p. 102781. ISSN: 0304-3878. DOI: https://doi.org/10.1016/j.jdeveco.2021.102781.

[18] Thomas Bossuroy, Markus Goldstein, Bassirou Karimou, Dean Karlan, Harounan Kazianga, William Parienté, Patrick Premand, Catherine C Thomas, Christopher Udry, Julia Vaillant, et al. "Tackling psychosocial and capital constraints to alleviate poverty". In: *Nature* 605.7909 (2022), pp. 291–297.

[19] Ihsaan Bassier. "Firms and inequality when unemployment is high". In: *Journal of Development Economics* 161 (2023), p. 103029.

[20] Erica Field, Robert Garlick, Nivedhitha Subramanian, and Kate Vyborny. "Why Don't Jobseekers Search More? Barriers and Returns to Search on a Job Matching Platform". Unpublished manuscript. Feb. 2023.

2 | Ignorance is bliss? Rejection and discouragement in on-the-job search

Abstract

I investigate the effect that experiencing repeated rejection has on on-the-job search. Using a lab-in-the-field experiment implemented with young workers in South Africa, I repeatedly ask subjects to choose between a high-return activity with frequent exposure to rejection signals and a lower-return activity with less frequent exposure to rejection signals. I ask whether subjects take costly action to avoid exposure to information on rejection by choosing the lower-return, lower-rejection activity. To do so, I experimentally vary both the rewards and the amount of rejection that subjects experience when choosing between the two tasks, holding other salient drivers of search behaviour constant, including eliminating the ability of players to learn about returns to search through experiencing rejection. I find that when exposed to (more frequent) rejection, subjects choose to trade-off expected earnings in order to avoid receiving rejection signals. I interpret these results as an example of active information avoidance.

Acknowledgements: I am very grateful to Christopher Adam, Kai Barron, Ihsaan Bassier, Joshua Budlender, Gauri Chandra, Stefan Dercon, Nilmini Herath, Vatsal Khandelwal, Lukas Lehner, Craig McIntosh, Muhammad Meki, Sandra Sequeira, Christian Meyer, Kate Orkin, Patrizio Piraino, and participants at the Jobs and Development Conference, the Oxford CSAE Workshop, Oxford Firms and Labour Workshop and Oxford Department of International Development DPhil Seminar for helpful comments and discussions. I am especially thankful to my supervisor Pramila Krishnan, whose unwavering support has made this paper and much of my other work over the last years possible. This research was possible thanks to partnerships with CapaCiTi, the Youth Employment Service, and Younglings Africa, with special thanks to Fiona Tabraham. I gratefully acknowledge funding for this research from the Center for the Study of African Economies, Oxford. Any errors are my own.

2.1 Introduction

Being rejected hurts. Certain activities, such as those involving searching for a job, for a graduate school, or for a partner, come with the potential for experiencing frequent rejection. At the same time, there may be high expected returns to engaging in these activities – finding a good job, going to a top graduate school, or finding a good partner. However, the experience of (repeated) rejection may impose a psychological cost high enough to create disincentives to engaging in such activities, making these positive outcomes less likely.

In the context of the labour market, the effect that frequent rejection has on job search is a hard problem to study: Search behaviour is determined by expected returns to search, by the costliness of the search process, and by outside options to search, among other things. All of these factors may co-vary in unknown ways with experiences of rejection in job-search, getting in the way of a clear identification of the role of rejection in job search. The psychological cost of rejection itself may vary depending on, for instance, the effort one puts into search, the expectations of a positive outcome, and past experiences of rejection. Because of these difficulties, we do not have a good understanding of the role that rejection qua rejection in job search has on search behaviour, especially in slack labour markets where rejection rates in job search are highest.

In this paper I ask whether workers are willing to pay – by forfeiting potential earnings – to avoid receiving information on rejection by searching less than they would search if they were rejection insensitive. In other words, do workers adjust their search behaviour downward to reduce their exposure to

rejection, even when they know that the expected returns to search are high? I also investigate the *dynamics* of this question: how do preferences to avoid receiving information on rejection evolve with past exposure to rejection?[1]

To answer these questions I need a means of isolating a rejection effect and netting out potential confounders. To do so, I develop and implement a lab-in-the-field experiment with a theoretically relevant population – young workers in South Africa. The experiment was implemented in Cape Town in August and September 2022, collecting data from 203 subjects using a within-subjects design.

By using a lab-in-the-field approach I wield a high degree of control over potential confounders, allowing me to isolate the effect of exposure to rejection on job search. The experiment consists of a series of incentivised multi-round games in which subjects are asked to choose between performing a task which has a high expected exposure to receiving a rejection signal, and a task which has a low expected exposure to receiving a rejection signal. In both tasks, the round-by-round offer rate is calibrated to be approximately four percent[2], yielding a round-by-round rejection rate of approximately 96 percent and a cumulative rejection rate of approximately 72 percent for each game. The crucial difference between the two tasks is that in the 'high rejection exposure' task, outcomes are revealed on a round-by-round basis, while in the 'low rejection exposure' task outcomes are withheld from the

[1] I frame rejection avoidance as a form of active information avoidance (Golman et al., 2017), so the precise sense in which I understand rejection avoidance is as information-on-rejection avoidance. When I use terms such as 'rejection avoidance' or 'rejection effects', they should be understood as shorthand for this more precise sense in which I understand rejection avoidance.

[2] Offer probability is sensitive to performance in a real-effort task. The game is calibrated so that the round-by-round offer rate is four percent *at median effort* in the real effort task.

subject until the game is over.[3] The high rejection signal activity is framed as 'searching for a new job', and the low-rejection signal activity is framed as 'focusing on your current job'. The probability of receiving a payoff at the end of the game is identical for both activities and is determined by chance and by performance on a real-effort task. Importantly, this fact is known by all subjects.

Thus, putting aside differences in payoffs and framing, the only distinguishing feature between 'searching for a new job' and 'focusing on your current job' is that information on outcomes is provided frequently in the former case, and infrequently and in a 'bundled' way in the latter. This allows us to interpret the game and results as a case of information avoidance or seeking behaviour.

Across versions of the game I vary the payoffs for the two activities ('search' or 'not search'), and the amount of rejection that subjects receive if they choose to 'search'. This allows me to measure the extent to which subjects respond to rejection signals by reducing their exposure to activities with frequent rejection exposure in favour of activities with infrequent rejection exposure, even where this comes at the cost of a lower expected payoff. We can interpret this as a willingness to pay to avoid receiving information. Using the multi-round feature of the games I am also able to track how rejection avoidant behaviour evolves with past exposure to rejection.

I find that subjects in the game are willing to reduce their expected earnings in order to avoid being exposed to rejection signals. In the presence of frequent rejection exposure, subjects systematically under-search (relative to hypothetical rejection indifferent players and observed players exposed

[3] In the 'low rejection exposure' task outcomes are still calculated on a round-by-round basis, and are simply not revealed to players until all rounds have been completed.

to less frequent rejection) in treatment conditions where there is a higher monetary return to searching than not searching. In the laboratory setting, exposure to rejection signals decreases search rates by between 18 and 25 percent compared to how much we would expect rejection indifferent players to search.

The dynamics of rejection avoidance are especially striking: In conditions where there is a monetary incentive to search (i.e. to choose the activity which exposes subjects to frequent rejection signals), players start out choosing to search as often as we would expect rejection-indifferent players to choose to search. That is, prior to receiving rejection messages, players' decisions are driven only by expected pecuniary returns and not by rejection avoidance. However, search rates drop as players who were initially induced to search by the monetary premium stop searching. After only five rounds, search rates drop to the point where they are indistinguishable from search rates in a condition in which there are monetary incentives *not* to search. That is, exposure to rejection signals when searching more than wipes out the monetary incentive to search, with search rates converging with search rates where there is a monetary incentive *not* to search.

However, since decisions to 'search' or 'focus on your current job' are framed using labour market language, it is possible that these findings are driven by framing effects. I investigate this by varying the amount of rejection while keeping framing constant, and find that the headline findings are robust to netting out framing effects. By varying the amount of rejection signals that players are exposed to when they decide to 'search', I isolate the effect of rejection exposure on search rates while holding the framing of the decision constant. This is important because players may carry norms and behaviours about search behaviour 'into the lab' from outside, potentially biasing results.

In treatment conditions where players are not exposed to rejection even when choosing to 'search', I find that search rates closely approximate search rates of a hypothesised rejection-indifferent player, showing that framing effects are not driving results.

While search rates drop after the first few rejections, I also show that this discouragement effect flattens out with subsequent rejections. That is, most players who are discouraged from searching are discouraged after the first few rejections, suggesting that effects of rejection on search appears to be diminishing in the number of past rejections.

Identification of these rejection effects require me to assume that I succeed in shutting down a 'learning channel' in the game. Specifically, if players believe that they are acquiring instrumentally valuable information through rejection – for instance, that the expected returns to effort in the high rejection-exposure 'search' task are lower than returns to effort in the low rejection-exposure 'stay' task – they may then be driven by this 'learning' rather than by rejection aversion to switch strategies. To shut down this learning channel from driving behaviour, I explicitly inform subjects about the returns to effort in both 'search' and 'focus on your current job' tasks, designing the game in such a way that subjects are made aware that the relationship between effort and the likelihood of receiving an offer is identical in both conditions. I also test subjects' understanding by using comprehension checks before each game.

To help interpret these results, I lay out a simple theoretical framework and propose that the rejection avoidant behaviour I observe in the lab may be interpreted as a case of active information avoidance (Golman et al., 2017; Andries and Haddad, 2020) in which individuals become increasingly averse

to receiving information on rejection which is discordant with their initial beliefs about ability. Intuitively, past signals reduce the 'ignorability' of new signals – repeated rejections are more difficult to brush off than just one. In turn, this affects individual incentives to avoid or seek out new information. Therefore, anticipating the possibility of receiving more rejections, individuals who have already received rejection(s) and who do not want to risk a downward revision to their self-esteem beliefs will become more motivated to change their behaviour to avoid further exposure to rejection.

This paper contributes to a large and growing literature on active information avoidance in economics and behavioural science (see Golman et al. (2017) for an review). Research on the topic has shown that people are willing to incur a cost to avoid receiving information that would threaten a valued aspect of their self-image or damage their self-esteem (Dana et al., 2007; Andreoni et al., 2017; Andries and Haddad, 2020; Momsen and Ohndorf, 2020; Kőszegi, Loewenstein, et al., 2021). The behavioural game I implement can be framed as an investigation into active and costly information avoidance, where rejection is merely a form of information, and choices *not* to search are costly in terms of expected earnings. While Andolfatto et al. (2009) and Kőszegi, Loewenstein, et al. (2021) have developed theoretical models which suggest that information avoidance may disincentivise job search, to my knowledge this paper is the first to experimentally test this prediction. In addition, current work on willful information avoidance largely approaches the issue as a static phenomenon, assuming that people already know their preferences for seeking or avoiding information (Dana et al. (2007) and Andreoni et al. (2017), inter alia). I add to this literature by approaching willful and costly information avoidance as a dynamic phenomenon in which information avoidant behaviour evolves as people, through exposure

to information, learn about and/or change their preferences for acquiring or avoiding information.

By studying the role of rejection in job search behaviour, I also contribute to the literature on job search, especially literature on discouragement in job search (Rosholm and Toomet, 2005; Bjørnstad, 2006; Burns et al., 2010; Wheeler et al., 2022) and recent work on determinants and consequences of inadequate on-the-job search (Spinnewijn, 2015; Abel et al., 2019; Jäger et al., 2021; Field et al., 2023). In slack labour markets rejection may be an important determinant of job search, with consequences for labour market turnover, especially among young workers (Kerr, 2018; Zizzamia and Ranchhod, 2019; Donovan et al., 2020). Most work on discouragement in job search focuses on unemployed workers, whose outside options are relatively poor – including research which has linked labour market optimism to under-searching (Spinnewijn, 2015). In the case of discouragement in on-the-job search, in which outside options are stronger, there is more scope for rejection avoidance to consequentially determine search behaviour.

A few potential limitations of the paper bear noting: The core results of this paper use a within-subjects design in which most participants receive all treatments, in an order which is randomised across experimental sessions. Randomising order will, in principle, eliminate ordering effects. Unfortunately, because of adaptations to the game design during implementation, the order with which subjects received different treatments was not randomised for all treatments. Nevertheless, for those treatments where order was randomised, results are robust to including sequencing fixed effects in the regression analysis. Results may also be biased if subjects make systematic errors in reasoning about probabilities, such as hot-hand or gamblers fallacies. I measure individual propensity to fallacious probabilistic reasoning,

and provide evidence that it is unlikely that such errors are driving results.

The paper sets out to answer a very narrowly defined question. This narrow scope forces a trade-off in which other important determinants of job-search behaviour are, by construction, set aside. For instance, this study does not consider game-theoretic determinants of search behaviour, or scenarios in which returns to search are correlated with underlying task-specific ability. In reality, as studied using models of occupational choice, (especially young) job-applicants may use the job-search process to learn about how their expected returns to effort differ across occupations. In contexts in which the distribution of returns to ability are right-skewed relative to the distribution of ability itself (Rosen, 1981; MacDonald, 1988), young job-seekers may actively expose themselves to high-rejection in search of high-returns. The present study intentionally does not consider such scenarios, focusing instead on the behavioural questions of how individual rejection aversion affects search behaviour.

In this paper I frame the discussion in terms of job search. However, this framework could be extended to any domain in which individuals face choices between high-rejection and low-rejection activities and where information on rejection carries a psychological cost that might affect behaviour. Of particular interest are classes of activities in which choices are made between high-rejection frequency, high-reward and low-rejection frequency, low-reward activities. Dating and university applications may fall into this category. For instance, the potential returns to being accepted into a top-ranked university may be very high, but applications to these universities carry with them a high potential for rejection. Rejection aversion may create disincentives to applying to top schools in favour of less competitive schools, with real welfare consequences for the individual. In such cases a similar framework as the one

used in this paper might be usefully applied.

The paper proceeds as follows: Section 2 provides empirical and theoretical context. Section 3 describes the protocol of the laboratory experiment in some detail. Section 4 describes the four treatment conditions. Section 5 presents the main results, starting with a descriptive sketch of search behaviour in the laboratory and proceeding to the causal estimation of exposure to rejection signals on search rates. In Section 6 I propose a simple theoretical framework through which I interpret the results of the game. In Section 7 I discuss the main results in the context of the theoretical framework I introduce in the previous section, and discuss the potential role of confounding behavioural explanations for the results I observe. Section 7 concludes.

2.2 Context

2.2.1 Understanding inadequate job search

Workers on fixed-term contracts face a difficult balancing act: How to optimally allocate time and effort to their current job versus searching for new jobs, in a context of uncertain returns to both. To the extent that there are diminishing marginal returns to effort in both activities, in many cases workers might find it optimal to allocate at least some time and effort to job search.

Inadequate on-the-job search may carry substantial individual costs. It is generally accepted that there is 'negative duration dependence' in unemployment spells (Kroft et al., 2013; Eriksson and Rooth, 2014; Mueller et al., 2021).[4] This fact helps explain the observation that job search is, by some

[4]Kroft et al. (2013) notes that duration dependence may be stronger in tight labour

estimates, up to three times more effective at the same level of effort for the employed than for the unemployed, and that on-the-job search results in higher quality offers than search while unemployed (Blau and Robins, 1990; Faberman et al., 2017). Rosholm and Toomet (2005) show that negative duration dependence interacts with motivation to search, since as an unemployment spell increases, the returns to search decline, leading to discouragement and labour market exit. If workers search from a position of relative 'strength' while on the job, this may have positive and lasting returns in the labour market and may mean workers are able to avoid potential unemployment traps.

To the extent that discouragement in job search has been studied in economics, this has tended to focus on discouragement among *unemployed* workers (Rosholm and Toomet, 2005; Bjørnstad, 2006; Burns et al., 2010; Berg and Uhlendorff, 2018; Mueller et al., 2021). However, discouragement may occur also for those who are searching on-the-job, with implications for employment-to-(un)employment transitions rather than unemployment duration.

If workers systematically under-search on the job, this may also have implications for monopsony power. Jäger et al. (2021) find that low-wage German workers anchor their wage expectations to their current wage, and that they consequently systematically underestimate their outside options. These beliefs lead to inadequate on-the-job search among low-wage workers. Employers are able to leverage this labour market behaviour to increase monopsony power, giving rise to depressed wages for low-wage workers and labour market segmentation. Other behavioural biases which lead to under-

markets, where unemployment duration may be a more informative signal of unobserved productivity.

searching – such as inadequate search because of rejection aversion – may have similar implications: if workers under-invest in their outside options, the terms and wages that employers can offer to keep workers on may be lower than they otherwise might be.

The amount that workers choose to search will be determined by their beliefs about the relationship between search effort and search outcomes. This paper was partly inspired by prior research I conducted among group of young workers in South Africa who were all registered as interns on a national internship programme, whom I surveyed about their labour market expectations. I found that youth from this population were (over)optimistic and (over)confident about their labour market prospects, systematically overestimating the likelihood that they would find employment after their internship ended – comparing their expectations to actual employment rates observed in the same sample (Appendix 2.G discusses these findings in more detail).

Over-optimism appears to be common in similar samples: Banerjee and Sequeira (2020) find that young South African job-seekers expect to earn almost twice the median salary of individuals with similar profiles, driven by over-optimism about finding high-wage jobs. Jones and Santos (2022) find that young Mozambican job-seekers also over-estimate expected earnings by a similar magnitude. Mueller et al. (2021) show that American workers have optimistically biased beliefs about the probability of finding jobs. While this optimism may have positive utility in the present (Brunnermeier and Parker, 2005) and may help sustain motivation (Benabou and Tirole, 2002), it may also lead to complacency and under-investment in job search. In addition, to the extent that this optimism may be 'fragile', it may also create incentives to avoid exposure to information that may force a downward revision in beliefs (Kőszegi, Loewenstein, et al., 2021).

While researchers can measure how much workers actually do search, what it would mean to say that a worker is searching 'too little' is a challenging question to answer, both conceptually and empirically. Individuals choose a level of job search which optimises not only for expected earnings, but also a host of other non-pecuniary sources of value. However, even assuming that individuals only optimise for expected earnings when choosing how much to search, understanding whether they search 'enough' depends also on the expected returns to their outside options to search, and about the relative costs incurred in search versus pursuing these outside options. Answering this question empirically in the field would therefore be a formidable challenge. The approach taken in this paper is to use the laboratory environment to exert control over the pecuniary returns to search and the outside options to search, as well as the effort cost of choosing one over the other. This allows me to answer the precisely defined question of whether the amount of search that individuals choose when exposed to information on rejection is lower than they would choose without exposure to this information.

2.2.2 Rejection avoidance as information avoidance

The rejection avoidant behaviour I study in this paper ought to be understood as an example of 'active information avoidance'. Golman et al. (2017) define active information avoidance as those cases in which individuals fail to acquire information even when *a)* they are aware that the information is available, and *b)* the individual has free access to the information. Active information avoidance has been well studied in various domains in economics and behavioural science, including in stock market investors displaying an 'ostrich effect' when markets are down (Karlsson et al., 2009), individuals avoiding potentially useful medical testing (Oster et al., 2013; Ganguly and

Tasoff, 2017), consumers avoiding information on the environmental impact of their choices (Momsen and Ohndorf, 2020), and altruistically-minded people being willing to incur costs to avoid being asked to donate to charity, thereby '[avoiding] empathic stimulation to regulate their giving and guilt' (Andreoni et al. (2017) p. 625). Dana et al. (2007) show that in a dictator game, players prefer to remain ignorant of the effects of their choices on another player.

One explanation for this information avoidance is that people strategically avoid being exposed to information where receiving this information would force them into a trade-off between taking costly action or revising a belief from which they derive positive utility. Both decisions incur a cost – in the one case monetary, and in the other psychological. Relevant beliefs may include those which are relevant to one's self-concept, such as self-esteem.[5]

Rejection in job search, like receiving an unfortunate medical diagnosis, is information which is both potentially useful, and which is unpleasant to receive. Rejection may contain instrumentally useful signals – for instance, on comparative advantage, match quality, and returns to certain search strategies – which may improve the efficacy of job search. However, the information contained in rejection may also be dissonant with beliefs from which individuals derive positive utility: for instance, these might include beliefs that individuals have of themselves as worthy, competent and in control of the outcomes they aspire to. To the extent that individuals derive utility from these beliefs, in a way that is analogous to medical screening, it is an empirical question whether, through rejection, the hedonic disutility of a possible

[5]Similar behavioural distortions aimed at maintaining a valued self-image have also been shown to occur ex post (Kőszegi, Loewenstein, et al., 2021). For instance, people have been shown to distort memories which are dissonant with a valued view they have of themselves (Huffman et al., 2022).

downward revision is greater than the utility of the potential instrumental value of the information. If the anticipated disutility of a downward revision of beliefs is great enough, we might expect individuals to strategically avoid exposure to rejection, even if this information avoidance comes at a material cost.[6]

We can think of information avoidance in the present context as belonging to a broader class of behaviours characterised by reference dependent preferences (Kőszegi and Rabin, 2006) and reference dependent risk attitudes (Kőszegi and Rabin, 2007). In the context of rejection, an individual's initial beliefs about ability serve as the reference point, and individuals are loss-averse around that reference point. In particular, we can understand rejection aversion as a form of 'disappointment aversion' (Gul, 1991; Gill and Prowse, 2012). Models of disappointment aversion take as references (for reference dependent preferences) an individual's expected payoff or outcome, where again 'agents are loss averse around their expected payoff, so losses relative to this expectation are more painful than equal-sized gains are pleasurable' (Gill and Prowse, 2012). Disappointment aversion provides a closely related behavioural explanation for behaviour in which individuals trade-off expected returns in order to reduce exposure to disappointment (Gul, 1991), or equivalently invest less effort in activities in which disappointment is more likely (Gill and Prowse, 2012).

[6]Repeated rejection may also make the information on rejection more *salient* than the explicit goal of the game – to maximise earnings. The salience of the rejection stimulus may interfere with the income-maximising goal, leading to suboptimal decision-making (Bordalo et al., 2022). It is possible that salience may be at play in this context, but, as I argue later, is not incompatible with the information-avoidance story.

2.2.3 The South African labour market

South Africa is experiencing a protracted unemployment crisis. As of the second quarter of 2022, the unemployment rate was estimated to be 35,3 percent – the highest since Statistics South Africa began collecting quarterly labour force survey data in 2008 (Statistics South Africa, 2022). If the expanded definition of unemployment is used, which includes 'discouraged work-seekers', the rate is 46,2 percent. Unemployment disproportionately affects Black South Africans (39,1 percent), women (38,2 percent) and youth, (66,5 percent for those between 18-24, and 43,5 percent for those between 25-34).

Using panel data, researchers estimate that between a quarter and a third of South Africans are stuck in chronic unemployment, with little chance of finding stable employment (Ingle and Mlatsheni, 2017; Zizzamia and Ranchhod, 2019). At the same time, South Africa has a very high rate of employment volatility, with many among the employed losing their jobs over time, and some among the unemployed moving into employment, even if temporarily (Banerjee, Galiani, et al., 2008; Kerr, 2018). Young workers' position in the labour market is especially precarious: Jobless and hence unable develop post-schooling human capital through work-based-learning, and struggling to signal the human capital they already possess to prospective employers, young workers are vulnerable to becoming trapped in a state of chronic joblessness.

It is in the context of these challenges that fixed-term employment opportunities, including internship-style programmes, hold some promise. From the firm's perspective and in the context of poor signals of worker quality, internships may be attractive as a regulatory mechanism to minimise risk

(through lowering dismissal costs) when hiring workers. This may allow firms to make more hires, especially of workers with noisier signals of human capital. From the worker side, simply getting a foot in the door may substantially alleviate several constraints, through the work-based learning led accumulation of human capital, networking effects and the ability to provide positive signals to future employers.

For workers, these short-term opportunities are stepping-stones, rather than the final destination. In order to sustain an upward career trajectory, workers need to proactively seek employment opportunities for when their current contract ends. While some workers may climb the job ladder at their current place of work – perhaps receiving a contract extension or the offer of a permanent position – many others will need to find new jobs, which will require on-the-job search if workers aim to avoid falling into unemployment.

How much individuals choose to search will be determined by the costs involved in searching and the expected benefits to search. The expected benefits to searching are evaluated against the expected benefits to not searching. For instance, if workers are very optimistic about their probability of climbing the job ladder within their current jobs, they may elect to search less than they would otherwise. In addition, search may be costly – in terms of time, money, and psychological costs. At present, little is understood about the determinants of job-search in these environments. Garlick et al. (2023) study how mistaken perceptions about job seekers' comparative advantage leads to misdirected search in South Africa – though this focuses on the quality rather than the quantity of search. Field et al. (2023) use a field experiment on a job matching platform in Pakistan to study how behavioural mechanisms – in this case present bias – affect the amount that people search. They find that alleviating these behavioural constraints to search increases the number

of applications submitted.

2.3 Empirical approach and data

2.3.1 Sample

The experiment was conducted in Cape Town, South Africa, in August and September 2022. Experimental sessions were hosted in-person in computer laboratories on the campus of my implementing partner (CapaCiTi). CapaCiTi is a digital skills training center which hosts young South Africans from low socio-economic status backgrounds on three to six month paid 'learnerships'. In these learnerships, participating youth are trained on a range of digital skills, such as the Microsoft Office suit and other more specialist applications, including accounting software and basic coding. Participants also receive some 'soft skills' training – for instance, how to approach interviewing for jobs and how to use LinkedIn. Because of the nature of the training, CapaCiTi attracts (and selects) candidates who are young, urban and who have some basic technical competence. Often, these candidates have some prior post-schooling education and training but are seeking ways to differentiate themselves in the labour market from their peers, with whom they are competing for jobs.

I deliberately selected this sample for its relevance to the research question: The participants I worked with were mostly employed on a short-term contract (in the form of a CapaCiTi learnership), and faced uncertain prospects in terms of finding a job or or being retained or 'placed' after their learnership. They were therefore grappling, in their own lives, with the problem being played in the game: how to allocate effort between focusing

on their current job in the hope of increasing their chances of being retained or placed, or to searching for a new job.[7] In total 203 subjects participated in the experiment. Of these subjects, 158 were drawn from the learnership programme at CapaCiTi and were thus in a form of temporary employment. The remainder were unemployed, and were drawn from CapaCiTi's alumni database or the networks of the employed subjects.

Table 2.3.1 presents basic descriptive statistics on the laboratory sample. Table 2.3.1 also puts the sample in the broader context of the South African labour market by including averages for a relevantly similar sample (column 2) and for national adult averages (column 3), using 2017 data from the National Income Dynamics Study – a large nationally representative survey.

The CapaCiTi sample is young – all are under 30 years old, and the average age is 23.5. 81 percent of the sample is black African, with the remainder being coloured[8]. Women make up slightly more than half of the sample. Most participants have some prior work experience (11 months on average) and have some post-schooling training. Among the 158 sample members who were employed at the time of their participation, wages were on average below the legal minimum wage – this was possible because participants receive a 'stipend' for training rather than a wage. This sample is typical of a particular class of urban youth: Educated enough to aspire to stable white collar work,but struggling to develop and signal their human capital so as to make this possible, with the result that they cycle between states of precarious short-term employment and unemployment.

Column 2 puts the lab sample into perspective. For comparability with

[7]It is worth noting that, perhaps for this reason, participants had little trouble understanding the set-up of the game or the choices being presented.

[8]Coloured is an official racial group in South Africa and the label has no derogatory connotations in this context.

the lab sample, the sample in Column 2 is restricted to urban non-white adults under the age of 28 who have at least completed secondary school. In terms of demographics, the lab sample is slightly less black (likely due to its location in Cape Town, which has a larger coloured population) and is substantially more educated, with approximately double the rate post-schooling training. In the labour market the two samples are fairly distinct, though this simply reflects the fact that selection into the lab sample was undertaken through a learnership programme. This explains the higher employment rate (78 percent compared to 51 percent), and the substantially lower earnings conditional on being employed (R3,436 compared to R10,292).

Column 3 reports averages for a nationally representative sample of the South African working age population (18-60). While the full NIDS sample is (by definition) less highly educated and less urban than column 2, it is also older and includes the white population – forces which pull in opposite directions in terms of aggregate well-being. Employment variables are fairly similar for columns 2 and 3, with the exception that earnings are slightly higher for the 'lab-comparable' sample. Taking a dynamic view, the 'lab comparable' sample has higher rates of chronic unemployment (42 percent of the unemployed were unemployed in both 2014 and 2017, compared to 35 percent for the full sample), and lower rates of stable employment (27 percent compared to 40 percent). However, as would be expected for a younger sample, rates of job finding (in many cases presumably being labour market entry) are higher (23 percent vs 14 percent). Economically, the 'lab comparable' sample live in smaller and economically better-off households than the full sample.

Table 2.3.1: Lab sample compared to National Income Dynamics Study samples

	Lab sample	NIDS sample (young, urban, educated)	NIDS sample (all adults)
Demographics			
Black African	0.81	0.88	0.88
	(0.39)	(0.32)	(0.33)
Woman	0.59	0.54	0.50
	(0.49)	(0.50)	(0.50)
Age	23.42	24.00	34.49
	(2.69)	(3.14)	(11.13)
Has some tertiary education	0.67	0.35	0.17
	(0.47)	(0.48)	(0.37)
Labour market			
Employed	0.78	0.51	0.54
	(0.42)	(0.50)	(0.50)
Not economically active	0.14	0.33	0.31
	(0.35)	(0.47)	(0.46)
Unemployed	0.08	0.16	0.15
	(0.27)	(0.37)	(0.36)
Real monthly earnings (ZAR)	3,436.57	10,292.10	9,405.30
	(971.43)	(8,106.11)	(12,188.40)
Past work experience (months)	10.86	-	-
	(12.88)		
Household variables			
Real hh monthly earnings (ZAR)	-	15,631.61	12,703.30
		(17,470.85)	(17,677.18)
Number of household residents	-	3.79	4.39
		(2.66)	(3.23)
Real p.c. hh monthly income (ZAR)	-	6,607.39	4,825.09
		(9,426.36)	(8,329.40)
Real p.c. hh expenditure (ZAR)	-	5,210.56	3,315.46
		(14,047.47)	(5,304.49)
Employment dynamics: 2014-2017			
Lost job between 2014-17	-	0.08	0.11
		(0.27)	(0.32)
Employed in 2014 and 2017	-	0.27	0.40
		(0.44)	(0.49)
Unemployed in 2014 and 2017	-	0.42	0.35
		(0.49)	(0.48)
Found job between 2014-17	-	0.23	0.14
		(0.42)	(0.35)
Observations	203	1988	16036

Notes: Standard deviations are reported in parentheses. The first column reports averages and standard deviations for the laboratory sample. Columns 2 and 3 report statistics from the 2017 Wave of the National Income Dynamics Study. For comparability with the lab sample, the sample in Column 2 is restricted to urban non-white adults under the age of 28 who have at least completed secondary school. Statistics reported in columns 2 and 3 are weighted using NIDS post-stratified weights. All monetary figures are deflated to August 2022.

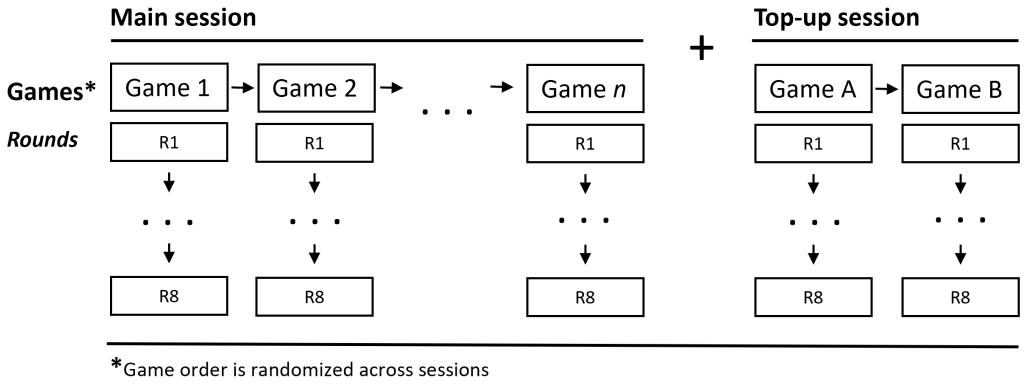
2.3.2 The structure of the experimental sessions

The experimental sessions have a simple hierarchical structure, represented graphically in Figure 2.1: Games – representing treatment variants described below – are nested within the overarching session, and rounds are nested within games. The overarching experimental session lasted on average 90 to 120 minutes. Each subject was eligible to play each game only once. Subjects were invited to complete two unsupervised top-up games, which they could access remotely, but only if they had already completed the main session in person. Details on this top-up session are provided in subsection 2.3.4. Henceforth, my definition of 'experimental session' will include top-up sessions for those who completed them.

Each experimental session was made up of between five and six games, representing the treatment conditions described in Table 2.4.1 and Appendix 2.C. In order to rule out possible sequencing effects, the order in which the games were played was randomised in each experimental session. I ran a total of 10 in-person sessions.[9]

[9]While in all of the in-person experimental sessions ordering was randomised, this could not be done for the top-up sessions, which only those who had already completed in-person sessions could participate in. This has implications only for the *SearchNoRej* and *SearchMoreRej* conditions, while order was always randomised for the *Search* and *Stay* conditions. Nevertheless, ordering effects may be important to account for. I do this by repeating the main analysis while controlling for game order where there is enough variation to do so (i.e. for the *Search* and *Stay* conditions), as seen in Table 2.A.1

Figure 2.1: Session structure



Each experimental session involved between 7 and 35 subjects.[10] Apart from a practice round[11], all decisions had monetary consequences. In each game, players were eligible to receive a pay-off of either ZAR55 (US$3.30) for the higher paying option (usually 'Search for a new job'), ZAR50 (US$3.00) for the lower paying option (usually 'focus on your current job'), or nothing. Since each session consisted of multiple games, the player's earnings over the session were simply the sum of earnings across all games, plus the fixed cash show-up fee of ZAR80 (US$4.50).[12] On average subjects received approximately ZAR80 (US$4.50).

All decisions were made on a computer screen. Participants were seated with an empty seat separating one another so as to maintain privacy and ensure they could not observe each others' decisions or the outcomes of their games. Verbal communication of decisions and outcomes was not permitted.

At the beginning of the experimental session, subjects were guided through

[10]Larger sessions were divided into smaller virtual 'rooms' of between 7 and 12 subjects.
[11]Practice data is excluded from the study.
[12]The relatively large show-up fee was chosen for fairness concerns – the experiment's design was such that I expected a substantial minority of players not to receive any payoff through the game itself, so guaranteeing an acceptable base payment was deemed to be important by myself and the implementation partner.

a set of pages which described the game and provided a comprehensive set of instructions. I personally guided all the experimental sessions. The introduction also included a shortened practice round in which players were guided through the choices in the game, and could practice the real effort task (described in subsection 2.3.5 below). Players also had the opportunity to ask questions during this introductory session.

In this introductory session I also explained the framing that would be used throughout the game – presenting the decisions which players were asked to make as a decision to 'Search for a job' or a decision to 'Focus on your current job'. Subjects were asked to imagine that they are employed on a fixed-term contract that is due to terminate in a determinate number of weeks, and that their decisions and the effort they exert would determine whether or not they are 'offered a new job', are 'retained in their current job', or 'fall into unemployment'.

Following the introduction, practice round and comprehension test, subjects could begin playing the games. In most sessions players completed five or six separate games. After each game, subjects were informed of their earnings for that particular game, and after the final game they were informed of their earnings for the full session. Total earnings is simply the sum of earnings across all games played.

After completing the games, players were asked to complete a survey collecting information on demographic characteristics, labour market history and behaviour, risk and time preferences, and psycho-social characteristics. After having completed the survey they received their compensation and pay-off and could leave.

The behavioural game and accompanying survey were designed and and

implemented in oTree (D. Chen et al., 2016)

2.3.3 Game protocol

For clarity of exposition, in this subsection I only describe the 'base game'. As described in the preceding subsection, in each experimental session participants play several variants of this base game, in each of which a key parameter is varied. These variations are described in the next section.

For a player, the (stated) objective of the game is simply to maximise earnings. In each round, players choose between two job offers: they can seek employment either by trying to find a new job *by searching*, or by trying to secure a contract extension *by focusing on their current job*.[13] In each version of the game, one choice strictly dominates the other in terms of expected pecuniary returns. Therefore, in expectation, players will achieve the payoff-maximising objective by always pursuing a strategy to secure the job offer with the highest payoff.

In each game, this choice is made multiple times. Specifically, each game consists of eight (or 16 in one variant) rounds in which the player will need to decide whether to *search* or *focus on their current job*. Having made the decision, the player then completes a real effort task, intended to represent the effort that they put into their chosen activity (described in subsection 2.3.5 below). Outcomes – that is, whether or not a player receives a job offer in their given choice – depends on the player's performance in the real

[13]The choice to frame the scenario explicitly in labour market terms is deliberate: I am seeking to understand labour market behaviour in the first instance, and hence prefer to frame the experiment in these terms rather than use neutral language. The choice of subjects (young entry-level white collar workers) is made for the same reason – for their relevance in exploring this particular labour market dynamic. However, by comparing *SearchNoRej* and *Search* conditions, I also seek to identify whether this explicit framing has an impact on behaviour in the laboratory setting.

effort task, and luck (described in more detail in subsection 2.3.5 below). In addition, players are informed that their chances of 'finding a new job' or 'securing a contract extension' are exactly equal for a given level of effort they exert in the real effort task.

Apart from differences in pay-off, the main difference between *choosing to search* and *choosing to focus on one's current job* is that if a subject *chooses to search* she receives a verdict on the outcome of her choice every round (at median effort this is a rejection in 96 percent of cases), compared to if a subject *chooses to focus on her current job*, in which case information on the outcome of her decision is withheld until the end of the game (at median effort this is a rejection in 72 percent of cases).[14] The *cumulative* probability of receiving an offer is, however, identical across the two activities (28 percent at median effort). There is no penalty to switching between activities during the game.

One of the key objectives of the game is to shut down the channel through which players learn, by experiencing rejection, about the mechanisms determining outcomes. To do this, I make sure that players in both the Search and Stay conditions have access to exactly the same information on the mecha-

[14]This round-by-round rejection rate may appear high, potentially raising concerns that if rejection rates in real life are substantially lower, behaviour in the game may display qualitatively different patterns than behaviour in the real world. I selected this high rejection rate for two reasons, the first empirical and the second pragmatic: *1)* At least in the South African context, a 96 percent rejection rate per application does not seem unrealistically high. One recent study on job search behaviour of a demographically similar sample in South Africa found that the average jobseeker submitted 10 job applications per week, and received only 0.17 job offers in the last 30 days (Garlick et al., 2023). This implies a success rate per application of only 0.4 percent. Even if we assume that this estimated success rate is a substantial underestimate, the 4 percent success rate in the game does not appear to be unrealistically low. *2)* I also chose this low acceptance rate for pragmatic reasons: I effectively 'lose' sample members when they are successful when searching for a new job, so a high rejection rate means that I am able to retain more of the sample.

nisms determining outcomes: I provide high frequency information on players' performance, the likelihood of success, and the relationship between performance and the likelihood of success in a way that is *identical* regardless of whether a player chooses to search or not to search. The only thing that differs is the information that is provided on the outcome – which in itself carries no additional information on the underlying probabilities or on the way the game works.

Each round consists of three steps (depicted in Figure 2.2):

1. **Choice:** Subjects are prompted to make a choice between *searching for a new job*, or *focusing on their current job*. Each option is associated with an identical cumulative probability of receiving a job offer, but a different payoff conditional on receiving an offer. This is communicated clearly to participants. Participants are tested on their comprehension of this concept before starting each game.

2. **Effort task:** Regardless of what subjects choose, they complete a real effort task which determines their likelihood of receiving an offer. The effort task is a simple slider activity based on Gill and Prowse (2019), where subjects have 40 seconds to complete a maximum of 30 sliders. This task was calibrated through piloting so that on average participants would complete 12 sliders, yielding a round-by-round job offer rate of 4 percent. Probabilities of receiving an offer were transparently determined by performance on the real effort task – each slider represented a 1/3 percentage point increase in the probability of receiving an offer, whether the decision was made to search or not to search. Comprehension checks ensured participants understood this relationship.

3. **Results:** Subjects are shown a results screen.

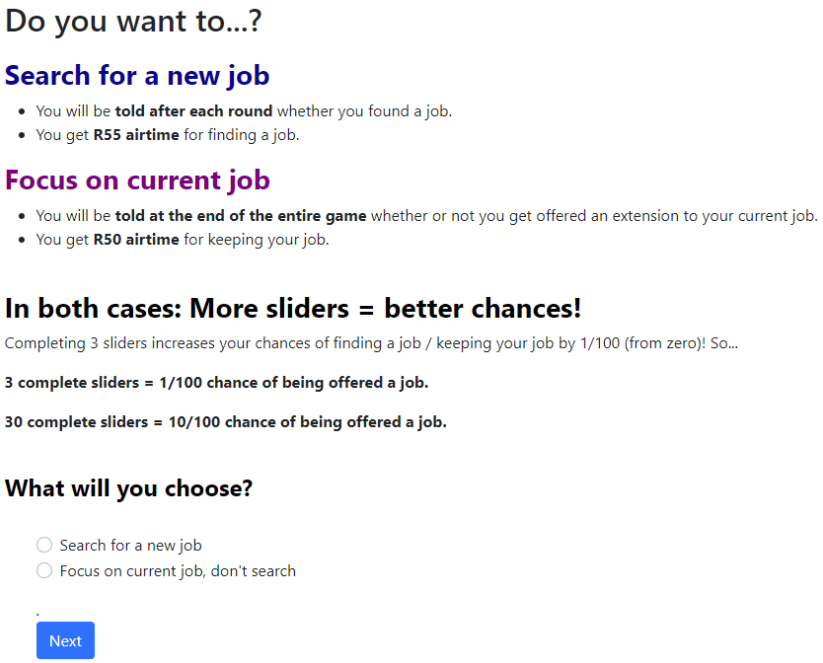
(a) If they had chosen *not to search*, they are informed only how many sliders they completed and the corresponding probability of receiving an offer. Information on whether or not they received an offer is withheld until the end of the game.

(b) If they had chosen *to search*, they are informed whether or not they received an offer, in addition to being told how many sliders they completed and the corresponding probability of receiving an offer.[15]

[15]The exception to this is the *SearchNoRej* condition, in which respondents are not informed in this step whether or not they received an offer. See Table 2.4.1.

Figure 2.2: Anatomy of a round

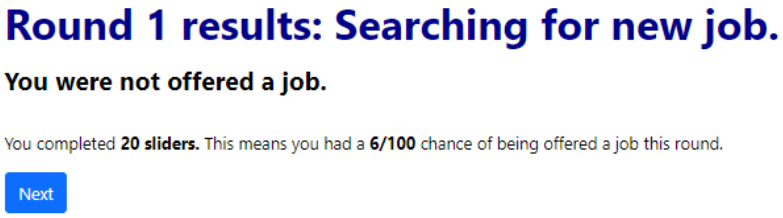
(a) Decision screen



(b) Real effort task



(c) Results screen



If players choose to search and are successful in receiving a job offer in a given round, they are no longer prompted to make decisions between searching and not searching. Rather, they are informed that they will retain their offer if they maintain a modestly high standard of performance in the effort task going forward.[16] For the purposes of the study, once players receive an offer, I effectively lose the chance to collect data from these individuals for the remainder of that particular game, since they are not longer prompted to make decisions and so I no longer observe the key outcome of interest for these individuals.

Before each game subjects are also informed of the *game specific characteristics* – described in detail in Section 2.4 – for instance, the expected pay-off to securing a new job versus keeping their current job, whether or not they will be informed immediately of the outcome of their job search if they decide to search (i.e. exposed to potentially frequent rejection), and over how many rounds the game is played. Because each game differs in these important respects from all other games, before beginning each game a comprehension test was used to ensure that each participant understood the distinctive characteristics of each game. Importantly, participants were required to correctly identify in which of the two tasks 'you can earn more'. While expected differences in earnings between the two tasks is very small in absolute terms, by prompting subjects to simply identify which of the two is more remunerative, I potentially increase the propensity for subjects to engage in heuristic reasoning, choosing as a default the higher-payoff activity. Subjects could not start the game until they had answered all questions

[16]Specifically, players keep their offer if they complete as many sliders as in the round in which they were made an offer, minus four. For instance, if a player receives an offer having completed 14 sliders, she will need to complete 10 in subsequent rounds to keep her offer. In the absence of this minimum performance standard, players might be incentivised to search so as to be able to 'slack off' once they receive an offer.

correctly. See Appendix 2.B for a full list of comprehension test questions.

2.3.4 Top-up rounds

After all in-person sessions had been held, subjects were invited to participate in an additional 'top-up' session. This top-up session was made up of the two treatment conditions *SearchNoRej* and *SearchMoreRej* (ordering randomised).[17] The *SearchMoreRej* condition was introduced in the final two in-person sessions as well as in a top-up session for those who had participated in the first eight in-person sessions.

The *SearchNoRej* condition had not been included in the original in-person sessions, or in the pre-analysis plan (see Appendix 3.G). During the experiment I became concerned that I would not be able to separate the rejection effect from potential framing effects, which motivated the design of the *SearchNoRej* condition. Like *SearchMoreRej*, this condition was introduced in the final two in-person sessions as well as in the top-up session.

In total, 122 out of a potential 203 participants completed the *SearchNoRej* or *SearchMoreRej* treatments in either the top-up session or in-person sessions.[18] There may be concerns that differential take-up of the top-up sessions may introduce bias into the comparisons between treatments implemented in the top-up sessions and those in the main sessions. Overall, apart from some gender imbalance, I do not find evidence for substantial imbalance on observables between these samples, as reported in Appendix 3.E.[19]

[17] *SearchMoreRej* had been included in the in-person sessions, but a coding error unfortunately invalidated results.

[18] Participants in the top-up session were eligible for incentives as for the in-person sessions, but did not receive a show-up fee.

[19] The gender imbalance is not a major concern. As reported in Table 2.E.1 in Appendix 2.E, gender does not appear to be correlated with search behaviour in the game.

2.3.5 The real effort task

The effort task plays an important function in the game. It is important that subjects perceive outcomes as being determined not only by luck but also by a variable over which they have some control and which is linked to performance or ability. This is because the way that an unsuccessful outcome is interpreted in a pure lottery may be psychologically very different from how an unsuccessful outcome is interpreted where performance of some sort is implicated. The real effort task allows for this. In addition, the relationship between effort and outcomes also helps subjects perceive the game as being at somewhat realistic. The effort task is introduced as being a task which 'is very different from searching or working in real life, but which represents the effort that you put into these tasks'.

Importantly, the relationship between effort and the probability of receiving an offer is transparently revealed to players: Specifically, that the probability of receiving an offer (holding effort constant) is exactly equal for the two choices of searching or focusing on one's current job. The assumption that players understand this feature of the game is important for the causal identification of a rejection effect. It is also important for the effort task to be identical for both 'search' and 'stay' tasks in order that preferences over the real effort task itself are not driving results.

In behavioural economics and experimental psychology many approaches to measuring effort have been used, including, among others, solving puzzles, doing data entry tasks and cracking walnuts (see Charness and Kuhn (2011) and Charness, Gneezy, et al. (2018) for reviews).[20] I use the Gill and Prowse

[20]Effort is typically measured using a 'stated-effort' paradigm, or a 'real-effort' paradigm Brüggen and Strobel, 2007. Where the psychological aspects of effort are relevant to the research design, real effort tasks are preferable (Charness and Kuhn, 2011). For this

(2019) 'slider task'. In this task, participants are instructed to move a slider on a computer screen to a specific target point. The screen contains multiple sliders, and the subject's objective is to correctly 'complete' as many sliders as possible within an allotted time, where 'completing' means aligning the slider with the target point. A screenshot of the slider task is provided in Figure 2.2 above, showing that 2/30 sliders have been completed with 21 seconds remaining.

The slider task was first used by Gill and Prowse (2012), and has since been used in a large number of behavioural experiments spanning a broad range of topics (see Gill and Prowse (2019) for a list). The slider task has several advantages which make it an appropriate choice for this experiment: The task is very easy to understand and requires little explanation. It is also able to generate a finely granular measure of effort within a short time frame, allowing for repeated observations for the same individual. It is also easy to implement in a game designed using OTree, and key parameters (such as the number of slides and the allotted time) can be finely calibrated. Importantly, there is little scope for 'learning' in the task, meaning that the cost of effort remains roughly constant within individuals over time (Charness and Kuhn, 2011).

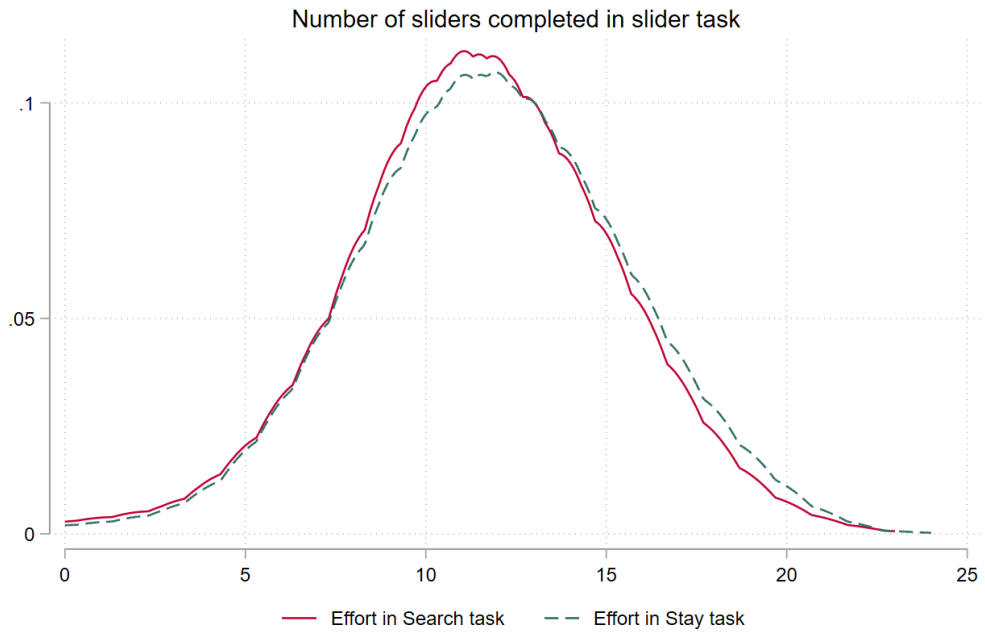
In this game, subjects had 40 seconds to complete up to 30 slider bars, determined through piloting to yield a median slider completion of 12, corresponding to the target 4 percent per-round offer rate.[21] Figure 2.3 plots the

reason, I only consider real effort tasks here.

[21]The probability of receiving an offer in any given round is determined by the number of sliders completed and random chance. Each slider represents a 0.33 percentage point increase in the likelihood of receiving an offer from a base of zero. Therefore, any subjects' probability of receiving an offer varied from 0 percent (if she completed no sliders) to 10 percent (if she completed 30 sliders). Thus, since the median slider completion rate was 12, the median subject had a per round offer probability of 4 percent.

probability density function for slider completion, distinguishing participants who chose to 'Search' from those who decided to 'Stay'. The difference in slider completion rates is not significant.

Figure 2.3: Distribution of effort in the real effort task



Notes: Subjects had 40 seconds to complete up to 30 slider bars (Gill and Prowse, 2019).

2.3.6 Survey

After completing the behavioural games, subjects were asked to complete a short survey. The survey collected data on subjects' basic demographic characteristics, labour market experience and history, labour market expectations, search behaviour, psycho-social characteristics [22], propensity to com-

[22]Self-efficacy measured using the New Generalised Self-efficacy Scale (G. Chen et al., 2001)

mitting hothand or gamblers fallacies[23], and time preferences[24].

2.4 Treatments

The four main treatments of the game create variation in two key parameters, which I use to isolate the effect of rejection on search behaviour. The first source of variation is in the *expected monetary returns to searching for a job*. The second source of variation is in the *amount of rejection that players are exposed to when searching for a job*. What is common to all games is that workers make multiple choices between 'searching' and 'focusing on their current job'. This choice – to search or not to search – is the outcome which I anticipate will respond to the treatment conditions. Table 2.4.1 summarises the four core treatment conditions.[25].

[23]Drawing on Ayton and Fischer (2004)

[24]Time preferences measured following the approaches developed by J. Rose and M. Rose (2019) (including OTree app) and Attema et al. (2016)

[25]This is not an exhaustive list of the treatment conditions implemented in the experiment. Other treatment conditions created variation in the *content* of the rejection messages. These are summarised in Appendix 2.C.

Table 2.4.1: Primary treatments

Treatment	**Payoff: finding job (ZAR)**	**Payoff: keeping job (ZAR)**	**Rejection signals when searching?**	**Round-by-round rejection rate (@ median effort)**	**n. rounds**
1. Stay	**50**	55	YES	96%	8
2. Search	55	50	YES	96%	8
3. SearchNoRej	55	50	**NO**	96%	8
4. SearchMoreRej	55	50	YES	**98%**	**16**

Notes: The round-by-round rejection rate is, for illustrative purposes, set at the (approximate) observed median effort (12/30 sliders completed). The probability of receiving an offer in any given round is determined by the player's performance in the effort task in that round, and ranges from 0% to 10% in the 8-round games, and 0% to 5% in the 16 round game. At median effort, the *cumulative* probability of receiving an offer over the course of the game is approximately equal in Treatments 1-3 (27.9%) and Treatment 4 (27.6%). T1, T2, and T4 were pre-specified, while T3 was added during fieldwork. See Appendix 2.C for a discussion.

In *Stay*, there is a monetary incentive to focusing on ones' current job – i.e. not searching for a job. As long as rejection signals are a net psychological cost, then both rejection effects and cash incentives pull in the same direction. Therefore, both rejection indifferent and rejection averse players would prefer not to search. Rejection aversion as an expression of active information avoidance should be inconsequential in this condition: Expected-payoff maximising players should never experience rejection because they should never choose to search.

In *Search*, there is a monetary incentive to search for a job. If information on rejection enters the individual's utility function negatively, then rejection information and cash incentives pull in opposite directions. This yields two possible outcomes: If an individual's aversion to information on rejection is strong enough, then rejection effects will dominate earnings incentives so that the individual will choose not to search. If, however, an individual's

rejection aversion is not strong enough for rejection effects to dominate earnings incentives, then the individual will choose to search. Thus, unlike in the *Stay* condition, now information avoidance becomes consequential to search behaviour.

If an individual's incentives to avoid information respond to past exposure to information, rejection effects may start out being dominated by pecuniary incentives, and in later periods may come to dominate pecuniary incentives. Because there is a well-defined monetary premium to searching, we can interpret decisions to stop searching as a willingness to forfeit the known monetary premium in order to avoid information on rejection. Because each game consists of multiple rounds, I will be able to investigate the presence of these dynamic information avoidance effects.

In the experimental set-up, rejection frequency is not the only difference between choosing to search and choosing to focus on one's current job. Framing effects may also matter, and the high rejection-frequency activity is explicitly framed as 'searching for a job' while the rejection-scarce activity is explicitly framed as 'focusing on your current job'. To the extent that that players carry behaviours, intuitions and norms from the real labour market into the game, this labelling may have a confounding influence on decisions to search or not to search, *independent of the experience of rejection.*

The *SearchNoRej* treatment provides a way of overcoming this challenge to the causal identification of a rejection effect. *SearchNoRej* is identical to *Search* in terms of incentives and framing, but differs in that if a player chooses to search, her rejection exposure is identical to the exposure had she chosen to 'focus on her current job'. The difference between *SearchNoRej* and *Search* can therefore be interpreted as the causal effect on search of being

exposed to some rejection compared to being exposed to no rejection, having experimentally netted out the framing effects which confound a comparison between *Stay* and *Search*.

SearchMoreRej is identical to *Search* except that the number of rounds is doubled and thus there is an increase from eight potential rejections to 16 potential rejections. Importantly, I hold the cumulative probability of receiving an offer roughly constant between *Search* (27.9 percent) and *SearchMoreRej* (27.6 percent). While the cumulative probability of receiving a payoff at the end of the game remains approximately unchanged, the likelihood of being rejected after any given round exactly doubles.

It is unclear ex ante whether the psychological cost of early rejections is higher than the psychological cost of later rejections. One could imagine either that individuals get a 'thick skin' after repeated rejection and become habituated, or that they get increasingly 'worn down'. Optimistic luck-based interpretations of rejection may be more difficult to sustain after repeated rejections, for instance. On the other hand, there is also likely heterogeneity across individuals in terms of their rejection aversion, with the result that highly rejection averse individuals will self-select out of search earlier than rejection insensitive individuals.

If the marginal effect of rejection on discouraging search is *decreasing* in exposure to rejection, we would expect the differences in search rates between *SearchNoRej* and *Search* to be greater than the differences in search rates between *Search* and *SearchMoreRej*. If the marginal effect of rejection on discouraging search is *increasing* in exposure to rejection, we would expect the opposite.

2.5 Results

2.5.1 Descriptive comparisons of search behaviour across treatments

2.5.1.1 Treatment 1: *Stay*

In this treatment subjects have a monetary incentive *not to search* (see Table 2.4.1). Players are informed that they will receive ZAR55 (US$3.30) if they keep their current job, and ZAR50(US$3.00) if they search and succeed in finding a new job, and that their chances of receiving an offer for either are equal (holding effort constant).

Figure 2.4 plots the evolution of the proportion of decisions made to search for the eight rounds over which the *Stay* game is played. The flat horizontal line plots the average search rate over the eight rounds.

The first thing to note is that players adopt a mixed strategy – switching between 'searching' and 'focusing on ones current job', so that search rates do not start at or approach zero.[26]

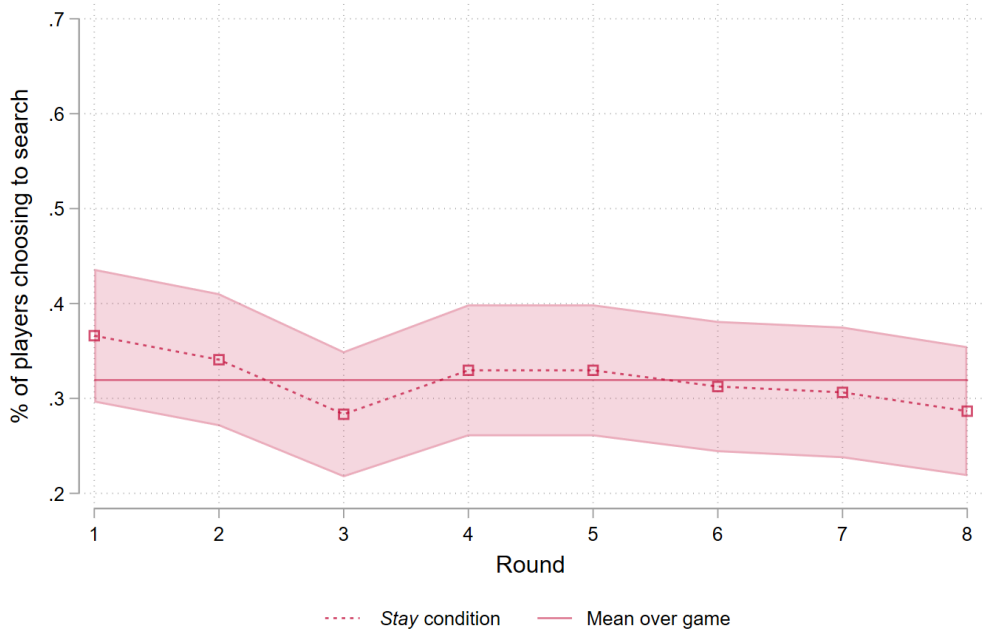
As described above, if the only determinant of players' decisions is the maximisation of expected payoffs, we would expect no players to choose to search in the 'Stay' condition. The result that we observe players to search in this condition approximately 30 percent of the time suggests that there are other determinants of subjects' decision making apart from payoff-maximisation.

[26]These non-zero search rates are driven both by within-subject switching, and between-subject heterogeneity in terms of game strategies – i.e. the presence of 'always searchers' inflates average search rates (see Figure 2.8). Removing these 'always searchers' does not bring average search rates to zero, pointing to the role played by within-subject switching strategies.

There may be many reasons for this: For instance, subjects may be curious about 'trying out' the search option – and given the very small differences in expected returns across the two conditions per round, this curiosity is not very costly in terms of forfeited expected earnings. Subjects may also simply be making random errors, perhaps because of an imperfect understanding of the game's instructions. It is also likely that subjects adopt strategies influenced by heuristic rules of thumb which are determined by search behaviour they engage in 'in the real world'. For example, strategies which may be sensible in the real world – such as diversifying risk by searching while on the job even while aspiring to be retained – may spill over into the game, despite the fact that there is no benefit to diversification in the game. These determinants of search behaviour introduce some noise into the results, but ought to apply equally across treatment conditions, and so are unlikely to be sources of bias.

Importantly, Figure 2.4 demonstrates that search rates remain flat for the eight rounds over which the *Stay* condition is played. While point estimates of search rates indicate a slight decrease in searching, these differences are not significant.

Figure 2.4: Evolution of search rates in the *Stay* treatment condition



Notes: Shaded area represents the 95% confidence interval of point estimates in each round. n = 183

2.5.1.2 Treatment 2: *Search*

In this treatment subjects have a monetary incentive *to search* (see Table 2.4.1). Players are informed that they will receive ZAR50 (US$3.00) if they keep their current job, and ZAR55 (US$3.30) if they search and succeed in finding a new job, and that their chances of receiving an offer for either are equal (holding effort constant).

Figure 2.5 plots the evolution of the proportion of decisions made to search for the eight rounds over which the *Search* game is played, comparing this to the trends in the *Stay* condition (as shown in Figure 2.4). The bottom two flat horizontal lines plot the average search rate over the eight rounds for the *Search* and *Stay* conditions respectively. The maroon flat horizontal

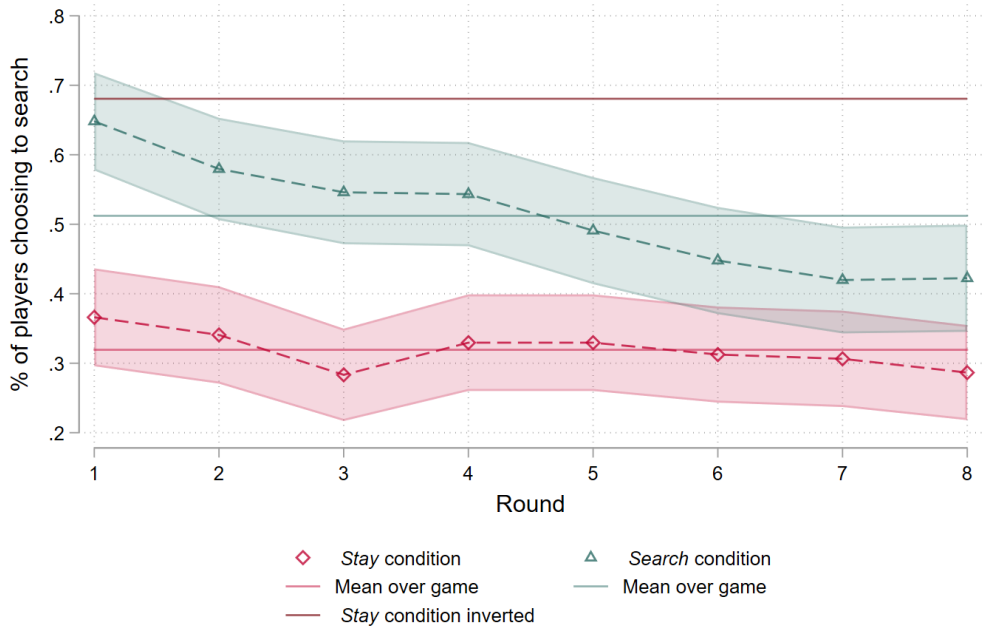
line plots what we can assume would be the average search rate in the *Search* condition *if subjects respond symmetrically to monetary incentives to search and not to search.* That is, the line is intended to represent how much rejection insensitive would players search if they responded symmetrically to inverted monetary incentives.

In the *Stay* condition, over the eight rounds players chose to search 30 percent of the time (solid red line). Thus, if players responded symmetrically to the inversion of the monetary incentives in the *Search* condition, we would expect search rates in the *Search* condition to be 70 percent (solid maroon line). However, we observe the average search rate in the *search* condition to be 50 percent (solid green line). This implies that the monetary incentive to stay 'bites' more than the same monetary incentive to search. This asymmetry suggests that there is a disincentive (by construction, rejection and framing effects) to searching that offsets the monetary incentive to search. Search rates in the *Search* condition are approximately 45 percent lower than would be expected under the assumption of a symmetrical response.

These averages hide important insights which are revealed by an analysis of the dynamics of search rates. Importantly, search rates in the first round in the *Search* condition are exactly what would be implied by symmetry with the *Stay* condition: 65 percent in *Search* round 1 versus 35 percent in *Stay* round 1. In other words, in the first round, the monetary incentive to search appears to 'bite' in the same way as the monetary incentive to stay. However, after the first round, there is a sustained decline in search rates in the *Search* condition (dotted green line) while search rates in the *Stay* condition remain relatively stable. Ultimately, by the sixth round the difference in search rates in the *Search* and *Stay* conditions are statistically insignificant. This convergence between the two search trend lines suggests that after

repeated rejections, the disincentive to search (plausibly an rejection information avoidance effect) dominates the monetary incentives to search, and induces those players initially induced by the monetary incentives to search, to switch strategies.

Figure 2.5: Evolution of search rates in the *Stay* vs *Search* treatment conditions



Notes: Shaded area represents the 95% confidence interval of point estimates in each round. *Stay*: n = 183. *Search*: n = 179

As discussed in Section 2.4, rejection aversion may not be the only thing that motivates a player to decide not to search – the choices are explicitly framed as 'search' and 'focusing on your current job'. To the extent that players bring norms, habits and intuitions from the labour market to the game, this labour market framing may affect search behaviour independently from the effect of rejection. Figure 2.5 already provides suggestive evidence that these framing effects are not strong: framing effects are presumably

stable over time, while rejection aversion may change with past experiences of rejection (as discussed in Section 2.6). The discouragement we observe in the *Search* condition in Figure 2.5 suggests that the average differences between search rates in *Search* and *Stay* are driven by time-varying factors, not stable factors such as framing effects. Nevertheless, by comparing *Search* and *SearchNoRej*, I provide a cleaner means of identifying the rejection effect.

2.5.1.3 Treatment 3: *SearchNoRej*

Comparisons between the *SearchNoRej* and *Search* treatments allow us to isolate the effect of rejection on job search. As in *Search*, in this treatment subjects have a monetary incentive *to search* (see Table 2.4.1). Players are informed that they will receive ZAR50 (US$3.00) if they keep their current job, and ZAR55 (US$3.30) if they search and succeed in finding a new job, and that their chances of receiving an offer for either are equal (holding effort constant).

The key characteristic of the *SearchNoRej* condition is that, if a player chooses to search, she does not receive information on the outcome of her decision to search in the Results page after each round (see Figure 2.2). The outcome on whether or not she received an offer is withheld until the end of the game – just as if she had chosen *not* to search. Since framing is held constant between *Search* and *SearchNoRej*, but the former condition exposes players to frequent rejection while the latter does not, a comparison between the search rates in these two conditions yields a clean identification of the rejection effect on search.

Figure 2.6 compares search rates in the *Search* and *SearchNoRej* conditions. The first panel presents results for the full sample, while the second

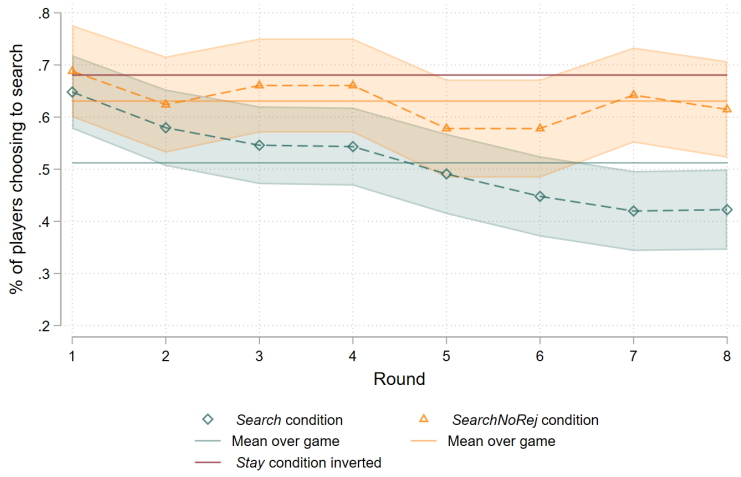
panel presents results for the sub-sample of those players who had decided to search in the first round. First, we observe that average search rates in the *SearchNoRej* condition (solid orange line) are higher than the search rates in the *Search* condition (solid green line). As we shall see in Table 2.5.2, these differences are economically and statistically significant. Differences between the average search rates in the *SearchNoRej* condition and the inverted *Stay* condition (solid red line) are smaller, and as we shall see in Table 2.5.3, are insignificant.

Search rates for the *Search* and *SearchNoRej* conditions are similar in the first round, but while search rates decline over the eight rounds in the *Search* condition until they fall to approximately 40 percent, in the *Search-NoRej* condition search rates remain stable between approximately 60 and 70 percent.

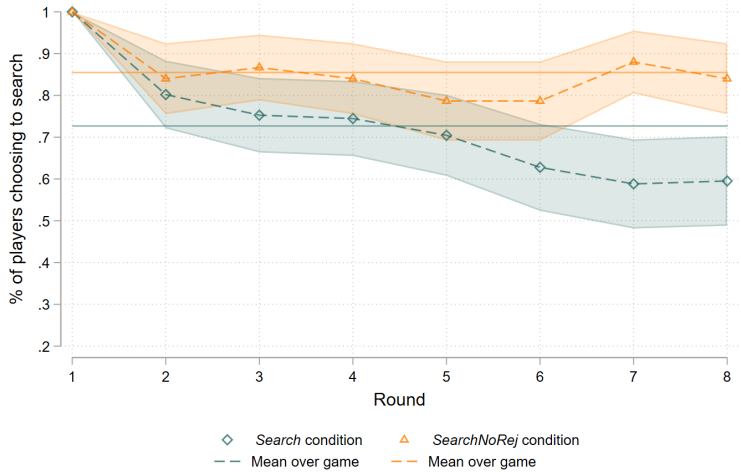
Figure 2.6b shows that, limiting the analysis to those who started out searching in the first round, we see that approximately 85 percent of those in the *SearchNoRej* condition were still searching in the final two rounds, while the search rate among for those in the *Search* condition had dropped to below 60 percent.

Figure 2.6: Evolution of search rates in the *Search* vs *SearchNoRej* treatment conditions

a) Evolution of search rate **for full sample**



b) Evolution of search rate **for those players who searched in R1**



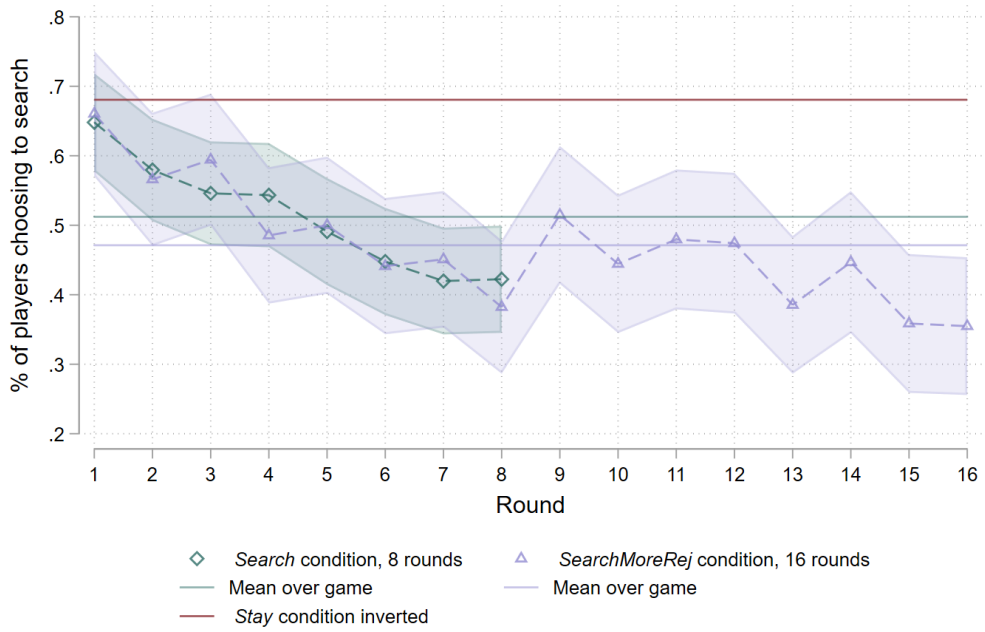
Notes: Shaded area represents the 95% confidence interval of point estimates in each round. *Search*: n = 179. *SearchNoRej*: n = 109

2.5.1.4 Treatment 4: *SearchMoreRej*

Comparisons between the *SearchMoreRej* and *Search* treatments allow us to determine whether the marginal effect of rejection on discouraging search is *decreasing* or *increasing* in exposure to rejection. That is, is discouragement sustained after the first eight rejections or does it flatten out?

As in *Search*, in this treatment subjects have a monetary incentive *to search* (see Table 2.4.1). The key characteristic of the *SearchMoreRej* condition is that the number of rounds is doubled and thus there is an increase from eight to 16 potential rejections. I hold the cumulative probability of receiving an offer constant between *Search*(27.9 percent) and *SearchMoreRej*(27.6 percent). Thus, while the cumulative probability of receiving a payoff at the end of the game remains approximately unchanged, the likelihood of being rejected in any given round doubles.

Figure 2.7: Evolution of search rates in the *Search* vs *SearchMoreRej* treatment conditions



Notes: Shaded area represents the 95% confidence interval of point estimates in each round. *Search*: n = 179. *SearchMoreRej*: n = 109

Perhaps surprisingly, average search rates in *Search* and *SearchMoreRej* are very similar, and as we shall see in Table 2.5.4 are statistically indistinguishable. Analysing trends, we see that search rates in *SearchMoreRej* track search rates in *Search* very closely for the first eight rounds as we should expect, but then show no evidence of a further decline.[27] This suggests that the discouragement effects of rejection are steepest in early rounds, after which point search rates flatten out – perhaps due to the more rejection sensitive subjects selecting out of search.

[27]Search rates in the *SearchMoreRej* condition are noisier than in the *Search* condition due to a smaller sample size (see Appendix 2.C), perhaps explaining the counter-intuitive apparent (though statistically insignificant) increase in search rates in the rounds immediately following round eight.

2.5.1.5 Comparing player strategies across games

The preceding analysis gives little insight into the strategies that players adopt in the game, and how these strategies respond to the different treatment conditions. Expected-payoff maximising and rejection insensitive players would be 'Always stayers' in the *Stay* condition, and 'Always searchers' in all the other conditions. In reality, there is heterogeneity both within players over time (as players start with one strategy and then switch to another during the game, or follow a mixed strategy), and across players (with some players being more disposed to search than others).

Figure 2.8 shows the distribution of the proportion of rounds in which players chose to search for each of the treatment conditions discussed above. We observe a bunching of search rates around 0 and 100 percent in all four treatments, but with expectedly pronounced bunching at 0 percent in the *Stay* condition, and at 100 percent in the other three conditions. It is also noteworthy that some players appear to adopt an inflexible 'always search' or 'never search' position, persisting in this strategy even in conditions where the monetary incentives pull in the opposite direction. One explanation might be that players apply simple heuristics (i.e. 'always search' or 'never search') for evaluating the decision to search / not to search in each round, which could help explain the bunching at 0 and 100 percent.

Differences between the average search rates across the conditions are therefore driven by those who are induced to switch strategies by the different treatments. The number of 'never searchers' is twice as large in the *Stay* condition (red line) as it is in the other three conditions, while this condition has about half as many 'always searchers' as the other three conditions. The change in incentives therefore induces some 'never searchers' to search, and

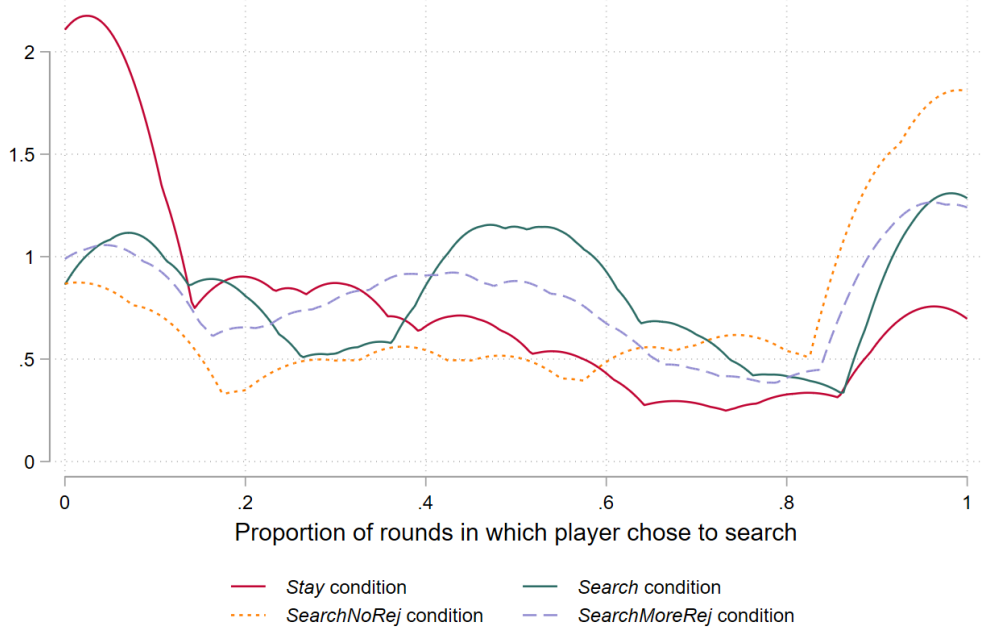
some ‘sometime searchers’ to become ‘always searchers’.

The difference between the *Search* condition and the *SearchNoRej* condition is slightly more complex: *Search* has a trimodal distribution with a high concentration of ‘sometimes searchers’, while *SearchNoRej* has a bimodal distribution. In other words, there are fewer ‘always searchers’ and more ‘sometimes searchers’ in the *Search* condition relative to the *SearchNoRej* condition. This suggests that fewer players who begin searching in the *SearchNoRej* condition are induced to switch strategies during the game than in the *Search* condition – a pattern which is consistent with the discouragement effect of rejection we observe in Figure 2.6.

Overall, the search patterns in *SearchNoRej* look roughly symmetrical to the search patterns in *Stay* – encouraging, given that fact that these conditions are identical apart from the their framing and the inversion of the pecuniary incentives. This gives us suggestive evidence that framing effects are not playing a major role on determining search behaviour in the laboratory setting – confirmed in Table 2.5.3 below.

It is also noteworthy that the number of ‘never searchers’ is very similar in the *Search*, *SearchNoRej* and *SearchMoreRej* conditions – which is consistent with the fact that for ‘never searchers’ these three conditions are identical – since ‘never searchers’ never search, they never experience rejection and therefore remain unaffected by the experimental variation which I have created in rejection across these three conditions.

Figure 2.8: Distribution of search behaviour across treatment conditions



2.5.2 Estimating the effect of rejection on job-search

In this subsection I present results of a series of fixed effects regressions to compare average search rates across the conditions discussed above. In all tables, the dependent variable is the binary Search/Not Search outcome observed at each round r in each condition c and for each individual player i. The independent variable of interest T is defined as a dummy equal to one if the individual was in the treatment condition of interest and equal to zero if the individual was in the comparison treatment condition.

The use of a within-player experimental design requires at least two additional measures for causal identification: *1)* Player fixed effects A_i are included so as to control for time-invariant player characteristics; *2)* At a given point in time, a player's experience of rejection and job finding is de-

termined by the choices she makes as well as the order in which she receives treatments. Thus, different players playing the same game may have had a different 'history' within the session prior to the treatment condition under consideration and therefore will have experienced different amounts of rejection and/or job finding in prior rounds. For this reason, I also control for past outcomes in previous sessions P_{ic-n}.

The order of treatments *Search* and *Stay* was randomised across experimental sessions, while *SearchNoRej* and *SearchMoreRej* were implemented later in an add-on session. This means that ordering was not randomly varied between *Search* and *SearchMoreRej* or *SearchNoRej*. For the sake of consistency, for this reason I exclude ordering fixed-effects in the main specifications. In a robustness test in Table 2.A.1 I investigate whether including ordering fixed-effects has an impact on results, and find that it does not.

Specifications thus take the following form:

$$Search_{rci} = \beta_0 + \beta_1 T_i + A_i + P_{ic-n} + \epsilon_{rci} \tag{2.1}$$

Since the unit of randomisation is individual, but the unit of analysis is individual-time period (i.e. in each regression there are two individual-time period observations), the error term ϵ_{rci} is clustered at the individual level (Abadie et al., 2017; Kim, 2022)

In a first specification I run regressions on the full sample of subjects. In a second specification I limit the analysis only to those subjects who were never made an offer in any of the two conditions under consideration. This is because successful searchers effectively select out of the sample since we no longer observe their choices once they receive an offer. In effect, this differ-

ential attrition may bias search rates downwards in the full sample relative to a specification which excludes job finders. However, since the same dynamic affects both the treatment condition of interest and the comparison condition, I do not expect this to be a consequential distortion.

Table 2.5.1: Treatment effects on search rates, comparing *Search* to *Stay* (inverted) conditions

	Full sample	Excluding job-finders
Search vs Stay (inverted)	-0.119**	-0.153***
	(0.053)	(0.056)
Constant	0.676***	0.701***
	(0.024)	(0.024)
Mean search rate for Stay (inverted)	0.68	0.68
% change, Search vs Stay (inverted)	-17.54	-22.55
Number of observations	2773	2600
Number of players	183	176
Indiv. Fixed Effects	Yes	Yes
Past outcomes in session	Yes	Yes
Excludes job-finders	Yes	No

Notes: *p < .1; ** p < .05; *** p < .01. Standard errors, clustered at the player level, are in parentheses. The outcome variable in both specifications is a binary variable equal to 1 if a player decides to search in a given round, and equal to 0 if a player decides to 'focus on her current job'. The independent variable of interest, *Search vs Stay (inverted)*, is defined as a dummy equal to one if the individual was in the treatment condition of interest (Search) and equal to zero if the individual was in the comparison treatment condition (Stay (inverted)). Coefficients can be interpreted as the treatment effect on search rates of being in the *Search* condition relative to the *Stay* (inverted) condition. The first column reports results for the full sample, while the second column reports results only for the subsample of players who did not search and receive a job offer in either condition under consideration.

Regression results confirm the narrative which emerges from the preceding analysis: Compared to what we would expect search rates to be if subjects responded symmetrically to inverted monetary incentives to search (*Stay (inverted)*), in Table 2.5.1 we observe search rates to be significantly lower in

reality when monetary incentives are inverted (*Search*). I interpret this difference as being driven by rejection information avoidance behaviour: Search rates in the *Search* condition are between 12 and 15 percentage points (or 18-23 percent) lower than what we would expect in the absence of rejection aversion. As discussed above, it is possible that framing effects contaminate these estimates.

A comparison between *Search* and *SearchNoRej* in Table 2.5.2 allows me to isolate the rejection effect from these possible framing effects. I find that, having netted out possible framing effects, the effect of rejection on search rates persists. A player's propensity to search is approximately 15 percentage points (or 25 percent) lower when a player is exposed to frequent rejection than when she is not exposed to frequent rejection, even when the choice is framed identically.

Table 2.5.2: Treatment effects on search rates, comparing *Search* to *Search-NoRej* conditions

	Full sample	Excluding job-finders
Search vs SearchNoRej	-0.156***	-0.150***
	(0.056)	(0.057)
Constant	0.678***	0.676***
	(0.033)	(0.034)
Mean search rate for SearchNoRej	0.63	0.63
% change, Search vs SearchNoRej	-24.80	-23.77
Number of observations	2225	2120
Number of players	160	154
Indiv. Fixed Effects	Yes	Yes
Past outcomes in session	Yes	Yes
Excludes job-finders	Yes	No

Notes: *p < .1; ** p < .05; *** p < .01. Standard errors, clustered at the player level, are in parentheses. The outcome variable in both specifications is a binary variable equal to 1 if a player decides to search in a given round, and equal to 0 if a player decides to 'focus on her current job'. The independent variable of interest, *Search vs SearchNoRej*, is defined as a dummy equal to one if the individual was in the treatment condition of interest (*Search*) and equal to zero if the individual was in the comparison treatment condition (*SearchNoRej*). Coefficients can be interpreted as the treatment effect on search rates of being in the *Search* condition relative to the *SearchNoRej* condition. The first column reports results for the full sample, while the second column reports results only for the subsample of players who did not search and receive a job offer in either condition under consideration.

A comparison between *Stay (inverted)* and *SearchNoRej* in Table 2.5.3 allows me to test directly for the presence of framing effects. We already have suggestive evidence that framing effects do not play a meaningful role in determining search behaviour, as evidenced by the roughly symmetrical search patterns in the SearchNoRej and Stay conditions in Figure 2.8 above. Table 2.5.3 provides further evidence that framing effects are small and statistically insignificant, with null effects fairly tightly centered on zero.

Finally, in Table 2.5.4 I test whether exposing players to additional rejection (while keeping cumulative probabilities of receiving an offer constant) further discourages job search. Results suggest that there is a small but insignificant (or weakly significant in the second specification) effect of additional rejection on job search. A direct comparison between the *Search vs SearchNoRej* results with the *Search vs SearchMoreRej* results suggests that the rejection effect on job search is steepest with the first rejections, after which point it flattens out, although I cannot rule out further declines in search rates of up to 9 percentage points (or 19 percent).

Table 2.5.3: Treatment effects on search rates, comparing *Stay* (inverted) to *SearchNoRej* conditions

	Full sample	Excluding job-finders
Stay (inverted) vs SearchNoRej	0.005	0.004
	(0.056)	(0.056)
Constant	0.683***	0.691***
	(0.035)	(0.036)
Mean search rate for SearchNoRej	0.63	0.63
% change, Stay(inv.) v SearchNoRej	0.87	0.63
Number of observations	2292	2224
Number of players	183	176
Indiv. Fixed Effects	Yes	Yes
Past outcomes in session	Yes	Yes
Excludes job-finders	Yes	No

Notes: *p < .1; ** p < .05; *** p < .01. Standard errors, clustered at the player level, are in parentheses. The outcome variable in both specifications is a binary variable equal to 1 if a player decides to search in a given round, and equal to 0 if a player decides to 'focus on her current job'. The independent variable of interest, *Stay (inverted) vs Search-NoRej*, is defined as a dummy equal to one if the individual was in the treatment condition of interest (*Stay (inverted)*) and equal to zero if the individual was in the comparison treatment condition (*SearchNoRej*). Coefficients can be interpreted as the treatment effect on search rates of being in the *Stay (inverted)* condition relative to the *SearchNoRej* condition. The first column reports results for the full sample, while the second column reports results only for the subsample of players who did not search and receive a job offer in either condition under consideration.

Table 2.5.4: Treatment effects on search rates, comparing *Search* to *SearchMoreRej* conditions

	Full sample	Excluding job-finders
Search vs SearchMoreRej	0.045	0.091*
	(0.047)	(0.050)
Constant	0.461***	0.430***
	(0.022)	(0.023)
Mean search rate for SearchMoreRej	0.47	0.47
% change, Search vs SearchMoreRej	9.56	19.28
Number of observations	2955	2704
Number of players	160	150
Indiv. Fixed Effects	Yes	Yes
Past outcomes in session	Yes	Yes
Excludes job-finders	Yes	No

Notes: *p < .1; ** p < .05; *** p < .01. Standard errors, clustered at the player level, are in parentheses. The outcome variable in both specifications is a binary variable equal to 1 if a player decides to search in a given round, and equal to 0 if a player decides to 'focus on her current job'. The independent variable of interest, *Search vs SearchMoreRej*, is defined as a dummy equal to one if the individual was in the treatment condition of interest (*Search*) and equal to zero if the individual was in the comparison treatment condition (*SearchMoreRej*). Coefficients can be interpreted as the treatment effect on search rates of being in the *Search* condition relative to the *SearchMoreRej* condition. The first column reports results for the full sample, while the second column reports results only for the subsample of players who did not search and receive a job offer in either condition under consideration.

2.5.3 Correlates of search behaviour and discouragement

Appendix 2.E investigates the demographic and behavioural correlates of the search behaviour I observe in the results presented above. Overall, I observe that older individuals are less likely to switch away from Search following a rejection, while black African workers are more likely than the coloured subsample to switch away from Search following a rejection. Interestingly, gender does not appear to be correlated with search behaviour in the laboratory game.[28] Those individuals who report struggling with their motivation to search for jobs in the real labour market also appear to be more easily discouraged in the game. There is some weak evidence of a correlation between a measure of self-assurance ('attributing success to one's own merit') and persistence in Search. Ranking above average on a self-efficacy scale (compared to below average) yields insignificant results. Importantly, I also find that individuals who are very optimistic about their chances of being retained in their (real) jobs, are also more likely to switch away from Search following a rejection.

It is difficult to determine whether the null result on self-efficacy is a true null, whether this is because such variables are difficult to reliably measure, or whether a deeper issue may be at play: Specifically, as suggested by work on fragile self-esteem (Kőszegi, Loewenstein, et al., 2021), we would expect individuals to avoid information on rejection if they have high but *fragile* self-esteem.[29] While it may be possible to measure self-esteem, it is

[28] Others have suggested that gender may be an important correlate of discouragement, with women more likely to be discouraged by rejection than men (Bapna et al., 2021).

[29] In Kőszegi, Loewenstein, et al. (2021)'s model, individuals have multiple 'personal self-esteem equilibria', and are either consciously or subconsciously aware that shocks to self-esteem may shift them from one equilibrium to another. Of particular relevance are

more difficult to measure the fragility of these beliefs. This gets in the way of a straightforward mapping of self-esteem beliefs onto rejection avoidant behaviour, and may help explain such null results.

2.6 A simple decision-making framework accounting for preferences for information on rejection

In the laboratory experiment I describe above, I find that: *a)* subjects appear to be initially naive with regard to their aversion to receiving information on rejection, *b)* subjects respond to early rejections by switching to an activity which has lower expected pecuniary returns but which also has lower exposure to high-frequency information on rejection, and *c)* that the drop-off in average search rates is steepest after early rejections and subsequently flattens out. In this section I propose a basic theoretical framework to help illuminate the mechanisms underlying these findings.

Let y_{itj} denote whether a positive payoff is received for individual i in each period t in activity j. The probability of obtaining a payoff in an activity

fragile high self-esteem equilibria – characterised by a large distance between positively biased beliefs about ability and true (low) ability. Kőszegi, Loewenstein, et al. (2021) show that these individuals face especially strong incentives to distort their choices so as to avoid a downward revision of self-esteem. Applying this model to job search, Kőszegi, Loewenstein, et al. (2021) propose that their model helps resolve an apparent paradox: One the one hand, the unemployed are systematically over-optimistic about finding jobs (Spinnewijn, 2015), while on the other hand, they appear to search surprisingly little (A. B. Krueger and Mueller, 2010; A. Krueger and Mueller, 2012). According to the 'fragile self-esteem' model, this paradox can be explained by the fact that individuals derive utility from occupying a high self-esteem equilibrium, but being aware that this equilibrium is 'fragile', avoid exposure to shocks – for instance, rejections in job search – which could move them on to a lower equilibrium. Kőszegi, Loewenstein, et al. (2021) imply that individuals may be willing to pay – in the form of foregone income and employment – to avoid these self-esteem shocks.

is given as $p_{it} = e_i + \epsilon_{it}$, where e_i denotes effort by individual $i, 0 \leq e_i \leq 1$, and ϵ_{it} ~ $N(0, \sigma^2)$ $\forall i, t$. In addition, the payoff w_{itj} is positive iff $p_{it} > T$ for some exogenous threshold T.[30]

Thus

$$\begin{aligned} y_{it} &= 1 \; if \; p_{it} > T \\ & \quad 0 \; otherwise \end{aligned}$$

Hence the payoffs $w_{itj} = w_j \cdot y_{it}$, $j \in \{J, S\}$

Rearranging, we can see that $p_{it} > T$ can be expressed as $(e_i - T) + \epsilon_{it} > 0$. Expressed in this way, and given the distributional assumptions made above, we can express $E(y_{it})$ as a probit model, where

$$Pr(y = 1|p_{it}) \equiv \Phi(e_i - T). \tag{2.2}$$

I assume that individuals draw utility from payoffs/wages w_{tj} and disutility from experiencing rejection ρ_{it}, both of which are informed by activities they undertake in each period. Individuals choose between two mutually exclusive activities $j \in \{J, S\}$ (as in the experiment) in each time period t: J, denoting effort in the current job and S, denoting effort in search for another job. The additively-separable intertemporal utility function is expressed as:

$$U_j(w_{itj}, \rho_{it}) = w_{itj} - \lambda_i(\alpha_{i0} \cdot \rho_{it})^2 \tag{2.3}$$

[30]The payoffs are identical in each period for a given activity, but differ across activities, so we suppress the subscript t.

where

$$\rho_{it} = \sum_{t=1}^{t} (1 - y_{it}). \tag{2.4}$$

w_{tj} denotes the individual's payoff in activity j. The weight parameter λ_i captures the extent to which individual i has a preference for not receiving information on rejection. If $\lambda_i = 0$, we return to the usual framework where preferences are determined over monetary returns w_{tj} alone. ρ_{it} is the cumulative number of rejections an individual has received up to time t. ρ_{it} is multiplied by α_{i0}, representing an individual's initial beliefs about ability (i.e. prior to receiving any rejection), so that the disutility of rejection increases in α_{i0}. This captures the notion that individuals with elevated beliefs about ability have 'more to lose' with experiences of rejection (Kőszegi, Loewenstein, et al., 2021). Note too that I use a quadratic form for the 'disutility of rejection' term, so as to capture increasing marginal costs of rejection.[31]

In line with the set up of the experiment, I assume that in activity S (the Search option), the individual is exposed to rejection after completing a task. In contrast, in activity J (the stay option), the individual does not receive rejection signals. I also assume that activity S yields higher returns than activity J in pecuniary terms, so that $w_S > w_J \, \forall t$.

Letting Φ denote $\Phi(e_i - T)$ (see equation 2.2 above), the expected utility in each activity can then be written as:

[31] This quadratic form captures the notion that the disutility if a later signal is larger than that of an earlier signal, as individuals get 'worn down' by rejection and become increasingly likely to avoid information following 'bad' news (Karlsson et al., 2009).

$$
\begin{aligned}
S : EU_S(w_{tS}, \rho_{it}) &= E(w_{iS}) - E(\lambda_i(\alpha_{i0} \cdot \rho_{it})^2) \\
&= w_S \cdot E(y_{it}) - \lambda_i \alpha_{i0}^2 E(\rho_{it}^2) \\
&= w_S \Phi - \lambda_i \alpha_{i0}^2 E(\rho_{it-1} + (1 - y_{it}))^2 \\
&= w_S \Phi - \lambda_i \alpha_{i0}^2 E((\rho_{it-1} + 1)^2 + 2(\rho_{it-1} + 1) y_{it} + y_{it}^2) \\
&= w_S \Phi - \lambda_i \alpha_{i0}^2 (\rho_{it-1} + 1)^2 + 2(\rho_{it-1} + 1)\Phi + \Phi)
\end{aligned}
$$

$$
\begin{aligned}
J : EU_J(w_{itJ}, \rho_{it}) &= E(w_{iJ}) - E(\lambda_i(\alpha_{i0} \cdot \rho_{it})^2) \\
&= w_J \cdot E(y_{it}) - \lambda_i(\alpha_{i0} \cdot \rho_{i\mathbf{t-1}})^2 \text{ since in J } \rho_{it} = \rho_{it-1} \\
&= w_{tJ} \Phi - \lambda_i \alpha_{i0}^2 (\rho_{it-1}^2)
\end{aligned}
$$

Simplifying, the choice of activity in each period can then be summarised by:

$$
\begin{aligned}
Activity = S \; if \; (w_S - w_J)\Phi > \lambda_i \alpha_{i0}^2 (2\rho_{it-1} + 1 + 2(\rho_{it-1} + 1)\Phi + \Phi) \\
J \; otherwise
\end{aligned}
\tag{2.5}
$$

Thus, individuals are more likely to search if λ_i is low, ρ_{it} is low, and/or $w_{tS} - w_{tJ}$ is high.

Note also that in the final period players will receive information on the outcome of their choice regardless of which activity they choose. For this reason, the 'disutility of rejection' term in EU_S and EU_J are identical in the final period, and only differences in earnings should affect choices. Since $w_{iS} > w_{iJ} \; \forall t$, individuals should revert to search in the final period.

Figure 2.9 graphically illustrates some of the characteristics of the model by plotting the search behaviour of two paradigmatic individuals over eight rounds of the game. In the representation in Figure 2.9, the model is param-

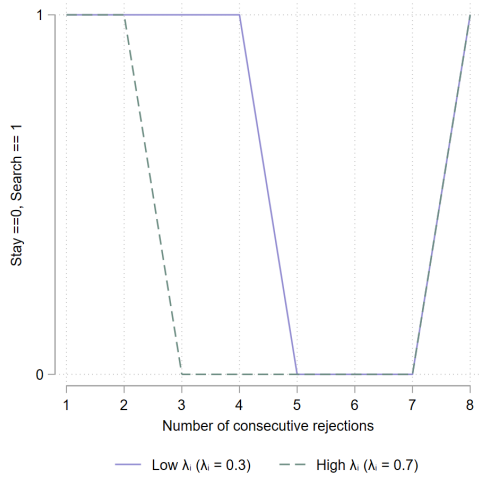
eterised so that in Round 1 both individuals search – that is, the pecuniary benefits of searching dominate the non-pecuniary benefits of notn searching. These two paradigmatic individuals are identical across all parameters apart from those that are specified to be different across the two individuals – λ_i in Figure 2.9a and α_{i0} in Figure 2.9b.

Individuals will switch from choosing Option S ('Searching') to Option J ('Staying') at the point at which the non-pecuniary benefits of not searching (the right hand side of expression 2.5) switch from being dominated by the pecuniary benefits to searching to dominating the pecuniary benefits of searching.

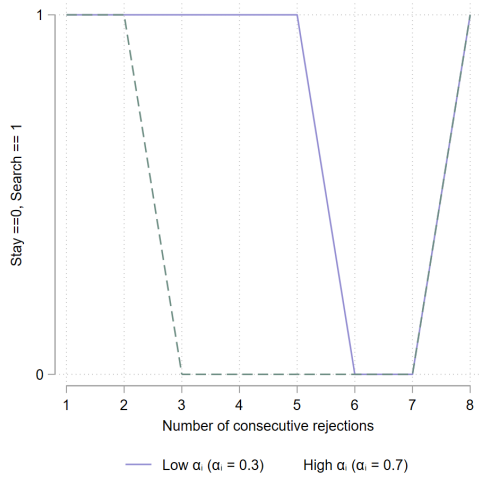
We can see in Figure 2.9a that individuals with low λ_i values – i.e. a low weight placed on rejection in the utility function – will persist in search longer than individuals with higher λ_i values. This makes intuitive sense – it will take more rejection for the pecuniary benefits of searching to be dominated by self-esteem concerns among individuals who place relatively little weight on rejection. Figure 2.9b illustrates that individuals with higher values of initial self-esteem (α_{i0}) will drop out of search sooner than individuals with initially lower levels of self-esteem. This is because the distance between the rejection signal and initial beliefs is larger for high self-esteem individuals than for low self-esteem individuals.

Figure 2.9: Comparing job search behaviour between individuals with different parameter values

a) High versus low λ_i values (preferences to avoid rejection)



b) High versus low α_{i0} values (initial level of self-esteem)



Notes: The Figure illustrates the search behaviour of two paradigmatic individuals over eight rounds of the game. Players search if $(w_S - w_J)\Phi > \lambda_i \alpha_{i0}^2 (2\rho_{it-1} + 1 + 2(\rho_{it-1} + 1)\Phi + \Phi)$. The model is parameterised so that in Round 1 both individuals search – that is, the pecuniary benefits of searching dominate the non-pecuniary benefits of not-searching. Apart from differing in the specified parameter values in each respective figure, the two paradigmatic individuals are identical across all other parameters. Individuals will switch from choosing Option S ('Searching') to Option J ('Staying') at the point at which the non-pecuniary benefits of not-searching (the right hand side of the above expression) dominate the pecuniary benefits to searching. Round 8 is assumed to be the final round. As described above, I show that all players ought to switch to Search in the final round.

It is worth limiting a few limitations of the model: First, the model does not explicitly account for the *fragility* of self-esteem beliefs. Following Kőszegi, Loewenstein, et al. (2021), we might expect that the fragility of self-esteem beliefs will be an important determinant of an individuals' propensity to actively avoid information that is not consonant with prior beliefs.

Second, the model does not account for the role of salience: Following Bordalo et al. (2022), we can contrast 'stimulus salience' as a *bottom-up* mechanism driving behaviour to 'goal relevance' as a *top-down* mechanism driving behaviour. In the case of the game played in the laboratory experiment, the 'goal' is to maximise earnings in the game, while (repeated) rejection may function as a salience stimulus in which 'bottom-up attention interferes with goals and causes suboptimal decisions' (Bordalo et al. (2022), p. 525). The role played by salience is simply that it changes the *weight* placed on a particular stimulus. In this case, the disutility of experiencing rejection may be weighted more heavily when it is more salient in high-rejection conditions. Importantly, this does not change the *direction* in which the stimulus pulls, but merely the weight which is placed on it. The role of salience could be formally accounted for in the basic model described above by multiplying the ρ_{it} rejection disutility term by a salience-weight parameter, and allowing this weight to increase with past exposure to rejection.

For the sake of simplicity and tractability, neither the role of salience nor the role of the fragility of prior beliefs is accounted for in this basic model.

2.7 Discussion

2.7.1 Discussion of main results

The framework I have laid out in the preceding section can help deepen our understanding of the mechanisms behind the three main findings of this paper. In the discussion that follows, I address each in turn.

Result 1: Prior to receiving any rejections in a given game, the average subject chooses the option that maximises the expected payoff rather than that which minimises exposure to rejection information.

In the first round of the game, the average subject appears to respond to monetary incentives as if they have no preferences for information – that is, observed behaviour (see Figure 2.5) closely approximates the paradigmatic cases presented in Figure 2.9. In other words, in the first round of a game, most subjects choose to search when there is a monetary incentive to search, and most choose to not search when there is a monetary incentive not to search (see Figure 2.5). Importantly, subjects in the first round respond symmetrically to pecuniary incentives to search and to pecuniary incentives not to search.

While in all games one option always strictly dominates the other in terms of expected monetary returns, the absolute differences in expected returns is very small once we account for the low probability of receiving a payoff. Therefore, if we account for some randomness in individuals' decision making, it is not clear ex ante that individuals would respond to the (small) pecuniary incentives in the way that we observe they do, even putting aside

preferences for information.

It therefore seems likely that individuals are engaging in heuristic reasoning to guide their decisions prior to exposure to rejection, where the simple rule of thumb is 'choose the option with the highest expected pay-off'. Two features of the game make the use of such heuristics more likely: First, players were encouraged to make decisions quickly: The other players in the session would only be able to proceed at the pace of the slowest player, which created pressure on all players to make quick decisions. Second, before beginning each game subjects were required to correctly identify in which of the two tasks 'you can earn more'. While expected differences in earnings between the two tasks is very small in absolute terms, by prompting subjects to simply identify which of the two is more remunerative, I potentially increase the propensity for subjects to engage in heuristic reasoning in which subjects default to the higher-payoff activity. As long as individuals are engaging in this heuristic reasoning, the role of random decision-making errors may be fairly weak, even if the absolute difference in expected returns between options is small.

Result 2: Exposure to rejection leads to a rapid and dramatic reduction in search rates: The average subject shifts from choosing the option that maximises the expected payoff to the option that minimises exposure to rejection information (when these are not the same).

Subjects respond to receiving rejection information by switching to an activity that has lower expected pecuniary returns but that also has lower exposure to high-frequency information on rejection. Since the disutility of rejection term in Equation 2.4 is not fixed but increases with exposure to

new information on rejection, we could expect to see individuals become more information avoidant after exposure to information which is discordant with their initial beliefs.

If this is the case, we may expect to see deviations from heuristically-based reasoning as information avoidance increases in importance within the individual's utility function relative to consumption utility. For some individuals the disutility of rejection will come to dominate the heuristic which dictates 'choose the option with the highest pecuniary return'. This may precipitate a switch to the activity which has lower expected pecuniary returns but which also has lower exposure to high-frequency information on rejection. As shown in Figure 2.9, we would expect that individuals who place a higher weight on avoiding rejection and individuals with higher initial self-esteem beliefs will switch earliest – these characteristics are unfortunately difficult to reliably measure in my sample.

Result 1 above implies that the anticipated costs of rejection information may initially be small as agents are confident about outcomes and/or may not know how sensitive their beliefs will be to new information. However, as rejection is repeated – representing compound shocks to self-esteem – the anticipated costs of risking further rejection may increase, leading to an increase in active information avoidance.

Result 3: The drop-off in search rates is steepest after the first rejections

The drop-off in average search rates is steepest after early rejections and subsequently flattens out. In the framework above I assume that individuals are heterogeneous in terms of their preferences for avoiding rejection and in terms of their initial beliefs about their ability. As graphically illustrated in

Figure 2.9a, this implies that a large negative shock to self-esteem will not enter strongly into an individual's utility function if the weight placed on self-esteem is small or if initial beliefs about ability are modest. The upshot of this is that some individuals will be slower to stop searching than others. The patterns observed in Figures 2.7 and 2.8 suggest that the drop-off in search rates over the rounds of the game is driven by selection out of search on the part of highly rejection-sensitive individuals.

2.7.2 Alternative explanations for observed behaviour

The identification of an effect of information avoidance on search behaviour hinges on being able to rule out alternative explanations. In this subsection I discuss several of these alternative potential drivers of search behaviour and the concerns they might pose to clean identification.

2.7.2.1 The salience of rejection

In treatment conditions in which subjects are exposed to more frequent rejection, the *salience* of rejection is higher than in conditions in which subjects are exposed to less frequent rejection. Rejection may therefore function as a highly salient stimulus which interferes with behaviour which would otherwise be directed to the primary objective of the game – to maximise earnings. This interference may lead to suboptimal decision making (Bordalo et al., 2022). On face value, the potential role of salience may suggest an alternative mechanism determining withdrawal from search in the game.

However, following Bordalo et al. (2022), the role played by salience is simply that it changes the *weight* placed on a particular stimulus. In this case, the disutility of experiencing rejection may be weighted more heavily

when it is more salient in high-rejection conditions. Importantly, this does not change the *direction* in which the stimulus pulls. In other words, for rejection to play a role, the experience of rejection will need to carry some disutility in the first place.

Bordalo et al. (2022) also argue that a stimulus is particularly salient if it is 'contrasting', 'surprising', and/or 'prominent'. For high-self-esteem individuals, rejection is dissonant with their self-concept – suggesting that rejection may be salient in this context because it is *surprising*. This mechanism is closely related to the information avoidance story I focus on. In Section 2.6 I discuss how salience might be formally incorporated into the information avoidance model I sketch.

2.7.2.2 Redirecting effort in response to (perceived) signals about comparative advantage

In the real world, rejection may carry instrumentally valuable information which may help job-seekers direct their search more profitably or refine their approach to signalling employability. For instance, (repeated) rejection in one type of job may signal to a job-seeker that her comparative advantage is elsewhere and so may motivate her to redirect her search or other efforts – which may include focusing on her current job.

This poses a potential problem to this study: If players in the game believe that they are learning instrumentally valuable information from rejections, when they choose to search and are rejected they may (mistakenly) interpret this as an indication that they stand a higher chance of receiving an offer in the 'stay' task. Therefore, if players believe that they are learning about the returns to search through rejections, this may explain some of the

switching that I observe from the 'search' to the 'stay' tasks. In reality, the likelihood of receiving an offer in each task is identical, conditional on effort.

I take several measures to minimise the possibility of a player making the mistake of believing that the likelihood of receiving a payoff in the 'search' task is lower than in the 'stay' task. I am fully transparent about the way in which outcomes are determined, emphasising that the underlying outcome generating process is identical for both tasks. I provide high frequency information on players' performance, the likelihood of success, and the relationship between performance and the likelihood of success in a way that is *identical* regardless of whether a player chooses to search or not to search. This is a major focus of the session introduction and the regular comprehension checks (see Appendix 2.B). In addition, the real effort task is identical for both the 'search' and the 'stay' activities, which eliminates the possibility of subjects believing that they are better at one task than the other. In both activities, immediately after the real effort task I provide information on performance (i.e. 12/30 tasks completed) and the corresponding likelihood of receiving an offer (4 percent chance of receiving an offer). This helps emphasise that the relationship between effort and outcomes is *identical* whether one chooses to search or not to search. The only thing that differs between conditrions is the information that is provided on the realisation of the outcome – which in itself carries no information on the mechanisms determining outcomes in the game beyond that which is already provided.

Of course, it remains possible that subjects do not fully understand that the relationship between effort and outcomes is identical in the two tasks, or that they do not believe that they are. Beyond the measures that I have taken, I cannot definitively rule out this possibility.

2.7.2.3 Relevance of self-esteem beliefs in a quasi-lottery

Outcomes in the game are always primarily determined by luck. On a round-by-round basis, the probability of receiving an offer varies from 0-10 percent.[32] A bad outcome is therefore only very marginally related to performance on the real effort task. One may wonder, therefore, whether failure would be perceived as a shock to self-esteem, or merely as a self-esteem-irrelevant disappointment in the same way that losing at a lottery might be disappointing but not hurtful. If this is the case, the information avoidance effect we observe simply collapses into disappointment aversion, rather than the more specific information-on-rejection avoidance mechanism I have proposed.

As noted in Section 2.2, rejection falls under the broader category of disappointment. To be rejected is to be disappointed, while to be disappointed is not necessarily to be rejected. Thus, the presence of disappointment aversion is not a threat to the causal story I tell as long as a part of this disappointment is self-esteem-relevant.

An observation (albeit anecdotal) from the experimental sessions provides evidence that that subjects systematically over-weighted the role of performance relative to luck in determining outcomes – suggesting a role for self-esteem-relevant beliefs:

As discussed elsewhere, at the beginning of the experimental session I personally went through a detailed set of instructions and explanations, which included a clear and precise explanation of how outcomes were determined partly by a player's performance on the real effort task, and partly by luck.

[32] In practice, the probability of receiving an offer varied from 0-8 percent since the best score on the real effort task (24/30) corresponds to an 8 percent chance of receiving an offer.

Before starting each game, players needed to complete a comprehension test. One of the questions was 'What determines whether you get offered a job or not?', and players were given the response options: 'How many of the slider tasks you complete in a given round', 'Luck', and 'Both answers' (See Appendix 2.B). Players were only able to proceed once they had answered all questions correctly. While supervising this comprehension test, I observed that the overwhelming majority of players got stuck on this question – answering that 'how many of the slider tasks you complete in a given round' determines game outcomes, rather than the correct answer that both luck and performance determines outcomes.[33] This suggests that, despite explicit explanations to the contrary, players systematically placed too much emphasis on the role of performance, and too little emphasis on the role of luck. While the comprehension test may have partially corrected this tendency to over-weight performance, it is likely that this phenomenon helps explain the strength of the behavioural response to rejection in the game.

In addition, the framing of the game itself – as 'rejection' in the context of 'job search' – may also have contributed to a more self-esteem-relevant response to failure than would have been the case in a pure lottery.

Ultimately, I cannot conclusively prove that subjects shifted away from search to protect themselves from shocks to beliefs about self-esteem and not only out of disappointment aversion, though the anecdotal evidence above and my sense from supervising the laboratory games and engaging with subjects is that rejections were experienced as rejections.

[33]Unfortunately, I did not anticipate responses to these comprehension checks to be of analytical interest, and so did not record responses and so I cannot give a quantitative sense of the relative frequency of these errors.

2.7.2.4 Contamination through peer effects

To cleanly identify the effect of rejection on search behaviour, it was important that subjects did not influence each other during the game to affect search rates. Minimising the amount of communication between subjects was therefore an important consideration in the set-up and management of experimental sessions. For this reason, participants were seated with an empty seat separating each other so that they could not view each others' screens[34], and were firmly instructed not to communicate outcomes in any way.

Nevertheless, in some experimental sessions in which participants were colleagues, there was communication which could have affected players' behaviour. In particular, in some cases players would audibly celebrate when they searched and received an offer. Only players who chose to search occasionally celebrated in this conspicuous way. Players who chose not to search did not celebrate during the game because they only received information on outcomes at the end of each game – that is, after the point at which they could have influenced other players' behaviour.

This imbalance in positive public feedback may have contributed to biasing expectations *upward* towards 'search' outcomes, leading non-searchers to start searching in fear of 'missing out' on the 'jobs' that their peers were audibly getting. This may have inflated search rates, thereby attenuating the 'rejection effect' which this paper estimates. Thus, while I cannot rule out the possibility that communication may have introduced bias in the results of the game, I believe that this bias is most likely to work against my results rather than inflating the effects I estimate.

[34]For this reason too, no colour coding was used in the 'Results' screens, so that outcomes could not be discerned at a glance by a neighbor but could only be read from up close by the subject herself.

2.7.2.5 Errors in probabilistic reasoning

I find no evidence that errors in *reasoning* are driving the discouragement results. Despite the comprehension checks and explicitly stating that probabilities of receiving an offer are identical whether a player chooses to search or not, one might still worry that subjects erroneously reason that a sequence of rejections may exhibit positive recency (i.e. that a rejection is more likely to follow a sequence of rejections) or may exhibit negative recency (i.e. that an offer is more likely to follow a sequence of rejections). Positive recency would be especially concerning, since a switch away from search following a string of rejections might be motivated by an erroneous belief that a rejection is more likely to follow a rejection. If this is the case, this positive recency might be confused for information avoidance.

In the post-session survey I measured subjects' propensity to commit hot hand fallacies (positive recency) or gambler's fallacies (negative recency).[35] One would be concerned if subjects who display a propensity to commit hot hand fallacies / display positive recency are less likely to search, since this may be a potential confounder which would undermine my ability to make causal claims about information avoidance. I see no evidence of this: If anything, those subjects who have a propensity to commit hot hand fallacies appear *more* likely to search.[36] This suggests that the rejection effect

[35] I used two questions motivated by Ayton and Fischer (2004)'s finding that 'where subjects construe a sequence of outcomes as reflecting human performance [they demonstrate] positive recency [while] where subjects expect that outcomes are due to aninanimate mechanism [they demonstrate] negative recency.' The two questions I used were: *1)* 'Imagine you flip a fair coin, and get heads five times. How likely is it that it will fall on heads when you flip it a sixth time?', and, *2)* 'Imagine you ask a friend to guess a coin flip, and flip it five times. The results are: H, H, T, H, T. Your friend guesses right every time! How likely is it that she will get it right the sixth time?'.

[36] The average search rate for those who have a propensity to commit hot hand fallacies is 53 percent, compared to 51 percent for those who do not display a propensity to commit hot hand fallacies. The proportion of those who stop searching following a rejection is also

we observe is robust to the potential presence of gambler fallacy / hot-had fallacy-type reasoning.

2.8 Conclusion

I use a behavioural game implemented with young workers in South Africa to study the effect that rejection signalling has on job search. In a multi-round game I ask subjects to make a choice between two activities – to 'search for a new job' or to 'focus on their current job'. At the end of the game, subjects will either have found a 'new job' by searching, have been offered a 'contract extension' by focusing on their current job, or neither. The two activities differ in the (known) probability that the subject will receive a rejection message after each round but not in the (known) overall likelihood of receiving a job offer (whether current or new). By experimentally manipulating the expected returns (in terms of compensation incentives) and the frequency of rejection signals across treatment conditions, I am able to isolate the effect of rejection signalling on job search, holding potential confounders, including framing effects, constant.

Two main results stand out. First, subjects display strong rejection aversion and are willing to trade-off expected earnings in order to reduce their exposure to rejection. This results in a reduction in search relative to search rates of a counterfactual rejection-indifferent agent. These results are driven by the rejection messaging, rather than by framing effects.

Second, the intensity of this effect increases with exposure to rejection. The design of the game allows me to interpret rejection-avoidant behaviour

similar – at 35 percent and 34 percent respectively. These differences are not statistically significant.

as a form of active information avoidance. The trends I observe in rejection avoidant behaviour show that this information avoidance is not a static phenomenon, but changes as people are exposed to rejection. To begin with, subjects respond to pecuniary incentives as we would expect them to if they were indifferent to information on rejection. However, I find that past exposure to rejection makes people more willing to pay to avoid further exposure to rejection.

These findings have implications for mechanisms to increase job search and on-the-job search: Cases in which applications are due approximately at the same time and where outcomes are communicated after all applications are submitted may increase the total number of applications submitted, compared to cases in which search is more 'sequential' in the sense that both search and outcomes are continuous. Examples of the former include, for instance, graduate school applications and some academic job markets, while examples of the latter are more typical in low-wage white and blue collar work.

Appendix

2.A Robustness to ordering effects

The order of treatments Search and Stay was randomised across experimental sessions so as to eliminate, in expectation, the role of ordering effects on outcomes. However, becasue the number of rounds was small, this randomisation may not have netted out ordering effects in practice.

Ordering effects might be concerning since behaviour in one game might be influenced by outcomes in a previous game. For instance, if a player chose not to search in one game, and at the end was informed of a positive outcome, they might be more likely not to search in the following game. Unfortunately, because *SearchNoRej* and *SearchMoreRej* were implemented later in an add-on session, ordering was not randomly varied between *Search* and *Stay*, and *SearchMoreRej* and *SearchNoRej*. However, in Table 2.A.1 I report results for a *Search* and *Stay*, the ordering of which was randomised, so as to test for whether results are sensitive to ordering effects.

Comparing thew estimates in Table 2.A.1 when ordering fixed effects are included (top panel) to when they are excluded (bottom panel) one can see that effects appear to be robust to ordering effects. Including ordering fixed effects seems to slightly inflate coefficients relative to effects estimated without ordering fixed effects, suggesting that the main results presented in the paper may be marginal under-estimates.

Table 2.A.1: Treatment effects of rejection on search rates: Search vs Stay (inverted), with and without sequencing fixed effects

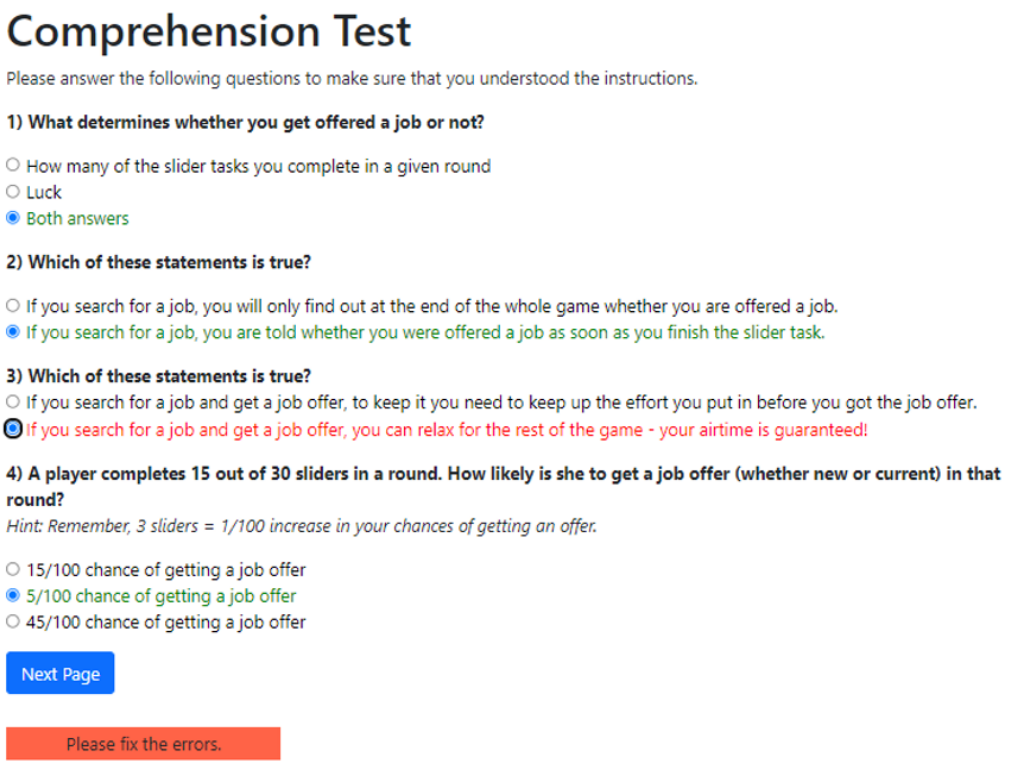
	Full sample	Excluding job-finders
	(1)	(2)
With order FE	-0.188***	-0.208***
	(0.062)	(0.064)
Constant	0.577***	0.616***
	(0.062)	(0.063)
Observations	2773	2600
Players	183	176
Without order FE	-0.119**	-0.153***
	(0.053)	(0.056)
Constant	0.676***	0.701***
	(0.024)	(0.024)
Observations	2773	2600
Players	183	176
Indiv. Fixed Effects	Yes	Yes
Past outcomes in session	Yes	Yes
Excludes job-finders	No	Yes

Notes: *p < .1; ** p < .05; *** p < .01. Standard errors, clustered at the player level, are in parentheses. The outcome variable in both specifications is a binary variable equal to 1 if a player decides to search in a given round, and equal to 0 if a player decides to 'focus on her current job'. The independent variable of interest (Search vs Stay (inverted)) is defined as a dummy equal to one if the individual was in the treatment condition of interest and equal to zero if the individual was in the comparison treatment condition. Coefficients can be interpreted as the treatment effect on search rates of being in the treatment condition of interest relative to the comparison condition. The first column reports results for the full sample, while the second column reports results only for the subsample of players who did not search and receive a job offer in the conditions under consideration.

2.B Comprehension tests

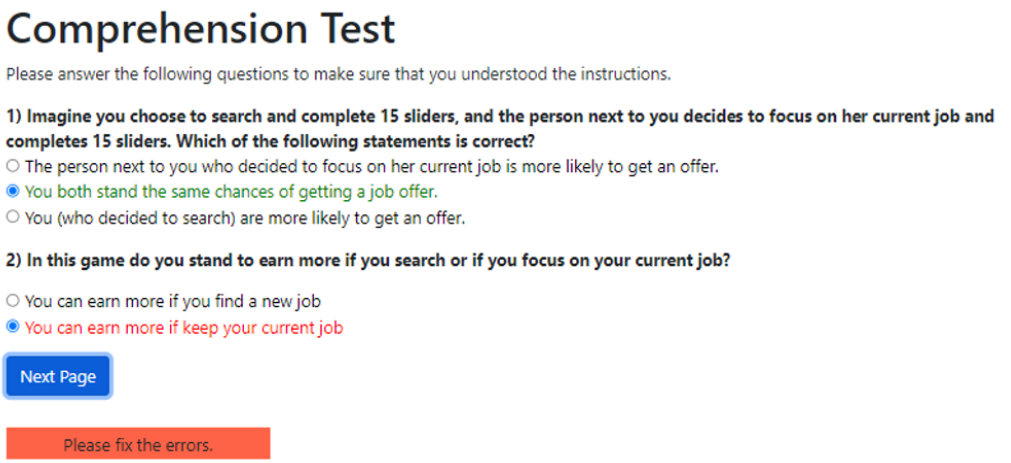
Players were required to complete the following comprehension test before beginning the *session*. As shown in the image, if players answered incorrectly, they were given the chance to correct their answers. After the comprehensions test I went over the questions, explaining why particular answers were correct and incorrect.

Figure 2.B.1: Main comprehension check



Players were required to complete versions of the following shortened comprehension test before beginning each *game*.

Figure 2.B.2: Supplementary comprehension check



2.C Full list of treatment conditions

In the table below I describe the full list of treatment conditions that were implemented during the experimental sessions run in 2022. The first four are those described in Table 2.4.1 and are those which provide the data that form the empirical basis of this paper. In addition to these four treatment conditions, I also implemented four other treatment conditions, with varying sample sizes. In these four additional treatment conditions, I intended to explore whether the provision of additional information along with information on rejection/acceptance might affect individuals' behaviour in the game. In all cases, there was a monetary incentive to *search*.

In *PeerEffortInfo*, when a player received the rejection message, she also received information on the performance of her peers (those playing the game in the same session) at the real effort task, and on whether her performance was above the peer-group average, or below the peer-group average. This was intended to provide a way to study whether information on peer performance

would be motivating or demotivating, and whether this effect might depend on whether one is above or below the peer average. In *PeerOutcomeInfo*, when a player received the rejection message, she also received information on the *outcomes* of her peers (those playing the game in the same session) – i.e. out of x number of players, how many received an offer in that round. This was intended to provide a way to study whether information on peer *outcomes* would be motivating or demotivating for those who are rejected in a given round. T7 and T8 contain no information was provided on peer outcomes or performance. Rather, the rejection messages were framed in 'kind' or 'firm' language: In the *KindRejection* treatment, players received the following message if they were rejected: 'We were impressed with your application! Unfortunately, we did not have a vacancy for you this round. Please consider applying again!'. In the *FirmRejection* treatment, players received the following message if they were rejected: 'Your performance was below this company's minimum standards. You were not our preferred candidate. You were not offered a job.'

Of the 8 treatment conditions listed below, *Stay*, *Search SearchMoreRej*, *PeerEffortInfo* and *PeerOutcomeInfo* were pre-specified in the pre-analysis plan (see Appendix). During the period in which the experiment was run, the remaining three treatment conditions were added (with the result that sample sizes for these are smaller). *SearchNoRej* was added to provide a means of testing for robustness to framing effects. The *KindRejection* and *FirmRejection* treatments were added to explore whether players' discouragement in response to rejection might be moderated or accentuated by the tone in which rejection is communicated.

However, data from the final four treatment conditions in the table below are not included in the main paper for two main reasons: First, I felt that

the question of the relationship between peer-information and behavioural responses to rejection was a distraction from the main focus of the paper, which is on the identification of a rejection effect. Secondly, I was not satisfied with the implementation of these treatment arms: While running the sessions, it appeared that participants quickly scanned the outcome information to determine whether or not they had been rejected. Beyond this, it did not appear that they paid any attention to the informational content of the rejection messages. Partly, this may have been because of the time pressure in the game, and partly because individuals were interested primarily in the outcome and not much else beyond this. The behaviour in these four 'information' arms was largely similar to that in the *Search* condition, though given the concerns raised above, it is difficult to interpret the null effect of information provision and so results are omitted.

Table 2.C.1: All treatments

Treatment	**Payoff: finding job (ZAR)**	**Payoff: keeping job (ZAR)**	**Rejection signals when searching?**	**Extra information when rejected?**	**Round-by-round rejection rate (median effort)**	**n. rounds**	**Players**
1. Stay	**50**	55	YES	No	96%	8	183
2. Search	55	R0	YES	No	96%	8	179
3. SearchNoRej	55	50	**NO**	No	96%	8	109
4. SearchMoreRej	55	50	YES	No	**98%**	**16**	109
5. PeerEffortInfo	55	50	YES	**On performance relative to peers**	96%	8	60
6. PeerOutcomeInfo	55	50	YES	**On offers recieved by searching peers**	96%	8	147
7. KindRejection	55	50	YES	**Rejection message is "encouraging"**	96%	8	38
8. FirmRejection	55	50	YES	**Rejection message is "discouraging"**	96%	8	56

Notes: The round-by-round rejection rate is, for illustrative purposes, set at the (approximate) observed median effort (12/30 sliders completed). The probability of receiving an offer in any given round is determined by the player's performance in the effort task in that round, and ranges from 0% to 10% in the 8-round games, and 0% to 5% in the 16 round game. At median effort, the *cumulative* probability of receiving an offer over the course of the game is approximately equal in Treatments 1-3 (27.9%) and Treatment 4 (27.6%).

In T5 T6 T7 and T8, players received "extra" information if they were rejected: In the *PeerEffortInfo* treatment players were told whether they had performed better or worse than their peers in the real effort task. In the *PeerOutcomeInfo* treatment players were told how many other players in their "room" (a group of between 8 and 12 players, grouped together in each session) had received an offer in the session. In the *KindRejection* treatment players received the following message if they were rejected: "We were impressed with your application! Unfortunately, we did not have a vacancy for you this round. Please consider applying again!". In the *FirmRejection* treatment players received the following message if they were rejected: "Your performance was below this company's minimum standards. You were not our preferred candidate. You were not offered a job."

2.D Pre-analysis plan

The experiment was pre-registered on the AEA RCT registry, and can be found here.

Of the four treatment conditions which form the empirical basis of this paper (listed in Table 2.4.1), three (*Stay*, *Search* *SearchMoreRej*) were pre-specified. One treatment condition analysed in the main paper that was not pre-specified, for reasons described in Appendix 2.C. There were also two treatment conditions listed in the pre-analysis plan (*PeerOutcomeInfo*) which were not analysed in the main paper, again for reasons described in Appendix 2.C.

2.E Correlates of search behaviour and discouragement

The survey data collected during the experiment allows me to explore some of the demographic, labour market and psychological correlates of search behaviour in the laboratory game. Table 2.E.1 reports multivariate regression coefficients of a range of demographic, labour market and psychological variables on two outcome variables representing the likelihood of being 'discouraged' by a prior rejection. The first outcome variable is defined as the proportion of all rounds *following a rejection in the prior round* in which an individual chose *not to search*.[37] For instance, suppose a particular subject was asked to make a decision between searching and not searching on 20 oc-

[37]Observations in the *Stay* and *SearchNoRej* conditions are excluded, since pecuniary returns incentivised not searching in the former, and no rejections were received in the latter.

casions in the game, 10 of which immediately followed a rejection in the prior round. If the subject chose *not* to search on six out of those 10 post-rejection occasions, her value for the outcome variable would be 0.6. A higher value therefore represents more 'discouragement'. The second outcome variable is simply the first outcome variable, weighted by the amount that subjects searched overall, this way taking into account the *duration* of the discouragement effect. If a subject is discouraged by rejection, but then switches back to searching quickly, the 'discouragement index' will be down-weighted heavily relative to the first outcome variable. If, on the other hand, a subject is discouraged by rejection and never returns to searching, the 'discouragement index' will be larger than the first outcome variable. This variable provides a more complete perspective on discouragement in the game, but has the disadvantage of being somewhat less intuitive to explain. For completeness, I include both.

Explanatory variables include basic demographic characteristics (race, gender, age and education), psychological variables – such as whether an individual attributes success to her own merit or to luck, and self-reported motivation to search – as well as labour market variables representing the subject's labour market behaviour in the real world – such as whether or not an individual is currently searching, and if employed, the individual's expectation of being retained in her current job. Variable definitions are provided in the table notes.

I run a simple OLS regression of the 'discouragement' outcome variables on these explanatory variables. Coefficients are reported in Table 2.E.1. The first specification (column 1) only includes the four demographic variables, the second specification adds the two psychological variables, and the third column includes the variable for whether or not the individual is searching

for jobs in 'real life'. The final column adds the control variable indicating whether or not respondents have an above or below median expectation of being retained in their current jobs. Because this variable is only asked of respondents who are employed in real life, this cuts the sample size by about 20%.

Compared to coloured respondents, black subjects appear to be approximately 14 percentage points more likely to stop searching following a rejection. However, this apparently large effect is somewhat diluted when we account for the fact that black subjects search more than coloured respondents overall, with the result that the coefficients on the 'discouragement index' are smaller and significant at a lower level. Older respondents appear to be more resilient to rejection, with lower rates of switching following a rejection. However, this effect disappears when we account for the total number of periods spent searching, suggesting that while older respondents are less likely to switch following a rejection, they are also less likely to come back to searching once they have switched. Gender and education do not appear to be correlated with discouragement in the lab game. This contrasts with other findings that suggest that there may be a correlation between gender and job-search persistence: for instance, in a field experiment Bapna et al. (2021) find that women are less likely to apply for further positions after having been exposed to rejection. On the other hand, my own previous work (Zizzamia, 2020) has suggested that young men in South Africa have a weaker attachment to the labour market and may therefore be more likely to be discouraged by disappointment in the waged labour market than women – who often have stronger family responsibilities.

Subjects who tend to attribute successes in their lives to their own merit – a concept related to self-efficacy and locus of control – appear to be less

likely to switch away from search following a rejection than subjects who subscribe to more luck-based explanations for success. However, this effect does not persist when we account for the number of search spells, with coefficients losing significance for the 'discouragement index'. Subjects who claimed to struggle to motivate themselves to search for jobs in real life are, unsurprisingly, more likely to be discouraged by a rejection. Interestingly, coefficients increase in size and significance when taking into account the total amount subjects search – indicating that subjects who struggle with motivation are not only more likely to cease searching following a rejection, but on average also take longer to (or are less likely to) return to search.

While a bivariate regression of the discouragement variables on search behaviour in real life ('Currently searching') returns a negative and significant coefficient (not reported here), once we include controls this result disappears, as seen in columns 3 and 4 – suggesting that other psychological and demographic factors, correlated with search behaviour in the real world, explain the bivariate association. Finally, restricting the sample to those who were employed in the real world, I find that those with high expectations of being retained at their real jobs display less resilience to rejection in the laboratory game. While it would be unsurprising to observe that those who have higher expectations of retention in real life also search less in real life, the fact that we see this search behaviour carry over into the laboratory game suggests that there may be a deeper psychological explanation for over-optimism and active avoidance of rejection information.

Table 2.E.1: Correlates of search behaviour and discouragement

	% of rejections followed by switch away from search				Discouragement Index			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Black African	0.132***	0.140***	0.140***	0.142***	0.092***	0.090**	0.090**	0.065*
	(0.039)	(0.043)	(0.044)	(0.043)	(0.033)	(0.035)	(0.035)	(0.038)
Female	-0.041	-0.047	-0.047	-0.019	-0.025	-0.029	-0.030	-0.005
	(0.035)	(0.035)	(0.035)	(0.037)	(0.030)	(0.029)	(0.029)	(0.031)
Age	-0.016***	-0.014**	-0.014**	-0.012*	-0.007	-0.005	-0.005	-0.005
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.005)
Some tertiary training	0.041	0.041	0.041	0.014	0.024	0.026	0.026	0.020
	(0.036)	(0.036)	(0.036)	(0.037)	(0.033)	(0.033)	(0.033)	(0.033)
Attributes success to own merit		-0.067**	-0.067**	-0.045		-0.046	-0.045	-0.002
		(0.033)	(0.033)	(0.035)		(0.029)	(0.029)	(0.031)
High self-efficacy		0.010	0.000	-0.034		-0.017	-0.019	-0.044
		(0.054)	(0.061)	(0.070)		(0.043)	(0.047)	(0.054)
Struggles with motivation to search for jobs (self-reported)		0.069**	0.069*	0.055		0.091***	0.095***	0.075**
		(0.034)	(0.035)	(0.038)		(0.031)	(0.033)	(0.036)
Currently searching			0.002	-0.003			0.032	0.016
			(0.048)	(0.054)			(0.040)	(0.044)
Higher than average expectation of being retained				0.082**				0.059*
				(0.034)				(0.030)
Constant	0.567***	0.536***	0.535***	0.435***	0.289**	0.231*	0.217*	0.169
	(0.133)	(0.137)	(0.138)	(0.142)	(0.125)	(0.127)	(0.128)	(0.128)
R^2	0.08	0.14	0.14	0.19	0.04	0.11	0.12	0.11
Number of observations	149	146	146	113	156	152	152	117

Notes: *p < .1; ** p < .05; *** p < .01. Robust standard errors are in parentheses. The *race variable* is a dummy equal to 1 if the subject is black and equal to zero if the subject is coloured (there were no whites in the sample). *Female* is a dummy equal to 1 if a respondent identifies as a woman, and zero if a respondent identifies as a man (there were no gender non-confirming individuals in the sample). *Age* is simply age in years. *Some tertiary education* is a dummy equal to 1 if a subject had completed any post-schooling qualification, and zero if the subject had not progressed beyond secondary school. All subjects had at least completed secondary schooling. *Attributes success to own merit* is a dummy equal to 1 if the subject is deemed to attribute instances of success in her life to her own merit. This variable is constructed as follows: Respondents were presented with the following two pairs of statements: 'I have a better chance of finding work than most of my peers. *vs* I basically have the same chance of finding a job as my peers - finding a job is mostly due to luck.' and 'The key to finding a job in South Africa is luck. *vs* The key to finding a job in South Africa is to not give up.'. Respondents were then asked to choose which between the pairs of statements they agreed with. If they agreed with the statements in which success was attributed to explanations centered on the individual, the 'success attribution' variable was coded as 1. If the subject agreed with the either of the statements in which success was attributed to explanations external to the individual, the 'success attribution' variable was coded as 0. *Self-efficacy* is measured using the New Generalised Self-efficacy Scale (G. Chen et al., 2001). *Struggles with motivation to search for jobs* is a dummy variable constructed as follows: Respondents were asked on a likert scale to indicate how strongly they agreed or disagreed with the following statement: 'I struggle to motivate myself to search for jobs'. Those who 'Agreed' or 'Strongly agreed' are coded as 1, and those who disagreed or were neutral were coded as zero. *Currently searching* is a dummy equal to 1 if a respondent was searching at the time of the survey, and zero otherwise. *Higher than expected expectation of being retained* is constructed as follows: Respondents who were employed at the time were asked how likely they believed, on a scale from 1-10, it would be for them to be retained in their current jobs. I used this variable to construct a dummy, with those above median expectations of being retained coded as 1, and those with below median expectations of being retained coded as zero.

2.F Balance

Table 2.F.1: Balance between sample that completed the top-up session and those that did not

Variable	N	(1) 0 Mean/SE	N	(2) 1 Mean/SE	T-test Difference (1)-(2)
Black African	82	0.793	121	0.826	-0.034
		(0.045)		(0.035)	
Female	82	0.476	121	0.669	-0.194***
		(0.055)		(0.043)	
Age (yrs)	82	23.537	121	23.347	0.189
		(0.329)		(0.269)	
Some tertiary education	82	0.671	121	0.669	0.001
		(0.052)		(0.043)	
Employment experience (mnths)	82	12.159	121	9.983	2.175
		(1.413)		(1.179)	
Currently employed	82	0.744	121	0.802	-0.058
		(0.048)		(0.036)	
Current earnings (ZAR)	42	3383.381	77	3464.987	-81.606
		(152.287)		(111.055)	

Notes: The value displayed for t-tests are the differences in the means across the groups. ***, **, and * indicate significance at the 1, 5, and 10 percent critical level.

2.G Self-image in the South African sample

In the first quarter of 2022 I collected data on concepts related to self-esteem and self-confidence from a group of young workers in South Africa who were all registered as interns on a national internship programme – the Youth Employment Service (YES).[38] In this survey I administered a number of modules measuring psycho-social characteristics such as grit and self-efficacy.[39] Both are related to other constructs such as self-esteem and self-confidence. I found that levels of grit and self-efficacy are extremely high in this sample, suggesting that youth have high levels of (stated) belief in their capacity to achieve goals which they set for themselves. Figure 2.G.1 shows the distribution of scores on grit and self-efficacy, illustrating that the majority of respondents score towards the higher values of both indices.[40]

These findings are lent further support by qualitative evidence collected using an exercises in which I prompted respondents to reflect on their goals and personal strengths. Dominant themes which emerged from interns' answers were resilience, patience, and optimism. Many youth emphasised narratives such as 'nothing is impossible if you set your mind to it'.[41] These

[38]Subjects who participated in the laboratory experiment were also registered as YES interns.

[39]Grit is defined as passion and perseverance for long-term goals (Duckworth and Quinn, 2009). Self-efficacy is defined as an individual's belief in his or her capacity to execute behaviors necessary to produce specific outcomes (Bandura, 1982).

[40]Neither grit nor self-efficacy exactly capture self-esteem, but I argue that alongside the qualitative data presented in this section, they help paint a picture of the self-image of individuals from which the sample is drawn.

[41]Typical quotations include: 'I learnt that I should always stay positive minded, I should be patient and trust myself.'; 'I learned that everything I wanna achieve in life I can achieve it if I give it all my 100 percent and that nothing is impossible – everything is possible if you really want it and you gotta get it'; 'You must never doubt yourself instead know what you want to achieve and by doing so you will be able to overcome every challenge you are facing. You must never give up.'; '[It is important to know] the type of person I am – hard working, persistent and hopeful.'; 'Perseverance, positive attitude, believing in myself and working hard helped me a lot.'

findings suggest that for the youth in this sample, self-belief and a conviction that hard work and perseverance beget good outcomes are central elements of their self-concept.[42]

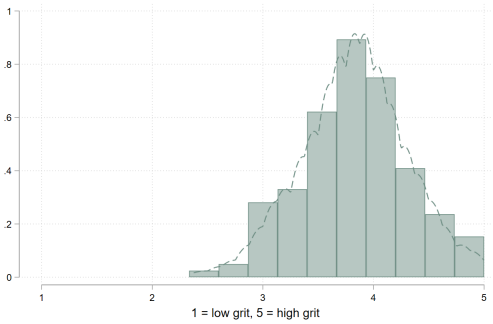
These high levels of self-confidence exist alongside very high levels of optimism about labour market prospects. Over-optimism appears to be common in similar samples: Banerjee and Sequeira (2020) find that young South African job-seekers expect to earn almost twice the median salary of individuals with similar profiles, driven by over-optimism about finding high-wage jobs. Jones and Santos (2022) find that young Mozambican job-seekers also over-estimate expected earnings by a similar magnitude.

The sample in this study is drawn from a group of interns who are employed on fixed term contracts with no guarantee of retention. In reality, less than 10 percent are kept on at their place of work in some capacity, and more than 60 percent fall into unemployment after completing the internship. Nevertheless, as show in Table 2.G.1, on average expectations for both retention and job search are very high, and either remain stable or decline only slightly. Concerns that youth are over-optimistic were also raised in interviews with high-level managers at YES. One manager explained that 'there is a strong sense from off-boarding sessions that many interns are under-prepared for the labour market: They have the attitude that 'I'm the best of the cohort so I'll be staying on when the programme ends, I'll be selected'.' To this, they added concerns that interns under-prepare for the labour market after the internship because they are present biased and treat the internship as a 'break' from searching for work – which could also be interpreted as seeking a 'break' from rejection.

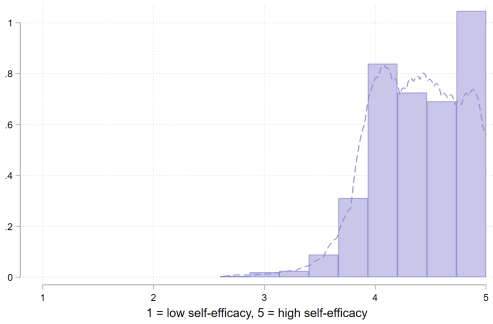
[42]This apparent over-confidence may also be a consequence of social-signalling (Burks et al., 2013).

Figure 2.G.1: Grit and self-efficacy: Distribution of scores

(a) Grit (YES sample)



(b) Self-efficacy (YES sample)



Notes: **Panel (a)** shows the distribution of grit scores for the YES baseline sample. Grit is defined as passion and perseverance for long-term goals. The Duckworth and Quinn (2009) short grit scale is used. **Panel (b)** shows the distribution of self-efficacy scores for the YES baseline sample. Self-efficacy is defined as an individual's belief in his or her capacity to execute behaviors necessary to produce specific outcomes (Bandura, 1982). The new general self-efficacy scale is used (G. Chen et al., 2001).

Table 2.G.1: Labour market expectations (confidence expressed on a scale of 10): 1 month vs 4 months after starting the internship

	Endline	Baseline	Diff. (EL - BL)	S.E.	Obs.
Retention	7.256	7.741	-0.485*	(0.275)	268
Job finding	8.639	8.647	-0.008	(0.219)	269

Notes: The table shows mean values for responses to the questions: a) 'How likely on a scale of 1 to 10 do you think it is that you would be offered a permanent position in the same organisation if you were to complete the internship?' and b) 'On a scale of 1 to 10, how much do you think this YES work experience will help you find other jobs once you have finished?'. Expectations were measured at baseline ($\tilde{1}$ month after the start of the internship) and at endline ($\tilde{4}$-6 months after the start of the internship). The third column reports the difference between baseline and endine mean values, and reports whether or not these differences are significantly different from zero (where *** $p<0.01$, ** $p<0.05$, * $p<0.10$.). The fourth and fifth columns report corresponding standard errors and the number of observations, respectively.

I propose viewing youths' puzzling and persistently high levels of optimism from this perspective: Optimism regarding job search and retention are sustained because youth are unwilling to revise their self-concept as powerful agents in their own labour market destinies. Moderating optimism downward would require an acknowledgement that despite hard work and perseverance, one may struggle to meaningfully determine one's own fate in the labour market. Apart from being profoundly demotivating, this alternative view is not consonant with what I suggest is a common and central element of the self-conception of the youth in the sample. As suggested by Kőszegi, Loewenstein, et al., 2021, this self-esteem may be 'fragile' and that individuals are consciously or subconsciously aware that major shocks to self-esteem may move them onto a lower self-esteem equilibrium. Where success is more important, Kőszegi, Loewenstein, et al. (2021) suggest that behavioural dis-

tortions to maintain a high but fragile self-esteem are likely to be largest. This interpretation is consistent with evidence that agents will actively avoid information (Golman et al., 2017) or distort memories (Huffman et al., 2022) to sustain motivated beliefs.

References

[1] Sherwin Rosen. "The economics of superstars". In: *The American economic review* 71.5 (1981), pp. 845–858.

[2] Albert Bandura. "Self-efficacy mechanism in human agency." In: *American psychologist* 37.2 (1982), p. 122.

[3] Glenn MacDonald. "The Economics of Rising Stars". In: *American Economic Review* 78.1 (1988), pp. 155–66.

[4] David M Blau and Philip K Robins. "Job search outcomes for the employed and unemployed". In: *Journal of political Economy* 98.3 (1990), pp. 637–655.

[5] Faruk Gul. "A theory of disappointment aversion". In: *Econometrica: Journal of the Econometric Society* (1991), pp. 667–686.

[6] Gilad Chen, Stanley M Gully, and Dov Eden. "Validation of a new general self-efficacy scale". In: *Organizational research methods* 4.1 (2001), pp. 62–83.

[7] Roland Benabou and Jean Tirole. "Self-Confidence and Personal Motivation". In: *The Quarterly Journal of Economics* 117.3 (2002), pp. 871–915.

[8] Peter Ayton and Ilan Fischer. "The hot hand fallacy and the gambler's fallacy: Two faces of subjective randomness?" In: *Memory & cognition* 32.8 (2004), pp. 1369–1378.

[9] Markus K Brunnermeier and Jonathan A Parker. "Optimal expectations". In: *American Economic Review* 95.4 (2005), pp. 1092–1118.

[10] Michael Rosholm and Ott Toomet. "A search model of discouragement". In: *Available at SSRN 742331* (2005).

[11] Roger Bjørnstad. "Learned helplessness, discouraged workers, and multiple unemployment equilibria". In: *The Journal of Socio-Economics* 35.3 (2006), pp. 458–475.

[12] Botond Kőszegi and Matthew Rabin. "A model of reference-dependent preferences". In: *The Quarterly Journal of Economics* 121.4 (2006), pp. 1133–1165.

[13] Alexander Brüggen and Martin Strobel. "Real effort versus chosen effort in experiments". In: *Economics Letters* 96.2 (2007), pp. 232–236.

[14] Jason Dana, Roberto A Weber, and Jason Xi Kuang. "Exploiting moral wiggle room: experiments demonstrating an illusory preference for fairness". In: *Economic Theory* 33.1 (2007), pp. 67–80.

[15] Botond Kőszegi and Matthew Rabin. "Reference-dependent risk attitudes". In: *American Economic Review* 97.4 (2007), pp. 1047–1073.

[16] Abhijit Banerjee, Sebastian Galiani, Jim Levinsohn, Zoe McLaren, and Ingrid Woolard. "Why has unemployment risen in the new South Africa? 1". In: *Economics of Transition* 16.4 (2008), pp. 715–740.

[17] David Andolfatto, Steeve Mongrain, and Gordon Myers. "Rational truth-avoidance and self-esteem". In: *Canadian Journal of Economics/Revue canadienne d'économique* 42.1 (2009), pp. 141–154.

[18] Angela Lee Duckworth and Patrick D Quinn. "Development and validation of the Short Grit Scale (GRIT–S)". In: *Journal of personality assessment* 91.2 (2009), pp. 166–174.

[19] Niklas Karlsson, George Loewenstein, and Duane Seppi. "The ostrich effect: Selective attention to information". In: *Journal of Risk and uncertainty* 38 (2009), pp. 95–115.

[20] Justine Burns, Susan Godlonton, and Malcolm Keswell. "Social networks, employment and worker discouragement: Evidence from South Africa". In: *Labour Economics* 17.2 (2010), pp. 336–344.

[21] Alan B Krueger and Andreas Mueller. "Job search and unemployment insurance: New evidence from time use data". In: *Journal of Public Economics* 94.3-4 (2010), pp. 298–307.

[22] Gary Charness and Peter Kuhn. "Lab labor: What can labor economists learn from the lab?" In: *Handbook of labor economics* 4 (2011), pp. 229–330.

[23] David Gill and Victoria Prowse. "A structural analysis of disappointment aversion in a real effort competition". In: *American Economic Review* 102.1 (2012), pp. 469–503.

[24] Alan Krueger and Andreas Mueller. "The lot of the unemployed: a time use perspective". In: *Journal of the European Economic Association* 10.4 (2012), pp. 765–794.

[25] Stephen V Burks, Jeffrey P Carpenter, Lorenz Goette, and Aldo Rustichini. "Overconfidence and social signalling". In: *Review of Economic Studies* 80.3 (2013), pp. 949–983.

[26] Kory Kroft, Fabian Lange, and Matthew J Notowidigdo. "Duration dependence and labor market conditions: Evidence from a field experiment". In: *The Quarterly Journal of Economics* 128.3 (2013), pp. 1123–1167.

[27] Emily Oster, Ira Shoulson, and E Ray Dorsey. "Optimal expectations and limited medical testing: evidence from Huntington disease". In: *American Economic Review* 103.2 (2013), pp. 804–830.

[28] Stefan Eriksson and Dan-Olof Rooth. "Do employers use unemployment as a sorting criterion when hiring? Evidence from a field experiment". In: *American economic review* 104.3 (2014), pp. 1014–39.

[29] Johannes Spinnewijn. "Unemployed but optimistic: Optimal insurance design with biased beliefs". In: *Journal of the European Economic Association* 13.1 (2015), pp. 130–167.

[30] Arthur E. Attema, Han Bleichrodt, Yu Gao, Zhenxing Huang, and Peter P. Wakker. "Measuring Discounting without Measuring Utility". In: *American Economic Review* 106.6 (June 2016), pp. 1476–94. DOI: 10.1257/aer.20150208. URL: https://www.aeaweb.org/articles?id=10.1257/aer.20150208.

[31] Daniel Chen, Martin Schonger, and Chris Wickens. "oTree—An open-source platform for laboratory, online, and field experiments". In: *Journal of Behavioral and Experimental Finance* 9 (2016), pp. 88–97.

[32] Alberto Abadie, Susan Athey, Guido W Imbens, and Jeffrey Wooldridge. *When Should You Adjust Standard Errors for Clustering?* Working Paper 24003. National Bureau of Economic Research, Nov. 2017. DOI: 10.3386/w24003. URL: http://www.nber.org/papers/w24003.

[33] James Andreoni, Justin M Rao, and Hannah Trachtman. "Avoiding the ask: A field experiment on altruism, empathy, and charitable giving". In: *Journal of political Economy* 125.3 (2017), pp. 625–653.

[34] Jason Faberman, Andreas Mueller, Ayşegül Şahin, and Giorgio Topa. *Job Search Behavior among the Employed and Non-Employed*. Working Paper 23731. National Bureau of Economic Research, Aug. 2017. DOI: 10.3386/w23731.

[35] Ananda Ganguly and Joshua Tasoff. "Fantasy and dread: The demand for information and the consumption utility of the future". In: *Management Science* 63.12 (2017), pp. 4037–4060.

[36] Russell Golman, David Hagmann, and George Loewenstein. "Information avoidance". In: *Journal of economic literature* 55.1 (2017), pp. 96–135.

[37] Kim Ingle and Cecil Mlatsheni. *The extent of churn in the South African youth labour market: Evidence from NIDS 2008-2015*. Working Paper 201. Southern Africa Labour and Development Research Unit, 2017.

[38] Gerard J van den Berg and Arne Uhlendorff. "14 Economic Job Search and Decision-Making Models". In: *The Oxford Handbook of Job Loss and Job Search* (2018), p. 243.

[39] Gary Charness, Uri Gneezy, and Austin Henderson. "Experimental methods: Measuring effort in economics experiments". In: *Journal of Economic Behavior & Organization* 149 (2018), pp. 74–87.

[40] Andrew Kerr. "Job flows, worker flows and churning in South Africa". In: *South African Journal of Economics* 86 (2018), pp. 141–166.

[41] Martin Abel, Ronelle Burger, Eliana Carranza, and Patrizio Piraino. "Bridging the intention-behavior gap? The effect of plan-making prompts on job search and employment". In: *American Economic Journal: Applied Economics* 11.4 (2019), pp. 284–301.

[42] David Gill and Victoria Prowse. "Measuring costly effort using the slider task". In: *Journal of Behavioral and Experimental Finance* 21 (2019), pp. 1–9.

[43] Julia Rose and Michael Rose. "Ready-made oTree apps for time preference elicitation methods". In: *Journal of Behavioral and Experimental Finance* 23 (2019), pp. 23–28.

[44] Rocco Zizzamia and Vimal Ranchhod. *Measuring employment volatility in South Africa using NIDS: 2008-2017*. Working Paper 246. Southern Africa Labour and Development Research Unit, 2019.

[45] Marianne Andries and Valentin Haddad. "Information aversion". In: *Journal of Political Economy* 128.5 (2020), pp. 1901–1939.

[46] Abhijit Banerjee and Sandra Sequeira. "Spatial mismatches and imperfect information in the job search". In: (2020).

[47] Kevin Donovan, Will Jianyu Lu, Todd Schoellman, et al. *Labor market dynamics and development*. Tech. rep. JSTOR, 2020.

[48] Katharina Momsen and Markus Ohndorf. "When do people exploit moral wiggle room? An experimental analysis of information avoidance in a market setup". In: *Ecological Economics* 169 (2020), p. 106479.

[49] Rocco Zizzamia. "Is employment a panacea for poverty? A mixed-methods investigation of employment decisions in South Africa". In: *World Development* 130 (2020), p. 104938.

[50] Sofia Bapna, Alan Benson, and Russell Funk. "Rejection Communication and Women's Job-Search Persistence". In: *Available at SSRN 3953695* (2021).

[51] Simon Jäger, Christopher Roth, Nina Roussille, and Benjamin Schoefer. *Worker Beliefs About Outside Options*. Working Paper 29623. National Bureau of Economic Research, Dec. 2021.

[52] Botond Kőszegi, George Loewenstein, and Takeshi Murooka. "Fragile Self-Esteem". In: *The Review of Economic Studies* 89.4 (Sept. 2021), pp. 2026–2060. ISSN: 0034-6527. DOI: 10.1093/restud/rdab060. eprint: https://academic.oup.com/restud/article-pdf/89/4/2026/44998593/rdab060.pdf. URL: https://doi.org/10.1093/restud/rdab060.

[53] Andreas Mueller, Johannes Spinnewijn, and Giorgio Topa. "Job seekers' perceptions and employment prospects: Heterogeneity, duration dependence, and bias". In: *American Economic Review* 111.1 (2021), pp. 324–63.

[54] Pedro Bordalo, Nicola Gennaioli, and Andrei Shleifer. "Salience". In: *Annual Review of Economics* 14 (2022), pp. 521–544.

[55] David Huffman, Collin Raymond, and Julia Shvets. "Persistent overconfidence and biased memory: Evidence from managers". In: *American Economic Review* 112.10 (2022), pp. 3141–75.

[56] Sam Jones and Ricardo Santos. "Can information correct optimistic wage expectations? Evidence from Mozambican job-seekers". In: *Journal of Development Economics* 159 (2022), p. 102987.

[57] Duk Gyoo Kim. "Clustering Standard Errors at the Session Level". In: *SSRN* (2022). URL: https://dx.doi.org/10.2139/ssrn.4139997.

[58] Statistics South Africa. *Quarterly Labour Force Survey: Quarter 2: 2022*. Statistical Report. Mar. 29, 2022. URL: https://www.statssa.gov.za/?p=15685 (visited on 12/03/2022).

[59] Laurel Wheeler, Robert Garlick, Eric Johnson, Patrick Shaw, and Marissa Gargano. "LinkedIn (to) job opportunities: Experimental evidence from job readiness training". In: *American Economic Journal: Applied Economics* 14.2 (2022), pp. 101–25.

[60] Erica Field, Robert Garlick, Nivedhitha Subramanian, and Kate Vyborny. "Why Don't Jobseekers Search More? Barriers and Returns to Search on a Job Matching Platform". Unpublished manuscript. Feb. 2023.

[61] Robert Garlick, Lukas Hensel, Andrea Kiss, and Kate Orkin. "Jobseekers' Beliefs about Comparative Advantage and (Mis)Directed Search". Unpublished manuscript. Mar. 2023.

3 | A push which is less big but just as good? Experimental evidence on the role of coaching within bundled ultra-poor graduation programmes

with **Narayan Das**, **Stefan Dercon**, **Imran Matin** and **Ashley Pople**

Abstract

We investigate the role that coaching plays within ultra-poor graduation programmes – specifically, how sensitive the welfare benefits of graduation interventions are to the intensity and the type of coaching inputs. We do so through a multi-arm field experiment implemented in Bangladesh in 2016, in which experimental variation was built into the amount of coaching (weekly vs fortnightly) and type of coaching (individual vs individual and group-based) that recipients received. We measure six-year impacts through an endline survey conducted in 2022, following the COVID-19 pandemic shock. While the programme itself does lead to large and sustained impacts on household asset holdings, earnings, land holdings, and hours worked in high-productivity activities, variation in coaching plays no discernible role in explaining the variation we observe in these outcomes. Reducing the intensity of coaching within ultra-poor graduation programmes could be an effective tool in bringing such programmes to scale.

Acknowledgements: We are very grateful to Christopher Adam, Muhammad Meki, Sandra Sequeira, David Evans, Vatsal Khandelwal, Pramila Krishnan, Verena Wiedemann, Rahpael Bradenbrink, Douglas Gollin, Dennis Egger, and participants at the Oxford CSAE Workshop and CREB ADE Conference for helpful comments and discussions. We gratefully acknowledge funding for this research from the ERINN Programme funded by the Wellspring Philanthropic Fund. Any errors are the authors'.

3.1 Introduction

One reason that people stay poor is that they start out poor (Kraay and McKenzie, 2014). If poverty is a trap, escaping it may require a 'big push' – lifting people out of poverty may enable them to sustain a trajectory that *keeps* them out of poverty. Interventions that attempt delivering such a 'big push' have included those which deliver large lump-sum cash grants (Haushofer and Shapiro, 2016), as well as 'graduation' programmes, which, to alleviate several constraints at once, bundle the transfer of a large productive asset with regular consumption support, savings facilities, and regular training/coaching. The graduation model was pioneered in 2002 by BRAC in Bangladesh with their Targeting the Ultra-Poor programme, now called the 'ultra-poor graduation programme' (UPGP). This paper studies the 2016 version of BRAC's UPGP.

BRAC's graduation interventions, and others like them, have received much scholarly attention over the last decade, with most of these studies finding large, positive and enduring impacts of the programme on outcomes ranging from health and education, to asset holdings and political involvement (Alatas et al., 2012; Banerjee, Duflo, Goldberg, et al., 2015; Bandiera et al., 2017; Bedoya et al., 2019; Banerjee, Duflo, and Sharma, 2021; Balboni et al., 2022; Banerjee, Karlan, et al., 2022). However, since these interventions are really a composite bundle, less is known about what is driving impact and through what channels. At the same time, a major draw-back of such programmes, with implications for scaling, is that they are expensive and logistically burdensome to implement. Coaching – which is traditionally delivered on a one-on-one basis at regular intervals over a period spanning up to two years – is one of the most expensive elements of many big push

interventions.

Understanding which elements of the programme are less cost-effective than others will be necessary in identifying where to scale back on the intensive margin in order to scale up on the extensive margin. At the same time, these knowledge gaps are not stopping governments and non-governmental organisations from bringing graduation programmes to scale. As of 2021, versions of the BRAC graduation model have been implemented in at least 75 countries around the world reaching more than 92 million people, and with implementation increasingly being undertaken directly by governments (Andrews et al., 2021). BRAC's own UPGP interventions have, as of 2021, reached over 2 million women in Bangladesh alone (BRAC, 2021). As long as graduation approaches are being scaled without a clear understanding of which elements of the programme are driving outcomes, implementing organisations remain blind to the trade-offs involved between impact on the intensive and extensive margin. If a 'big push' intervention could be just as good while being somewhat less 'big', this might allow implementers to reach more people, thereby increasing the programme's net impact.

This study seeks to understand the role that the form and intensity of the coaching element of BRAC's UPGP has in explaining the effects of the programme on welfare outcomes. In traditional big push interventions, one-on-one coaching and mentoring is administered to recipients through regular in-person house visits. On the original BRAC UPGP programme, the coaching component accounted for approximately 40 percent of the US$450-US$550 programme cost per recipient (Andrews et al., 2021) – a proportion of the total cost that remained approximately unchanged until 2016. In this paper we evaluate the impact of the 2016 UPGP programme, studying a cohort of approximately 8,000 eligible households in rural Bangladesh. We

ask three questions:

First, did the 2016 intervention have lasting effects on the welfare of recipient women and households when we followed up with them six years later? While other studies in Bangladesh and elsewhere (Banerjee, Duflo, Goldberg, et al., 2015; Bandiera et al., 2017; Banerjee, Duflo, and Sharma, 2021; Balboni et al., 2022, *inter alia*) have found lasting positive impacts, none studied programme effectiveness in the context of a large covariate shock such as that represented by the COVID-19 pandemic, with the result that we did not know whether big push programmes are 'fair weather' interventions, or if impacts can be sustained in the context of coordinated shocks.

Second, does changing the mode of delivery of the coaching component or reducing the amount of coaching recipients receive affect the magnitude of welfare impacts? If these less resource-intensive interventions are comparably effective at a lower cost, offering a version of the UPGP with less coaching may hold the promise of increasing impact on the extensive margin without compromising on the intensive margin. From 2016 onward, BRAC has in fact adopted a less coaching-intensive version of the UPGP as their basic model – making this paper a retrospective evaluation of the consequences of that decision.

Third, we investigate some of the potential mechanisms of action of coaching, attempting to discern whether we can detect effects of coaching within the UPGP programme on some of the intermediate outcomes through which coaching has been hypothesised to operate – for instance, through skills transfers and through encouraging programme compliance.

To answer these questions we exploit variation which was introduced in the 2016 version of the BRAC UPGP in a multi-arm field experiment. Three

treatment arms received a version of the UPGP, with variation in the type and amount of coaching delivered, while a control arm did not receive the programme. The first treatment group received the 'full' UPGP programme – including weekly coaching. The second treatment group received the same UPGP programme, but with fortnightly instead of weekly coaching, while the third treatment group received fortnightly coaching alongside fortnightly group-based coaching.

Comparing treatment arms to the control arm, we are able to evaluate the overall impact of the programme on a range of welfare indicators. By exploiting the randomised variation in the intensity (weekly vs fortnightly) and form (individual vs individual and group-based) of the coaching component, we are also able to study the marginal impact of these coaching variations by comparing outcomes across the three treatment groups. We also collect detailed survey data at endline on variables designed to measure impacts on a range of intermediate outcomes relevant to several key hypothesised mechanisms of coaching action.

We find that the UPGP programme has sizeable impacts on asset holdings, consumption, earnings, land holdings, and labour supply in higher-productivity economic activities. These impacts are sustained six years after the programme was introduced, despite the impact of COVID-19 from 2020 to 2022. At endline, depending on which of the three the UPGP treatment arms we consider, household asset holdings for the treatment groups were between 95-107 percent larger than for the control group, and the value of land owned was between 63 and 109 percent higher. Real monthly earnings of the main beneficiaries were between 26 and 31 percent higher for the treatment group compared to control, likely driven by the 57-78 percent increase in hours dedicated to livestock rearing – a higher-return economic activity than

forms of casual labour which predominated prior to the asset transfers. These aggregate effects appear be driven by a successful minority who experience large benefits in the programme – an effect which is especially evident when considering impacts on land and livestock holdings. Effects on consumption are somewhat more muted, with per capita household consumption in the treatment group being between 9 and 19 percent higher than in the control group at endline.

Where comparable indicators can be constructed, we also find that the magnitude of the effects in our study are comparable to the three-year effects estimated by Banerjee, Duflo, Goldberg, et al. (2015) in their five-country pilot study, and to the four-year effects estimated by Bandiera et al. (2017), who evaluate a previous version of the BRAC UPGP in rural Bangladesh. This is encouraging given both that the time horizon is substantially longer in our study (six compared to three or four years respectively) and that our endline data collection was undertaken after the COVID-19 pandemic. While it is not possible to estimate the effect that the COVID-19 pandemic had on programme impacts, the fact that impacts were sustained and remain comparable to those of prior studies is cause for optimism that the BRAC UPGP is not merely a 'fair weather' intervention. None of the 2016 cohort – in either treatment or control arms – were eligible for COVID-specific support through BRAC. Differential support during the COVID period should therefore not be contaminating any of the effects we observe.

However, when comparing effects *across* treatment conditions, we find little evidence that variation in the coaching component of the UPGP has any effect on welfare outcomes. Treatment effect point estimates are similar across the three treatment arms, and – apart from a significant difference in consumption between the 'weekly coaching' and the 'fortnightly coaching'

treatment groups – we fail to reject the null of no differential impact across treatment groups for the full range of welfare indicators. There is no evidence that the 'full' UPGP treatment with weekly one-on-one coaching (T1) has larger effects on asset holdings, earnings, land holdings or labour supply relative to the treatment with fortnightly coaching (T2) or the treatment which replaces half of the one-on-one coaching with group-based coaching.

Overall, our findings suggest that the marginal impact of coaching is small and does not justify the substantial costs required. To the extent that in-person coaching is a bottleneck to bringing graduation programmes to scale, reducing the quantity of coaching may be an important tool. A rough cost-benefit analysis does, however, suggest that rather than simply halving the amount of one-on-one coaching provided, it may be preferable to replace these with more cost-effective group-based coaching sessions.

Since we introduced variation only in the quantity and type of coaching received and not on whether or not a recipient received any coaching, this paper cannot answer the question of whether or not receiving any coaching has an impact on welfare outcomes. However, our results do suggest that, if coaching is effective, it displays rapidly diminishing marginal returns.

While our paper cannot speak directly to it, evidence from other studies does appear to be converging on the conclusion that coaching *is* effective. A fairly well-developed and growing body of evidence suggests that in-person coaching is an important determinant of programme impact – when coaching has been experimentally removed in similar interventions, programme impacts have tended to dissipate (Blattman et al., 2016; Burchi and Strupat, 2018; Sedlmayr et al., 2020; Banerjee, Karlan, et al., 2022; Bossuroy et al., 2022). However, while it seems that coaching does work, the channels

through which coaching operates remain less well understood.

Proponents of coaching – including BRAC itself – have often emphasised a skills-transfer hypothesis (Macmillan, 2022). According to this hypothesis, coaching plays a complimentary role to the asset transfer by training recipients on how to make productive use of the asset that they have received through the programme. Beyond business skills, coaching may also transmit life skills which may promote, for instance, improved education and health practices. Others have also proposed that coaching may have positive psycho-social benefits and may affect hope and aspirations, by, for instance, exposing recipients to mentors and role-models (Macmillan, 2022). It is also plausible that coaching simply functions as a commitment device to increase programme compliance among recipients (for instance, disincentivising the premature liquidation of the asset that was transferred) and that coaches act as a link between recipients and broader programme ecosystems which carry other benefits, such as access to healthcare and financial services (Blattman et al., 2016). We investigate whether we can find evidence for any of these potential coaching mechanisms, and provide suggestive evidence that coaching in this context may operate more as a mechanism promoting programme compliance than as a transfer of human capital through skills development.

Our paper makes a contribution to addressing one of the major bottlenecks in bringing big push interventions to scale. While the literature on the effectiveness of big push interventions in general is now fairly mature, researchers and policy-makers are now confronting the challenge of how best to bring these interventions to scale and are seeking to understand how to overcome the major bottlenecks to doing so (Andrews et al., 2021; Banerjee, Duflo, and Sharma, 2021). Coaching, being both expensive and reliant on scarce human resources, represents a major obstacle within this scaling

challenge (Al-Ubaydli, List, and D. L. Suskind, 2017; Al-Ubaydli, List, and D. Suskind, 2019). Our findings suggest that marginal reductions in the amount of coaching delivered through big push interventions may not result in reduced impact on key welfare indicators on the intensive margin, but by freeing up resources to implementers may increase extensive margin benefits. We also provide some suggestive evidence on why this might be the case: Coaching may not work primarily through transferring hard or soft skills, but may simply work because coaches provide a vital link between the individual recipient and the implementing organisation. Crucially, as long as some form of coaching is regular and reliable, the benefits of coaching may be less sensitive to the frequency and intensity with which coaching is delivered than we would expect if coaching functioned primarily as a skills-transfer mechanism.

This paper also contributes to the literature on the long-term effectiveness of big push interventions (Banerjee, Duflo, and Sharma, 2021). Specifically, this is the first paper to assess the long-term impacts of a big push intervention in the presence of a large covariate shock.

Unfortunately, in the 2016 UPGP BRAC did not collect detailed and disaggregated cost data, with the result that we cannot precisely calculate per-recipient programme costs across the three treatment arms. The cost-benefit analysis of the three treatment arms that we conduct therefore relies on fairly demanding assumptions. In addition, the estimated programme benefits rest heavily on consumption effects, which are somewhat anomalous in that (unlike other welfare outcomes we measure) the weekly coaching arm appears to perform better than the other two arms. The result is that the benefits of the weekly treatment arm are inflated relative its costs. For these reasons, we choose not to focus on the relative cost-effectiveness of the three

treatment variations and instead focus on treatment effects on a range of welfare-relevant outcomes.

The paper is structured as follows: In the following section we provide context on the graduation model in general, the 2016 BRAC Targeting the Ultra Poor programme in particular, and review evidence on the role of coaching within graduation approaches. In Section 3 we describe our experimental design, provide details on the intervention and the three treatment conditions, and describe our data. Section 4 describes our empirical strategy. In Section 5 we present our main results, comparing outcomes across treatment and control, and across treatment groups. In Section 6 we focus on the effects of the intervention and coaching variations on intermediate outcomes to test various hypotheses relevant to the role of coaching. Section 7 concludes.

3.2 The graduation model and the role of coaching

3.2.1 The graduation model

Finding cost-effective ways of reducing poverty among the ultra-poor continues to be a major policy challenge. One approach that has shown promise across different contexts is the ultra-poor 'graduation' model. These graduation approaches typically comprise of a package of interventions designed to help the ultra-poor alleviate several constraints simultaneously, in order to develop stable and sustainable livelihoods. The model typically consists of the transfer of a productive asset, regular consumption support, access to financial services, and 'coaching' in the form of entrepreneurship, health and social awareness training. Recurring elements, such as consumption support

and training, typically last for between 18 and 24 months.

The graduation model was pioneered by BRAC in rural Bangladesh in 2002. The original model, called the Targeting the Ultra Poor (TUP) and now known as the Ultra-Poor Graduation Programme (UPGP), was an intervention aimed at reaching the poorest women in rural areas who are not eligible for most other interventions, such as microfinance. Following the early success of the programme in Bangladesh, versions of the BRAC graduation model have been implemented in at least 75 countries around the world reaching more than 92 million people, and with implementation increasingly being undertaken directly by governments (Andrews et al., 2021). BRAC's own UPGP interventions have, as of 2021, reached over 2 million women in Bangladesh alone (BRAC, 2021).

Many of these interventions and early pilots have been the subject of rigorous experimental impact evaluations. An early evaluation by Banerjee, Duflo, Goldberg, et al. (2015) found that a series of big push intervention pilots led to lasting improvements in the lives of the very poor in six different countries. The program had positive effects on consumption, income, and other indicators of well-being when measured three years after the initial asset transfer. Bandiera et al. (2017), studying a scaled up version of the graduation model implemented by BRAC in Bangladesh, found that the programme had positive impacts on a range of well-being indicators, including increasing annual income and consumption, and had large positive effects on asset accumulation which were sustained (and even increased) after four and seven years. They showed that these poverty-reducing effects were driven in particular by the effect that the intervention had on recipient women's' labour supply: Access to productive assets moved women from low return occupations (maid work and casual agricultural labour) to higher return work

which required capital inputs (livestock rearing). Results like these point to a poverty-trap interpretation of the effects of the programme, in which the intervention functions as a 'big push', moving poor households onto a positive economic trajectory in which they are able to sustain an escape from poverty. Balboni et al. (2022) provide direct evidence in support of this hypothesis, finding that the benefits of the programme were especially strong and enduring for those who were above an asset threshold before receiving the intervention, but that benefits dissipated for those who had been below this threshold.

Several other studies have added to the growing consensus that graduation models can have lasting positive impacts on a range of well-being outcomes. These include, for instance, experiments in Uganda (Blattman et al., 2016), Ghana Banerjee, Karlan, et al. (2022), Malawi (Burchi and Strupat, 2018), Niger (Bossuroy et al., 2022), and Afghanistan (Bedoya et al., 2019), as well as a 10 year follow-up study in India (Banerjee, Duflo, and Sharma, 2021).

However, despite this headline evidence on the effectiveness of graduation programmes, there are two recurring (and closely related) open questions on the role of graduation approaches in sustainable poverty reduction efforts.

The first set of concerns is that, since the graduation model is a bundled intervention, evaluations such as those discussed above tend to say little about which elements within the bundle are driving impacts. Banerjee, Duflo, Goldberg, et al. (2015) and Bandiera et al. (2017) acknowledge these limitations, noting that without being able to identify the relative importance of specific elements within the graduation bundle (or interactions between elements), the graduation model itself remains a black box. Bandiera et al.

(2017) and Balboni et al. (2022), for instance, choose to focus in particular on the effect that the productive asset element of the programme has on labour supply, and pay relatively less attention to the role of coaching or consumption support. However, by some estimates, the cost of coaching is just as large as the cost of the asset transfer per recipient, making our ignorance regarding the individual drivers of impact consequential in terms of bringing graduation programmes to scale. For instance, in the 2007 BRAC UPGP cohort, the asset transfer and the coaching package were both valued at $560 per recipient (in 2007 $PPP) (Andrews et al., 2021). Our own back-of-the-envelope calculations based on 2016 programme cost data suggest that coaching accounted for approximately 40 percent of the per-recipient programme cost (see Appendix 3.C). [1]

The second set of concerns pertain to challenges in bringing graduation programmes to scale. One of the major concerns in scaling up the graduation model is its relatively higher cost compared to typical livelihood interventions, such as cash transfers (Sulaiman, 2018). Despite these challenges, governments and non-governmental organisations are pushing ahead and bringing graduation approaches to scale regardless, with graduation-style models reaching over 92 million people in 75 countries as of 2021 (Andrews et al., 2021).

As long as graduation approaches are being scaled without a clear understanding of which elements of the programme are driving outcomes, implementers remain blind to the trade-offs involved between impact on the intensive and extensive margin. Interventions which are on net cost-effective

[1]Unfortunately, BRAC did not collect detailed and disaggregated cost data in 2016, with the result that we cannot precisely estimate the relative per-participant cost of assets and coaching within the programme. Assumptions used to estimate the cost of the coachign component are described in Appendix 3.C.

may nevertheless contain elements which are ineffective. As long as implementers do not know which elements could be trimmed, they unintentionally and blindly trade-off spending resources effectively on the extensive margin for spending resources ineffectively on the intensive margin.

In addition, in many contexts – especially early and well-studied interventions – the graduation models were implemented as small-scale interventions or pilot studies (Banerjee, Duflo, Goldberg, et al. (2015), for instance). It is well known that small-scale interventions are often more effective than scaled-up versions of the same programmes (Muralidharan and Niehaus, 2017; Al-Ubaydli, List, and D. L. Suskind, 2017; Al-Ubaydli, List, and D. Suskind, 2019). This may be because maintaining the quality of an intervention as it scales can be challenging, especially if it involves human resources – which graduation programmes rely on in the form of coaching (Al-Ubaydli, List, and D. L. Suskind, 2017; Al-Ubaydli, List, and D. Suskind, 2019). Small scale interventions are also typically implemented by NGOs, while scaled-up versions are more often implemented by governments, which often face a different set of constraints and political priorities relative to NGOs. This too can affect the success of the transition from pilot to fully-scaled interventions (Bold et al., 2018).

A recent wave of research on graduation models has focused on these scaling issues. For instance, some recent studies evaluated larger-scale interventions implemented directly by government agencies – for instance, Bossuroy et al. (2022) study an intervention which was implemented by the government in Niger (although a pscho-social training element of the programme was implemented by an NGO in one of their treatment arms).

A number of other studies have attempted to 'unbundle' the graduation

model to determine which elements of the programme are driving impacts. Bossuroy et al. (2022) also contribute to this recent wave of papers, which provide evidence on the cost effectiveness of elements within the graduation model by selectively unbundling the intervention. In their case, they study a graduation intervention in Niger. In the base intervention that they study, recipients receive regular consumption support, savings groups, coaching and entrepreneurship training. In addition to this, a first treatment group also receives a lump-sum cash grant (T1), a second treatment group also receives a set of psychosocial interventions, such as socioemotional skills training and community sensitization on aspirations and social norms (T2), and a third treatment group receives both (T3). The control group only receives consumption support.[2] 18 months after the intervention, all three treatment arms displayed significant and positive effects on a range of economic and social outcomes, with no significant differences in the outcomes *across* treatment groups. However, because the psychosocial arm (T2) is substantially cheaper to implement than the other two arms, the internal rate of return of the psychosocial treatment is substantially higher than that of the capital arm (T1) or the full treatment (T3). These results suggest that a form of psychosocial coaching is an important driver of programme impact on its own terms, and not merely as a means of 'unlocking' the benefits of the asset or lump-sum cash transfer.

Banerjee, Karlan, et al. (2022) conducted an experiment in Ghana to compare a 'full' graduation programme with just an asset transfer. They found that the transfer of the asset alone does not lead to similar impacts on a range of outcomes as those observed for the full intervention, suggest-

[2]Importantly, most of the intervention, barring some of the training, was implemented by the government, with implications for the feasibility of bringing such programmes to scale (Al-Ubaydli, List, and D. L. Suskind, 2017; Al-Ubaydli, List, and D. Suskind, 2019).

ing that coaching plays an important and complimentary role within the programme.[3] Burchi and Strupat (2018), who also conduct an experiment which 'unbundles' the training and asset elements of a graduation intervention in Malawi and find similar results: a 'full' version of the intervention has a positive impact on a range of outcomes, including consumption, non-farm business activities, financial inclusion and agricultural production, while the 'asset only' arm only (and unsurprisingly) finds positive impacts on number of livestock. Sedlmayr et al. (2020) provide further evidence supporting the important role of coaching: Studying a graduation intervention in Uganda, they find that the positive effects of the programme were eroded when training and other forms of support were experimentally removed.

Blattman et al. (2016) also study a programme in post-war Uganda in which ultra-poor women receive a lump sum cash grant and once-off business training. In addition to this, one 'treatment-plus' group also received ongoing supervision (which represented 50 percent of programme costs). Ongoing supervision did increase business survival and investment in productive assets, but had no discernible impact on income or consumption relative to the regular intervention. Importantly, comparing a treatment condition which involved supervision *without advice* to a treatment which offered more frequent supervision *with advice*, no discernible impacts emerge – suggesting that the supervision effect is driven by encouraging programme compliance rather than by the content of the supervision.

In sum, recent evidence suggests that asset transfers alone are not sufficient to explain the range of positive outcomes associated with graduation

[3] At both two and three years after the intervention, Banerjee, Karlan, et al. (2022)'s 'full' graduation programme had positive and significant impacts on five out of five welfare relevant outcomes which were examined, compared to zero out of five for the asset-only treatment.

models. Training (or coaching) does appear to have some role in unlocking the productive value of the asset transfer. However, the mechanisms through which coaching operates remain unclear. Blattman et al. (2016)'s evidence appears to provide reason to believe that the 'supervisory' element is more important than the 'training' element – that is, coaching may operate more through functioning as a commitment device than as a means of transferring hard and soft skills. On the other hand, Bossuroy et al. (2022)'s results suggest that some form of psycho-social coaching add value independently of the productive assets transferred through the programme.

Improving the cost-effectiveness of graduation programmes is currently one of the key priorities for implementing institutions and funders, and reducing operational costs through less intensive coaching is considered an important avenue for cost optimisation. BRAC has pioneered some of these cost-effectiveness motivated adaptions to the graduation model: Since 2016, BRAC has halved the amount of coaching in the form of home visits that recipients receive (from weekly to fortnightly), and introduced a fortnightly group meeting. Along with other modifications to the programme[4], these changes have reduced the cost of the intervention per beneficiary from $550 in pre-2016 versions of the programme to $350 in post-2016 versions (Andrews et al., 2021).

[4]These include a general shift towards credit and away from transfers. In addition, matched savings (1:1 at up to £1.20 per month) have replaced regular consumption support, and the asset transfer is now implemented through an interest-free partial loan, to be repaid over the course of the programme. See Andrews et al. (2021) for more details.

3.2.2 The effects of coaching: Possible mechanisms and hypotheses

In this paper we use 'coaching' to encompass various instruction, training and supervision elements within graduation approaches. These elements are sometimes called 'training', 'mentoring', 'supervision', or 'coaching'. We deliberately use the latter term because it can be construed as encompassing the other elements, including training (i.e. skills transfers), mentoring (i.e. exposure to roll-models) and supervision (i.e. compliance checks), or some combination of these. We therefore use the term 'coaching' as a fuzzy concept which allows us to remain agnositc ex ante about which element is driving impact.

While coaching is a core component of the graduation model, the mechanisms through which it operates are not clear. This is partly because coaching through household visits serves multiple functions – providing information and hard skills training, being a source of encouragement, providing access to role-models, and serving as a commitment device to sustain engagement in other aspects of the programme. Coaches may also function as the link between individual recipients and the broader ecosystem of support which are provided by programme providers – allowing recipients to access other forms of support, savings and credit facilities, and so on. While experiments which have offered variants of the graduation programme without coaching have not not seen similarly positive impacts, it is unclear how much or what sort of coaching is required to unlock the benefits. For instance, it is not clear whether there are sharply diminishing returns to the amount of coaching received, or if the relationship between coaching and programme outcomes is more linear.

Two key mechanisms through which coaching is hypothesised to affect programme outcomes are the following: First, coaching may function as a means of transferring valuable skills, hard and soft, to recipients. Second, coaching may serve as a mechanism that increases engagement in and compliance with other elements of the intervention, which in turn drive impact. In BRAC's own motivation for the role of coaching within the UPGP, these two mechanisms are predominant. While there may be other mechanisms through which coaching affects welfare outcomes – such as through effects on social and financial inclusion or through psycho-social effects – in this paper we focus on the skills transfer and programme compliance hypotheses (though we do discuss some other mechanisms in Appendix 3.D).

According to the skills transfer hypothesis, coaching serves as a means of transferring human capital to recipients. Under this hypothesis, the *content* of the coaching, delivered through a structured syllabus, matters. The graduation model thus functions not only as a transfer of financial capital in the form of a productive asset, but also of human capital which is complimentary to the asset and which facilitates its productive use.

If the skills transfer hypothesis obtains, then it is plausible that the benefits to coaching will be sensitive to the *amount* of coaching received, though not necessarily in a linear fashion. Repetition, for instance, may be pedagogically valuable, while more contact time may also allow a more comprehensive syllabus to be delivered.

According to the programme engagement and compliance hypothesis, coaching serves as a mechanism which increases engagement in and compliance with other elements of the intervention. Through regular 'check ups' by programme staff, recipients of the graduation model may be less likely

to liquidate the productive asset that was transferred to them. Absent regular check-ups, present-biased and ultra-poor households may be tempted to liquidate productive assets to meet basic needs in the present, especially following negative welfare shocks, with negative long-run consequences for earnings and asset accumulation. Under this hypothesis, coaching serves as a commitment device to hold onto assets.

In addition, coaches may serve as intermediaries between the individual recipient, and the broader programme and other social and financial services. The coach may be able to provide information, advice and encouragement to recipients to make use of other services which individuals would not know about or be equipped to access without the coach. If this hypothesis obtains, it is more likely that the benefits to coaching will *not* be highly sensitive to the *amount* of coaching received as long as coaching visits are regular and anticipated.

We bring data to bare in determining which of these hypotheses may be most important in driving impacts. Because we do not have a treatment group in which no coaching was delivered, our investigation into these mechanisms remains, by necessity, somewhat speculative. In a way that is described in Section 3.6, we exploit variation in the amount of coaching between treatment groups, and variation in whether or not coaching was received between treatment and control (which is not, of course, the only difference between these groups) to say something suggestive about the mechanisms through which coaching may operate.

3.3 Experimental design

3.3.1 Study Sites and Participants

This study draws a sample from the 2016 cohort of the BRAC UPGP intervention.[5] A subsample of the women identified as eligible for the 2016 intervention were selected for the experimental evaluation, and these households were randomly allocated to treatment and control groups.

Initial selection into the programme proceeded in the following steps: First, the poorest rural *upazilas* (subdistricts) in which BRAC operates were identified using poverty mapping tools developed by the World Food Programme (Ahmad et al., 2010). Within these *upazilas*, communities with the highest concentration of poverty were identified based on the knowledge of existing BRAC field staff already operating in these areas (for instance, in BRAC's microfinance, health, and education programmes). Next, BRAC officers produced a list of the poorest households in these communities using participatory wealth ranking (PWR) exercises in which all households within a community are ranked in terms of material well-being by villagers in an open meeting [6] The PWR serves to classify all households into categories

[5]This was the 4th phase of BRAC's UPGP.

[6]The process starts with rapport building where a programme staff walks around a village with a few villagers to decide whether there is a substantial concentration of poor households in the village. Once a village is chosen, the programme staff engages with more villagers to identify a venue for the PWR. Villagers were invited for the PWR, which takes the form of a group discussion, usually on the following day. While the PWR may have differed from village to village, it usually included a group of about 40-50 villagers. A map of the village was drawn on the ground and villagers discussed the economic condition of each household as it was pointed out on the map. In the process of ascertaining economic conditions, the participants were asked to keep in mind issues such as the roofing material of the household, if any children were out of school, whether members had a steady income source, had any productive assets, et cetera. This was done for every household in the village.

such as very poor, poor, middle-class, and non-poor. Based on this classification, BRAC programme staff then administer a questionnaire to those households in the village who were classified as 'very poor' or 'poor' to generate a preliminary list of ultra-poor households. This survey serves to verify whether particular households met the inclusion and exclusion criteria set by BRAC.[7] The primary female adult in those households who are verified as eligible through this survey were then registered as beneficiaries.

The 2016 UPGP was implemented in 126 branches from 26 districts. Of these, 11 districts where the programme had never been implemented before were selected for inclusion in the experimental evaluation (see Figure 3.3.1).[8] The programme then selected eight branches from each of these selected districts. These branch offices were then randomly assigned to each of the four treatment or control groups (see Figure 3.3.1).[9] In total, 8,468 women and their households competed baseline surveys, and constitute the study sample.

Table 3.3.1 provides a descriptive snapshot of the study sample at base-

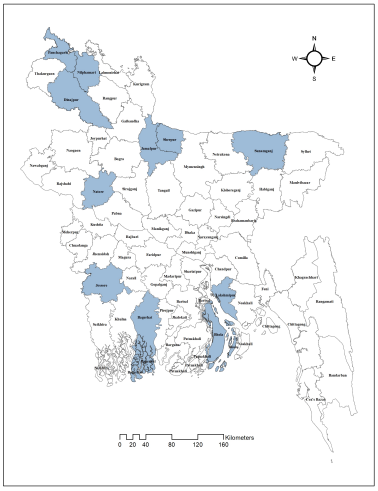
[7]The programme has five inclusion and two exclusion criteria. To be eligible, a household must fulfill three of the inclusion criteria and none of the exclusion criteria. *The inclusion criteria*: 1. The household depends on female day labour for income 2. It has less than or equal to ten decimal land 3. It has no productive assets 4. The household has children of school-going age (6-14 years) who have to work 5. The household has no adult male members. *The exclusion criteria*: 1. The household has no female members capable of working. 2. The household has one or multiple members that are participants in other government/non-government development projects.

[8]Those not included in the experimental evaluation received a version of the intervention that is described below as 'Treatment 3', which included fortnightly one-on-one coaching as well as fortnightly group meetings.

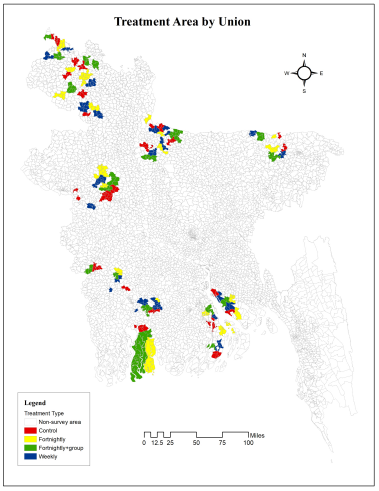
[9]The fourth phase of the intervention included two kinds of beneficiaries, 1) Special-TUP (STUP) and 2) Other-TUP (OTUP). The two packages (STUP and OTUP) differ in terms of their selection criteria and benefit provided. Both OTUP and STUP participants were selected in each branch office. The selection criteria for STUP target the poorest households while those for OTUP target relatively less poor households, and applies a lower bar in terms of eligibility criteria. The intervention evaluated in this paper concerns only the STUP beneficiaries.

Figure 3.3.1: Selection of study areas and allocation of branches to treatment and control

(a) Study districts



(b) Treatment allocation within districts



line in 2016. The average age of beneficiaries was 43 years. Half of beneficiaries' households had female heads, and the average household size was just over three members. Half of the respondents were married, half of the respondents' households owned the homestead they lived in, and seven percent of the households owned land. 23 percent of the households had an outstanding loan, and 60 percent of the households could afford two meals per day. Importantly, almost none of the respondents owned livestock, and correspondingly spent few hours a year in livestock rearing, with the majority of their labour dedicated to maid services and casual agricultural employment.

Table 3.3.1: Baseline sample statistics

	Mean	SD
Respondent age (yrs)	43.45	14.00
Hh'ld head is female (%)	0.52	0.50
Hh'ld size	3.20	1.72
Respondent is married (%)	0.53	0.50
Hh'ld owns homestead (%)	0.52	0.50
Hh'ld owns land (%)	0.07	0.26
Hh'ld has outstanding loan (%)	0.23	0.42
Hh'ld can afford 2 meals/day (%)	0.60	0.49
Number of cattle	0.01	0.16
Hours worked in livestock rearing (/year)	51.38	164.72
Hours worked in maid services (/year)	367.83	560.26
Hours worked in casual agri. employment (/year)	207.19	395.64

Notes: Statistics are based on data collected at baseline from 8,468 women included in the study in treatment and control groups.

3.3.2 The intervention

Since the BRAC UPGP is a bundled intervention, I describe below the key components of the programme. Where treatment is at the individual level, the recipient is always the main woman in the household. Unless explicitly stated otherwise, the treatment elements described below apply to the interventions received by all treatment groups in the study.

1. *Asset Transfer*:

   At the start of the programme, the main woman in the household is presented with a menu of assets, each of which can be used productively in an income generating activity. These include, for instance, livestock, land, and capital equipment. 97 percent of recipients chose a package with either cows and/or goats, and 76 percent chose a package containing cows. On average, the assets received by the woman were

valued at approximately $494 (2016 PPP). Less than 3 percent received a package that included land or non-agricultural capital goods. Details are provided in Appendix 3.A.

Prior to the asset transfers, BRAC staff provide recipients with classroom training on how to manage the asset for income generating activities. Assets are transferred within a month of this training session, and BRAC encourages beneficiaries not to liquidate the transferred assets for at least two year after receipt, although there are no hard sanctions if the asset is liquidated prematurely.

2. *Coaching*

All beneficiaries receive regular one-on-one coaching through home visits by BRAC programme staff. As described in Section 3.3.3, BRAC introduced experimental variation in whether this coaching was received on a weekly or a fortnightly basis, as well as whether a group coaching element was included.[10] For all treatment groups, these coaching sessions included life skill training, such as guidance on health and education related issues, enterprise management skills, and guidance on how to care for the asset transferred through the programme, among other things. Health topics included: family planning, prevention and treatment of waterborne diseases, non-transmittable diseases and hookworm, vaccination and nutrition. Coverage of social issues included, for instance, guidance on how to further the education of children. The participants were also given coaching meant to boost confidence, were prompted to set goals, taught about their rights and responsibilities, and were encouraged to save and to invest in simple businesses such

[10]Details on the differences in coaching between treatment groups are also described in Section 3.3.3.

as retail trading. They were also informed about different government and non-government services, such as access to credit and eligibility for support through various government rural development schemes.

An important part of the home visits was to follow-up on the status of the asset. For instance, coaches observe whether the asset is being used for income generation, and discuss issues that the beneficiary may be having with putting the asset to productive use. If the asset was livestock, coaches also helped ensure that it was healthy and receiving appropriate care. For instance, coaches typically provided training on animal husbandry best practices, such as insemination processes, vaccination, feeding, milking, and so on.

In summary, coaching comprises several elements serving different purposes and potentially operating through different mechanisms: Training is provided to transmit hard skills along with soft skill development, beneficiaries are exposed to mentors who might serve as role models, beneficiaries are able to access other social services through the intermediation of the coach, and the regular visits serve as a compliance check, ensuring that beneficiaries are using the transferred assets as intended.

Coaching sessions lasted on average approximately 30 minutes, and were continued for 18 months after the initial asset transfer. Coaches, also called BRAC programme officers, are young and highly educated: Requirements are that they have a masters degree and be under the age of 35. Each branch had on average three programme officers serving between 80 to 120 households.

3. *Group meetings (T3 only)*:

For the group meetings, a BRAC programme officer / coach convened a gathering of between 3 and 8 programme beneficiaries in a village. These group meetings were typically held on the same week as the fortnightly home visit, and were generally held in one of the programme beneficiary's houses. Beneficiaries of the T3 variant received the subsistence allowance stipend once every two weeks during these group meetings instead of during home visits. Group meetings were not explicitly set up as rotating savings and credit associations, though participants may have used the forum to set up similar informal institutions.

The topics covered during the group visits were the same in substance as those covered during the one-on-one coaching home visits. Group-based instruction and the moderation of these sessions was also undertaken by the same BRAC coaches who were responsible for the one-on-one sessions. However, while BRAC programme officers also use the group meetings to provide instruction on various social and health related matters, the group setting provided beneficiaries with a less structured forum in which to discuss concerns and raise questions. The group meeting was also a potential opportunity to create social networks among the beneficiaries. In addition, group-based coaching may be complementary to one-on-one house visits in other ways, for instance, representing a potentially more interactive and engaging approach to hard- and soft-skill acquisition, a more binding form of commitment device, and allowing recipients to learn from the experiences of other recipients.

4. *Subsistence Allowance*:

   All beneficiaries receive regular consumption support in the form of

small unconditional cash grants. These cash grants are delivered in person by a BRAC officer during either weekly or fortnightly in-person coaching sessions. The weekly value of these transfers is $6.60 in 2016 PPP. [11] Consumption support in the form of this subsistence allowance lasted for the first 40 weeks of the programme. Recipients were encouraged to save a small amount of the transfer, with a target set at approximately 10 percent.

5. *Health and Miscellaneous Social Services*:

   The beneficiaries are provided with access to medical support and are encouraged to seek the assistance of a BRAC health worker in event that a family member is sick. BRAC health workers are able to provide basic medicines, give medical advice, and refer household members to local hospitals and health centers.

6. *Community Mobilisation*:

   Informal institutions, called 'Village Poverty Reduction Committees' were set up in treatment villages.[12] These committees were created with the vision of community mobilization, linking villagers across social strata and creating mechanisms for better-off villagers to support the ultra-poor. The stated goal of these committee structures was to

[11] The total value received did not differ between beneficiaries who received weekly versus those who received fortnightly training – the latter simply received double the amount half as frequently. Since the amounts are small in absolute terms, we do not expect that the differences in the frequency of transfers would have any impact on investment or enduring impacts on consumption.

[12] Some experimental variation in this element of the UPGP bundle was built into the research design: Village Poverty Reduction Committees were *not* established in approximately half of those villages in the fortnightly coaching + group meeting treatment group (T3), with randomisation at the village rather than the subdistrict level. Details are provided later in the paper. This paper will not examine the effects of setting up these Committees, but relevant results will be tested for the robustness of including or excluding the villages without Committees.

support the economic security of UPGP beneficiaries and create an enabling environment for the beneficiaries in the village. Committees consisted of 9-11 villagers and were made up of village elites, non-elites, and UPGP beneficiaries. These committees were designed to be established within three months of the asset transfer.

Note that this is a description of the intervention design rather than the *implementation* of the intervention. Compliance was in some cases imperfect, inputs not delivered, or village institutions not set up. For this reason, our evaluation estimates intention-to-treat effects. Our endline data show that more than 90 percent of the selected beneficiaries received assistance from BRAC. The remaining 10 percent may have ceased to be eligible or may have refused to participate. Our data shows that those who did not receive benefits owned, on average, higher-valued land than those who did receive benefits – suggesting that non-compliance was correlated with baseline wealth.

On average, approximately 5 percent of households in treated villages received the intervention, though this proportion was likely somewhat variable. Because we do not collect data on non-recipients, it is not possible to estimate precisely either the general equilibrium effects of the programme or the exact proportion of a village that was covered. While we cannot rule it out, given that the proportion of villagers who received treatment was fairly small, and given prior research on general equilibrium effects of the BRAC UPGP, we do not expect the intervention to have had adverse effects on non-beneficiaries within treatment villages.[13]

[13]Bandiera et al. (2017) examine general equilibrium effects by surveying ineligible women in the villages they study. They find that the intervention moves recipients into a different occupational class, with no discernible negative earnings or labour supply effects for ineligible villagers in the *destination* occupational activities (livestock rearing), but with some positive earnings effects (10 percent increase in wages) for ineligible villagers in the *origin* occupational activities (casual maid and agricultural labour).

3.3.3 Experimental design

In the 2016 cohort, BRAC introduced experimental variation in the form and intensity of the coaching element of the bundled UPGP. BRAC's interest in overcoming barriers to bringing their programmes to scale led them to explore whether reducing the intensity of in-person coaching might alleviate a potential scaling bottleneck, and to explore what trade-offs such a modification to the programme might necessitate.

A baseline survey was conducted between April and August 2016, prior to randomisation or the implementation of the intervention. These baseline surveys were completed before disruptions associated with the seasonal monsoons. Baseline surveys were completed by 8,468 recipients and their households. The baseline survey drew a sample of between 50 and 200 eligible ultra-poor households from each branch office, which is the smallest unit in BRAC's administrative structure.

Our analysis is based on an endline survey which began in June 2022 and was completed in August 2022, which, like the baseline, was completed before the 2022 monsoon season and its associated disruptions. Excluding split households, 7,421 respondents completed the endline survey, yielding an attrition rate of 13 percent.[14]

Figure 3.3.2 summarizes the experimental design. The study covers 88 branch offices from 11 districts. These 88 branches were then randomly assigned to three treatment groups and a control group of 22 branches each

[14]We impose the most conservative definition of continuing sample members, excluding 'split households'. A 13 percent attrition rate is comparable to other TUP studies in Bangladesh, such as Bandiera et al. (2017) and Balboni et al. (2022) which use similar definitions of continuing sample members. Attrition is lower, at 11 percent, if a less conservative definition of household continuity – which would include 'split households' – is adopted.

(see Figure 3.3.1). Randomisation was stratified by district.

The intervention offered to all three treatment groups was identical in all aspects described in subsection 3.3.2 above, apart from variation introduced in the coaching element (and the nested variation built into the Village Poverty Reduction Committees in T3, described below). The variation in coaching and the size of the respective treatment arms is described below:

a) **Weekly Coaching (T1)**: Coaching in treatment arm 1 is the same as that in the pre-2016 standard UPGP intervention. In this treatment, coaching is delivered on a *weekly* one-on-one basis. 2,101 beneficiaries who were surveyed at baseline were assigned to this treatment arm, and of these, 1,855 were both successfully surveyed at endline and met the criteria we set for them to be included in the final analytical sample (i.e. once we had excluded split households).

b) **Fortnightly Coaching(T2)**: Coaching in treatment arm 2 is a stripped down version of the coaching delivered in the pre-2016 standard UPGP intervention, with *fortnightly* one-on-one coaching instead of weekly one-on-one coaching. 2,358 beneficiaries who were surveyed at baseline were assigned to this treatment arm, and of these, 2,086 were both successfully surveyed at endline and met the criteria for inclusion in the final analytical sample.

c) **Fortnightly+ Group Coaching(T3)**: Coaching in treatment arm 3 is the same as that adopted from 2016 onward, with *fortnightly* one-on-one coaching and *fortnightly group meetings*. 2,297 beneficiaries who were surveyed at baseline were assigned to this treatment arm, and of these, 1,994 were both successfully surveyed at endline and met the criteria for

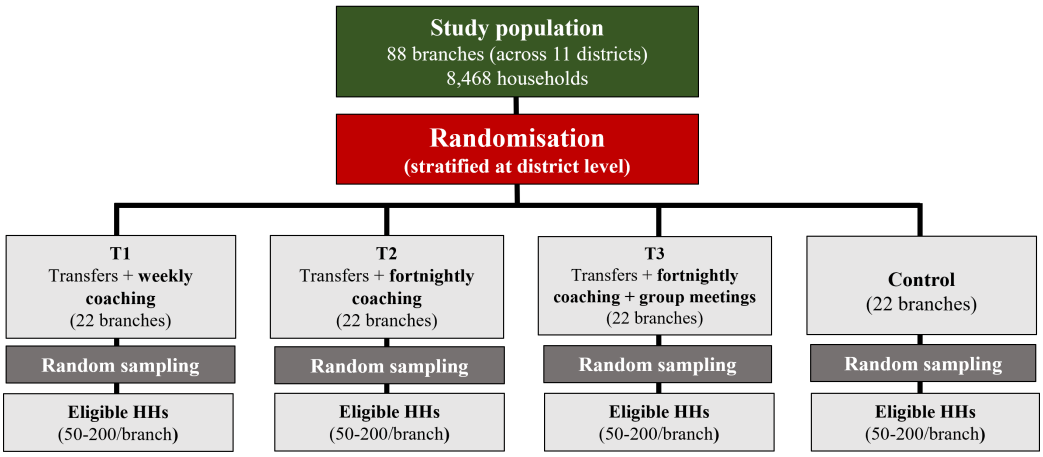
inclusion in the final analytical sample.[15]

d) **Control (C)**: The control group received no intervention and no coaching. 1,712 control group women were surveyed at baseline, and of these, 1,486 were both successfully surveyed at endline and included in the final analytical sample.

In Appendix Table 3.E we show that, apart from some significant baseline imbalance in monthly expenditure between T3 and other treatment and control conditions, randomisation achieves balance across a full range of baseline variables across all six possible treatment-vs-control and treatment-vs-treatment comparisons.

Figure 3.3.2 presents a visual representation of the experimental design and randomisation.

Figure 3.3.2: Experimental design



Notes: Experimental variation in whether or not Village Committee structures were established was nested within T3, meaning that approximately half of the endline sample in this treatment arm (1,135/2,034) need to be dropped for a strict like-for-like comparison across T1, T2 and T3.

[15]If we exclude those households in T3 who live in villages in which Village Committees were not established (as a part of the nested experiment described in 3.3.2, the sample for T3 is reduced to n = 1,029 at baseline, n = 899 at endline. We are clear throughout whether or not we apply this restriction to T3.

Among the three coaching variations, the weekly visit (T1) has the highest intensity and operational cost. This variant is the approach to coaching that was used in BRAC graduation programmes prior to 2016. This was therefore the coaching variant which was used in the interventions studied by Bandiera et al. (2017) and Balboni et al. (2022).

The second treatment arm (T2) is designed to test whether reducing the coaching intensity by half (from weekly to fortnightly home visits) results in a reduction in impact across a range of relevant welfare-related outcomes. Of the three variants, it has the lowest intensity and operational cost.

Finally, the third treatment arm (T3) includes fortnightly household visits along with a fortnightly group meeting. In terms of intensity and operational cost, T3 offers mid-point between T1 and T2. In 2016, the variation of the programme represented by T3 was the version of the programme which was delivered to those recipients who were not included in this study's sample.

The *content* of the coaching sessions remained the same across treatment arms, with there simply being more repetition of content in the weekly versus the fornightly treatment arms. In addition, the *duration* of each meeting was also the same across treatment arms. Each BRAC programme officer was responsible for visiting a fixed number of households each day, and this number was the same across treatment arms – the key difference being that the number of staff servicing households was smaller in the fortnightly coaching treatment arms compared to the weekly coaching treatment arm.

The reader will notice that no treatment arm was offered in which recipients received no coaching – which unfortunately means we can make no causal claims about the value of receiving some coaching versus receiving

the programme without coaching. BRAC chose not to offer an arm without coaching because it is their strong belief was that in-person coaching is an essential driver of programme impact, and that the benefits of the programme to beneficiaries would be substantially compromised if no in-person coaching were offered in one treatment arm. Based on it's own field research, BRAC believes that in-person coaching delivered through home visits facilitate asset accumulation by allowing programme officers to monitor assets, 'troubleshoot' issues which may arise, providing necessary guidance on the care of the asset, and creating disincentives to prematurely liquidating the asset to support consumption. BRAC was confident enough in this assumption that they did not consider a formal test of the value of the coaching element to be worth the substantial cost that would be required. In particular, BRAC believed that some of the key benefits of the coaching element specifically require *home visits*, since they provide the BRAC programme officer the opportunity to directly observe the recipients' living conditions, presence of the asset, and the standard of care the recipient is able to provide to the asset. Because home-visits were considered so important, BRAC also did not consider a treatment arm in which only group-based coaching would be offered, since this would forego the home-visit dimension of coaching.

3.3.4 Data

Since there is a lag of six years between the implementation of the programme and the evaluation, this project measures the longer-term effects of the programme on recipient households (as per Bandiera et al. (2017), Banerjee, Duflo, and Sharma (2021), and Balboni et al. (2022)). Data collection was undertaken in person by the BRAC Institute of Governance and Development. A random 10 percent sample was selected for in-person back-check

surveys.

Appendix 3.E includes tests for differential attrition across the exhaustive combination of treatment groups. There does not appear to be differential attrition between treatment groups and control, or across treatment groups. We do, however, detect some differential *panel* attrition, as shown in Appendix 3.F. Women who at baseline were younger, lower earning, and in larger households are more likely to attrit – potentially through marriage-related mobility. Since attrition is uncorrelated with treatment status we are not concerned about the implications for our analysis.

All monetary values are deflated to 2016 prices using the Bangladesh Bank's rural consumer price index and converted to 2016 purchasing power parity dollars ($PPP).

Several outcomes in our data, most notably asset holdings, are non-negative, right-skewed, and have a large probability mass at zero. Others, such as Bandiera et al. (2017) and Balboni et al. (2022) address this feature of the data by transforming the outcome variable using the natural logarithm plus a constant – $ln(X+1)$ so as to avoid dropping zero values. As Balboni et al. (2022) note, this transformation is arbitrary and may be biased, and so use an inverse hyperbolic sine transformation in robustness checks (Bellemare and Wichman, 2020). However, Mullahy and Norton (2022) show that in such cases both methods are biased, and suggest either forgoing the transformation altogether and simply using OLS or two-part models (i.e. Belloni et al. (2014)). Since Mullahy and Norton (2022) show that two-part models and OLS perform equivalently well on data of this type, I present simple OLS results for the sake of transparency. As a robustness check, in Appendix 3.B.3 I repeat the analysis using OLS on variables using the natural loga-

rithm plus a constant – $ln(X+1)$, as per Belloni et al. (2014) and Bandiera et al. (2017).

In most cases, outcome variables are Winsorized at the 99 percent level. Where Winsorization is inappropriate, this is explained in table notes. In Appendix 3.B.4 we replicate our main results with outcomes Winsorized at the 95 percent level to test for sensitivity to the level of Winsorization.

3.4 Empirical approach

Our objective is to measure *a)* the impact of (coaching variations of) the UPGP intervention on a set of welfare-related outcomes compared to the control group, and *b)* test whether there are significant differences in the impact of coaching variations of the UPGP programme (T1, T2 or T3) on the same set of outcomes compared to other variations of the programme (T1, T2 or T3).

Specifications take the following form:

$$Y_{ibt} = \beta_0 + \beta_1 T1_b + \beta_2 T2_b + \beta_3 T3_b + \gamma + \nu_d + \epsilon_{itd} \tag{3.1}$$

Where Y_{ibt} is the outcome of interest for individual i at time t in branch b. $T1_b$, $T2_b$ and $T3_b$ are dummy variables taking the value of 0 when the programme was offered not offered in a particular branch (Control), and 1 where the relevant treatment condition was offered (on a weekly basis in T1, on a fortnightly basis in T2 and on a fortnightly basis with a group meeting element in T3). A first set of results is estimated on the treatment variables with location fixed-effects and controls for date of interview.[16] γ are date-

[16]Baseline and endline data collection is implemented so as to guarantee balance between

of-interview fixed-effects for both baseline and endline surveys and ν_d are district fixed effects. The error term ϵ_{itb} is clustered by BRAC branch office, the unit of randomization.

To increase the precision of our estimates and soak up any imbalance across treatment groups, in Appendix 3.B.1 we also present a second set of results including baseline covariates. In this case, $\mathbf{X}_{t-1}$ is included as the vector of baseline controls. Following Belloni et al. (2014), we adopt the 'post-double-selection' method for selecting regressors, including first-order interaction and quadratic terms.

The β_1, β_2 and β_3 coefficients measure the effect of the relevant treatment relative to the control group. For every primary outcome, we test the null hypothesis that the programme and specific coaching delivered as a part of it has no impact, and present p-values indicating whether or not we can reject the null of equivalence for comparisons of coefficients across the treatment arms.

3.5 Main results

3.5.1 Effects on primary outcomes

Table 3.5.1 presents our main treatment effect estimates for the three treatment arms. Regression results are presented for a set of seven outcome variables: Columns one to four present results for welfare proxies – total asset holdings, consumption, earnings, and land holdings. Columns four to

treatment and control in terms of date of interview, so as not to introduce seasonality-related differences in outcomes across treatment groups. Nevertheless, we control for date of interview so as to account for any seasonality-related variation which may have been introduced.

six present estimates of effects on labour supply: yearly hours dedicated to livestock rearing, yearly hours dedicated to casual farm and non-farm (maid) labour, and total yearly hours worked. Overall, all treatment arms have a positive and significant effect across the range of outcomes after six years.

Column (1) shows evidence that six years after the initial asset transfer, asset holdings remain significantly higher in treatment compared to control conditions. Beneficiaries who receive weekly coaching (T1) increase asset holdings by 95 percent relative to control, compared to 98 percent for fortnightly coaching (T2), and 107 percent for fortnightly+group coaching (T3). These estimated treatment effects are not significantly different across treatment groups, as indicated by the p-values on T-tests of equivalence.

Column (2) reports treatment effects on consumption. Effects on consumption are somewhat more muted overall, through remain positive and significant for beneficiaries who receive weekly coaching (T1) (19 percent increase in consumption relative to control), fortnightly coaching (T2) (9 percent increase in consumption relative to control) and fortnightly+group coaching (T3) (14 percent increase relative to control). The treatment effect for T1 (19 percent) is significantly larger than the treatment effect for T2 (9 percent). The differential treatment effect across T1 and T2 is robust to the inclusion of PDS controls, as shown in Appendix 3.B.1. Apart from this, the estimated treatment effects are not significantly different across treatment groups.

Column (3) reports treatment effects on earnings. Earnings are significantly higher when comparing treatment to control, at 26 percent, 31 percent and 30 percent respectively for T1, T2, and T3. As per Bandiera et al. (2017), it is likely that this increase in earnings is explained by a shift from

relatively less productive to relatively more productive forms of work – specifically, decreasing reliance on forms of casual labour (primarily maid labour and agricultural day labour) and increasing hours worked in livestock rearing. Once again, estimated treatment effects are not significantly different across treatment groups.

The intervention also has an effect on landholding (Column (4)). We consider here only land which is owned (rather than rented) and which excludes any title that recipients have over the land on which their homestead is constructed. This land is therefore most likely to function as a capital good. The programme increases such landholdings by 109 percent, 63 percent and 83 percent respectively for T1, T2, and T3 relative to control, though once again, differences between treatment groups are not significant.

Columns (5), (6), and (7) show the effect of the intervention on labour supply: Column (5) shows that the programme had large and positive effects on hours dedicated to livestock rearing, with increases of 57 percent, 69 percent and 78 percent respectively for T1, T2, and T3 relative to control. Bandiera et al. (2017) find similar, though substantially larger effects of the intervention on hours dedicated to livestock rearing, observing increases of over 200 percent. They also find that this increase in hours dedicated to livestock rearing was the result of a reallocation of labour supply away from forms of casual labour, observing significant decreases in maid and agricultural day labour. Overall, they find that increases in hours worked overall are modest (17 percent). Our results are less striking. Column (6) shows that the intervention did not lead to a reduction in hours dedicated to casual labour, and for the fortnightly coaching group (T2), lead to a marginally significant increase in hours dedicated to casual labour, with point estimates close to zero for T1 and T3. Overall, therefore, in Column (7) we find slightly

larger point estimates of the effect of the intervention on total labour supply compared to Bandiera et al. (2017) – with effects ranging from 25 percent for the fortnightly coaching arm (T2), to 31 percent in the fortnightly coaching and group meetings arm (T3).

Table 3.5.1: Overall treatment effects of the BRAC UPGP with weekly, fortnightly, or fortnightly + group coaching

	(1) Assets	(2) Consumption	(3) Earnings	(4) Value of land	(5) Livestock hrs	(6) Casual hrs	(7) All hrs
Weekly coaching (T1)	297.780	8.349	38.909	499.919	158.281	-4.015	227.991
SE	44.975	1.960	9.763	156.679	28.551	53.292	60.473
p-value	[0.000]***	[0.000]***	[0.000]***	[0.002]***	[0.000]***	[0.940]	[0.000]***
q-value	[0.001]***	[0.001]***	[0.001]***	[0.002]***	[0.001]***	[0.156]	[0.001]***
Fortnightly coaching (T2)	307.440	4.133	47.278	285.825	191.611	72.561	193.432
SE	42.748	1.879	10.907	119.817	34.964	49.486	59.691
p-value	[0.000]***	[0.030]**	[0.000]***	[0.019]**	[0.000]***	[0.146]	[0.002]***
q-value	[0.001]***	[0.009]***	[0.001]***	[0.006]***	[0.001]***	[0.032]**	[0.002]***
Fortnightly + group coaching (T3)	336.213	6.086	45.476	381.510	216.538	-4.278	237.808
SE	44.285	1.746	9.548	128.169	38.357	51.391	64.560
p-value	[0.000]***	[0.001]***	[0.000]***	[0.004]***	[0.000]***	[0.934]	[0.000]***
q-value	[0.001]***	[0.001]***	[0.001]***	[0.002]***	[0.001]***	[0.156]	[0.001]***
Control mean at endline	314.23	44.35	152.25	457.00	276.20	546.85	769.82
% change, T1 v C	95	19	26	109	57	-1	30
% change, T2 v C	98	9	31	63	69	13	25
% change, T3 v C	107	14	30	83	78	-1	31
T1-T2 p-value	0.82	0.03	0.43	0.13	0.29	0.08	0.54
T1-T3 p-value	0.37	0.23	0.55	0.42	0.11	1.00	0.87
T2-T3 p-value	0.42	0.29	0.89	0.42	0.50	0.07	0.38
Adjusted R-squared	0.04	0.08	0.04	0.03	0.05	0.16	0.08
Number of observations	7421	7416	7445	7421	7445	7421	7445
Number of clusters	88	88	88	88	88	88	88

Notes: *p or q < .1; ** p or q < .05; *** p or q < .01. Total asset holdings is defined as the total value of all durable *household* assets. Consumption is defined as total monthly *per capita* food and non-food consumption. Earnings is defined as total monthly *household* earnings from paid economic activity. Value of land is defined as the total value of land owned by the *household*, excluding rented land and the value of the land on which the household's homestead is built. Livestock hours is defined as the total yearly hours dedicated by the respondent to livestock rearing and to other forms of agricultural self-employment. Casual hours is defined as the total yearly hours dedicated by the respondent to casual agricultural day labour and to casual maid labour. Total hours is defined as the total yearly hours dedicated to all remunerative work, including 'casual hours', 'livestock hours', and 'other hours', meaning it is not the simple sum of casual and livestock hours. All outcome variables are Winsorized at the 99 percent level. Standard errors, in parentheses, are clustered at the branch level. All regressions control for date of interview and include district fixed effects. Following Benjamini et al. (2006), we use false discovery rate (FDR) corrections and report sharpened q-values to account for multiple hypothesis testing across our primary outcome variables. While asset holdings and consumption were pre-specified as the two primary outcome variables (see Appendix 3.G), we conservatively apply FDR corrections to all seven outcome variables. Sample sizes in the four treatment and control groups are as follows: C = 1723 ; T1 = 2115 ; T2 = 2389 ; T3 = 2328

Overall, while we detect significant treatment effects across a range of economic outcomes for treatment-versus-control comparisons, when comparing effects across treatment arms we fail to detect a significant marginal impact on any of the economic outcomes we consider in Table 3.5.1 – with the exception of consumption in T1 compared to T2.

These results suggest that, on the margin, decreasing the intensity of the coaching component of the UPGP intervention has no discernible effect on the overall effectiveness of the UPGP programme. Specifically, halving the amount of coaching that beneficiaries receive relative to the the 'full' UPGP intervention with weekly coaching does not lead to lesser impact on the key economic outcomes the programme is intended to affect. The exception is that we detect lower impacts on per capita consumption for T2 compared to T1. However, effects on consumption in general are fairly modest. When a group-based coaching element is introduced, the differential effects on consumption disappear.

It bears repeating that, since all interventions included some coaching, we cannot definitively say anything about the overall effects of 'some coaching' versus 'no coaching'. However, the results above do suggest that 'some coaching' is not discernibly worse than 'more coaching'.

3.5.2 Putting effect sizes in context

Putting these effects into context, in Table 3.5.2 we replicate a set of regression results from comparable programmes as reported in Banerjee, Duflo, Goldberg, et al. (2015) and Bandiera et al. (2017). Bandiera et al. (2017) study a previous iteration of the UPGP intervention implemented at scale in Bangladesh (with weekly coaching), while Banerjee, Duflo, Goldberg, et

al. (2015) study graduation programme pilots in six countries – Ethiopia, Ghana, Honduras, India, Pakistan and Peru. Bandiera et al. (2017) replicate Banerjee, Duflo, Goldberg, et al. (2015) in order to facilitate a comparison, which allows us to simply add further 'rows' to their table of comparisons. While we observe treatment effects after six years, Bandiera et al. (2017) report results from a four year endline, and Banerjee, Duflo, Goldberg, et al. (2015) study effects after three years. Both studies assess effects of the overall programme using a single treatment-versus-control comparison. Since we have multiple treatment arms, in Table 3.5.2 we pool all treatment groups to facilitate a single treatment-versus-control comparison, while also presenting estimates specific to each treatment arm.

Regression specifications and variable definitions are, wherever possible, the same as those used by Banerjee, Duflo, Goldberg, et al. (2015) and Bandiera et al. (2017). We consider outcomes across eight dimensions: per capita consumption, food security, asset holdings, labour supply, earnings, physical health, mental health, and political awareness. All variables are index variables and are standardised as z-scores. Appendix 3.H discusses in detail the empirical approach we adopt, describes variable definitions, and pays particular attention to where there are differences between variables used here and those used by Banerjee, Duflo, Goldberg, et al. (2015) and Bandiera et al. (2017).

Overall, two main findings emerge: On the 'economic' outcomes, we find effects which are mostly situated somewhere between Banerjee, Duflo, Goldberg, et al. (2015) and Bandiera et al. (2017). Effects on per capita consumption (1), food security (2), asset holdings (3), labour supply (4), and earnings (5) are all positive and significant. Effect sizes for consumption and food security are slightly lower than those shown in Bandiera et al.

(2017) and slightly higher than those in Banerjee, Duflo, Goldberg, et al. (2015). Point estimates of the effect on overall labour supply are higher in our study than in Bandiera et al. (2017), for reasons discussed above. Our point estimates on assets are also slightly higher than in Banerjee, Duflo, Goldberg, et al. (2015) and Bandiera et al. (2017), while estimates of effects on income and revenue are slightly lower.

However, on the final three indicators, our estimated effects are small and insignificant. While Banerjee, Duflo, Goldberg, et al. (2015) and Bandiera et al. (2017) do find effects on these indicators, it should be noted that effect sizes are smaller and less significant than for the economic indicators. Bandiera et al. (2017) find effects for all three, although their mental health estimates are only marginally significant. Banerjee, Duflo, Goldberg, et al. (2015) find small effects for all three, although effects on physical health lack significance. Unlike the main economic outcomes, our mental health and political awareness indices are also not directly comparable to those used in Banerjee, Duflo, Goldberg, et al. (2015) and Bandiera et al. (2017), which may partly explain the differences we observe (see Appendix 3.H).

In sum, Tables 3.5.1 and 3.5.2 show that the 2016 version of the BRAC UPGP had positive and significant effects on a range of welfare-relevant outcomes. For many of the outcomes considered, these effects are large in absolute terms and are in the range of effects observed in other similar programmes. Importantly, effects are sustained over six years, even when compared to four or three year estimates. This is especially noteworthy in the context of a large covariate shock to rural livelihoods in the form of the COVID-19 pandemic. Previous studies, especially those in Bangladesh, were undertaken in a relatively stable growth phase, and it was not obvious ex ante whether similar effects could be expected in a more volatile economic

landscape.

Table 3.5.2: Comparison with Bandiera et al. (2017) and Banerjee, Duflo, Goldberg, et al. (2015)

	(1) Total p.c. consumption	(2) Food security index	(3) Asset index	(4) Total time spent working, main woman	(5) Income and revenue index	(6) Physical health index	(7) Mental health index	(8) Political awareness index
Overall 6 year in treatment effect	0.206***	0.116*	0.386***	0.232***	0.193***	-0.020	-0.028	-0.068
	(0.052)	(0.068)	(0.055)	(0.057)	(0.048)	(0.063)	(0.042)	(0.068)
Weekly coaching (T1)	0.291***	0.144*	0.346***	0.251***	0.194***	0.011	-0.044	-0.125
	(0.074)	(0.084)	(0.063)	(0.075)	(0.061)	(0.075)	(0.074)	(0.089)
Fortnightly coaching (T2)	0.171**	0.056	0.335***	0.196***	0.181***	0.001	0.003	-0.046
	(0.073)	(0.095)	(0.064)	(0.065)	(0.064)	(0.067)	(0.042)	(0.092)
Fortnightly + group coaching (T3)	0.164**	0.152*	0.474***	0.253***	0.204***	-0.071	-0.047	-0.037
	(0.064)	(0.077)	(0.066)	(0.064)	(0.061)	(0.068)	(0.054)	(0.063)
4 year treatment effect in Bandiera et al. (2017)	0.314***	0.256***	0.327***	0.122*	0.627***	0.108***	0.077*	0.269***
	(0.034)	(0.079)	(0.029)	(0.065)	(0.074)	(0.027)	(0.043)	(0.091)
3 year treatment effect in Banerjee, Duflo, Goldberg, et al. (2015)	0.120***	0.113*	0.249***	n/a	0.273***	0.029	0.071***	0.064***
	(0.024)	(0.022)	(0.024)		(0.029)	(0.020)	(0.020)	(0.019)

Notes: *p < .1; ** p < .05; *** p < .01. Standard errors, in parentheses, are clustered at the branch level. Following Bandiera et al. (2017) and Banerjee, Duflo, Goldberg, et al. (2015), we estimate ITT by regressing endline outcomes on baseline outcomes and randomization strata (districts). We construct indices first by defining each outcome so that higher values correspond to better outcomes. We then standardize each outcome into a z-score, by subtracting the control group mean and dividing by the control group standard deviation at the corresponding survey round. We then average all of the z-scores, and again standardize to the control group within each round. The variables used for each index are described in detail in Appendix 3.H. All indices but financial inclusion, mental health and political awareness are directly comparable, though in some cases (all noted in the appendix), the index is constructed with a more limited number of variables. Column (1) reports standardized total per capita consumption per month. The food security index in column (2) is based on survey responses regarding whether the household had a food surplus or deficit, enough food to eat over the last month, and could afford to have two meals per day most of the time during the last year. The asset index in column (3) is constructed based on the total value of productive and household assets measured in terms of a numeraire asset (goats) and standardized. Column (4) reports a standardized measure of the total time the main female household member spent in productive activities on a typical day during the past year. The incomes and revenues index in column (5) is constructed based on monthly household livestock revenue and income from agriculture, nonfarm micro-enterprises and paid labor as reported by the main female respondent. The physical health index in column (6) is constructed based on respondents' self-reported ability to perform physical tasks and the respondent's self-perception of her current health. The mental health index in column (9) is constructed based on self-reported happiness. The political awareness outcome in column (10) is simply a dummy for whether the respondent voted in the most recent election, converted to a z-score as described above. Prior to the construction of indices, outcome variables are Winsorized at the 99 percent level. The second row reports the 4-year endline estimates from Table IV in Bandiera et al. (2017), based on a sample of 6,732, while the third row reports the 3-year endline estimates from Table 3 in Banerjee, Duflo, Goldberg, et al. (2015), based on a sample that varies from 9,482 to 9,508.

3.5.3 Effects on asset holdings

A key mechanism through which the UPGP intervention is thought to operate is through setting recipients onto a trajectory in which they can sustain an escape from extreme poverty by engaging in higher productivity economic activity. The initial asset transfer is a part of this big push, but given the six year lag between the initial transfer and the endline survey, in most cases we would not expect that the original asset that recipients received (which consisted in most cases of livestock) would still be present. Rather, we are interested in how individual asset accumulation has evolved and diversified over the six intervening years.

In Table 3.5.3 we present treatment effect estimates for a range of asset classes – land holdings, livestock (goats and cows) and non-livestock assets. We find that all three treatment arms have a positive and significant effect on asset holdings, both on the extensive margin (the likelihood of holding assets in a particular class) and intensive margin (the value of assets held in each class).

However, it is also apparent that the intervention does not straightforwardly transform livelihoods by converting casual labourers into livestock farmers. While treatment households are approximately 10 percentage points more likely to own cows than control households, since the control mean is 23 percent, this means that only one third of treatment households possess cows at endline. Since cows were the most frequently transferred asset, this shows that most treatment households who received cows no longer possessed cows six years after the initial transfer. However, treatment effects on the number and the value of cows are much higher in the treatment compared to the control groups – with effects ranging between 150 and 200 percent in

both cases. This suggests that the overall effects on livestock assets are being driven by a minority of households who were able to retain and increase the number and value of their herd.

A similar dynamic is true of land holdings: Table 3.5.3 shows that treatment households are approximately 20 percent more likely to own land (or ten percentage points compared to a control mean of 57 percent), while Table 3.5.1 shows that the value of land owned is between 63-109 percent higher among treatment households than it is for control households. For these two key asset classes – land and livestock – effects are larger on the intensive than on the extensive margin – suggesting that effects are being driven by the successful minority (Balboni et al., 2022).

The value of other non-livestock assets is also higher in treatment groups, ranging from 36 percent in the fortnightly group (T2) to 67 percent in the fortnightly+group training group (T3). Households in T2 and T3 are also more likely to own goats than control households. In sum, we find evidence for significant diversification in asset holdings *away* from cows.

When comparing differences across treatment groups we find little evidence of a differential impact: Households in the weekly coaching group (T1) appear less likely to own goats (a lower-value livestock asset than cows) than those in T2 or T3, while households in the fortnightly coaching group (T2) appear to have fewer non-livestock assets than those in the other treatment groups.

Table 3.5.3: Treatment effects on assets

	(1) Owns land	(2) Owns cow(s)	(3) Number of cows	(4) Number of goats	(5) Value of cows	(6) Value of non-livestock assets
Weekly coaching (T1)	0.122***	0.127***	0.251***	0.091	222.839***	71.710***
	(0.041)	(0.019)	(0.044)	(0.055)	(37.399)	(14.328)
Fortnightly coaching (T2)	0.116***	0.165***	0.300***	0.184***	234.478***	54.901***
	(0.041)	(0.021)	(0.042)	(0.054)	(34.936)	(15.140)
Fortnightly + group coaching (T3)	0.105**	0.141***	0.276***	0.218***	215.269***	97.609***
	(0.044)	(0.021)	(0.042)	(0.052)	(37.446)	(14.569)
Control mean at endline	0.57	0.12	0.15	0.35	141.88	134.97
% change, T1 v C	22	108	164	26	157	53
% change, T2 v C	20	140	195	53	165	41
% change, T3 v C	19	120	180	63	152	72
T1-T2 p-value	0.84	0.05	0.27	0.05	0.74	0.17
T1-T3 p-value	0.61	0.44	0.56	0.01	0.84	0.03
T2-T3 p-value	0.71	0.17	0.51	0.47	0.51	0.00
Adjusted R-squared	0.06	0.05	0.04	0.05	0.03	0.05
Number of observations	7534	7421	7445	7445	7445	7421
Number of clusters	88	88	88	88	88	88

Notes: *p < .1; ** p < .05; *** p < .01. 'Owns land' is a dummy variable indicating whether or not the respondent owns land (excluding rented land and land on which the household's homestead is constructed). 'Owns cow' is a dummy variable indicating whether or not the respondent owns any cows. 'Number of cows' is a count variable indicating the number of cows owned by a household, inclusive of zeros. 'Number of goats' is a count variable indicating the number of goats owned by a household, inclusive of zeros. 'Value of cows' is a variable indicating the value of cows owned by a household in 2016 $PPP terms. 'Value of non-livestock assets' is a variable indicating the value of non-livestock assets owned by a household in 2016 $PPP terms. These assets include, for instance, agricultural capital equipment, vehicles, sheds, mobile phones, boats, fishing nets, and trees. Count and dummy variables are not winsorized, but 'Value of cows' and 'Value of non-livestock assets' are Winsorized at the 99 percent level. Standard errors are in parentheses and are clustered at the branch level. All regressions control for date of interview and include district fixed effects.

3.5.4 Cost-effectiveness

BRAC did not collect detailed and disaggregated cost data in the 2016 UPGP, with the result that we cannot precisely calculate per-recipient programme costs across the three treatment arms. The cost-benefit analysis of the three treatment arms that we conduct (see Appendix 3.C) therefore relies on fairly demanding assumptions. In addition, the estimated programme benefits rest heavily on consumption effects, which are somewhat anomalous in that (unlike other welfare outcomes we measure) the weekly coaching arm appears to perform better than the fortnightly coaching arm. Differences in consumption effects between the weekly and the fortnightly-plus-group coaching arm are not statistically significant. Nevertheless, the high point estimates of consumption in the weekly coaching treatment arm mean that the benefits are inflated relative the costs. For these reasons, we choose not to focus on the relative cost-effectiveness of the three treatment variations and instead focus on treatment effects on a range of welfare-relevant outcomes, as we have done above.

3.6 Effects of variation in coaching intensity on intermediate outcomes

In this section we more speculatively investigate whether we can learn anything about the mechanisms through which coaching does (or does not) operate. Specifically, we investigate whether coaching may operate through the transfer of hard skills and/or as a driver of compliance and programme engagement. Other secondary mechanisms are explored in Appendix 3.D.

Because we do not have a treatment arm in which the full programme

is delivered without any coaching, we cannot make direct inferences about the role of receiving any coaching versus receiving no coaching. However, by exploiting treatment versus control and treatment versus treatment comparisons, we can make some tentative inferences on the role played by coaching.

3.6.1 Hard skills

According to the 'skills transfer' hypothesis, coaching serves as a means of transferring valuable skills, hard and soft, to recipients. Under this hypothesis, the graduation model is not only a transfer of financial capital in the form of a productive asset, but also of human capital which is complimentary to the asset and which facilitates its productive use. We ask two questions: First, can we detect whether the intervention succeeds in transferring human capital? And second, if it does, can we determine whether certain coaching variations are more effective than others?

Table 3.6.1 examines eight outcome variables intended to reflect certain hard skills which respondents may have developed through the intervention, and through coaching in particular. These questions are closely related to the content of the coaching intervention and are directly based on topics covered in the coaching syllabus.

To construct these variables, respondents were asked whether they know how to perform the following tasks: Managing one's own business, finding new business opportunities, protecting businesses from harm/loss by someone else, accessing medical care, resolving an interpersonal dispute, furthering the education of children in one's household, collecting debts, and increasing the number of cows one's household owns.[17] For simplicity, variables are defined

[17]Respondents could answer: 'No, cannot', 'Maybe' or 'Definitely'. We re-coded the variables so that 'No' = 0, and 'Maybe' or 'Definitely' = 1.

so that 1 represents some degree of confidence in a respondent's ability to perform a task, and 0 represents a belief that the respondent cannot perform the task.

Table 3.6.1 displays results: For all three treatment groups and for the pooled treatment-versus-control comparisons, we find that receiving the UPGP intervention has no effect on respondents' confidence in the tasks considered. We do observe positive point estimates for the treatment effect of the pooled treatment-versus-control comparisons on six out of eight outcomes, though effects are insignificant. For fortnightly coaching (T2) we observe positive treatment effects on all outcomes and large (though insignificant) effects for confidence in 'Accessing healthcare' and 'Educating children' – surprising given the *lower* coaching intensity of this group relative to the other two treatment groups. However, none of these estimates are significant.

Treatment-versus-treatment comparisons also yield inconclusive results: Apart from a marginally significant result that fortnightly coaching (T2) is more effective than fortnightly and group coaching (T3) in imparting certain business skills, treatment-versus-treatment comparisons are significant. The result that fortnightly coaching (T2) is more effective at imparting some skills than fortnightly and group coaching (T3) is counter-intuitive, given the lower levels of coaching in T2 relative to T3. We therefore interpret these results with caution.

Not only does varying the intensity of coaching have no discernible impact on hard skill acquisition, but the UPGP intervention itself does not appear to move the needle on these outcomes relative to the control group. This is especially striking because we would expect that peoples' hard skills in specific domains increase not only through training, but also through learning-by-

doing, and treatment groups would have had more exposure not only to training, but also to learning-by-doing in business domains. Given this, the lack of treatment effects is especially stark.

Note that the dummies in Table 3.6.1 are constructed from questions in which competence is self-reported. They therefore reflect not only competence, but also *confidence*. It may be that the intervention has no effects on either competence or confidence, or, less plausibly, that it affects both in opposite directions, such that they cancel each other out.

Table 3.6.1: Likelihood of being confident in ability to perform various tasks

	(1) Collecting debts	(2) Educating children	(3) Growing herd	(4) Accessing medical care	(5) Starting business	(6) Protecting business	(7) Resolving dispute	(8) Running business
All treatment groups	0.027	0.055	0.030	0.062	-0.001	0.039	0.014	-0.020
	(0.064)	(0.065)	(0.048)	(0.067)	(0.052)	(0.065)	(0.069)	(0.054)
Weekly coaching (T1)	0.038	0.045	0.026	0.063	0.017	0.067	0.025	0.008
	(0.077)	(0.077)	(0.059)	(0.081)	(0.060)	(0.076)	(0.083)	(0.063)
Fortnightly coaching (T2)	0.070	0.128	0.053	0.107	0.040	0.056	0.085	0.016
	(0.079)	(0.079)	(0.058)	(0.080)	(0.064)	(0.078)	(0.084)	(0.067)
Fortnightly + group coaching (T3)	-0.015	0.014	0.017	0.031	-0.048	-0.003	-0.048	-0.074
	(0.070)	(0.069)	(0.056)	(0.073)	(0.055)	(0.073)	(0.075)	(0.057)
Control mean at endline	0.58	0.52	0.70	0.56	0.36	0.54	0.51	0.46
% change, all T v C	5	11	4	11	0	7	3	-4
% change, T1 v C	6	9	4	11	5	12	5	2
% change, T2 v C	12	24	8	19	11	10	17	3
% change, T3 v C	-2	3	2	5	-13	-1	-9	-16
T1-T2 p-value	0.62	0.23	0.61	0.54	0.66	0.86	0.42	0.89
T1-T3 p-value	0.46	0.66	0.87	0.67	0.20	0.32	0.33	0.13
T2-T3 p-value	0.23	0.10	0.48	0.28	0.09	0.40	0.08	0.10
Adjusted R-squared	0.06	0.05	0.05	0.08	0.04	0.03	0.05	0.04
Number of observations	7534	7534	7534	7534	7534	7534	7534	7534
Number of clusters	88	88	88	88	88	88	88	88

Notes: Outcome variables are dummy variables indicating whether or not a respondent considers herself capable of performing a particular task. From left to right, these are: Collecting debts, furthering the education of children in one's household, increasing the number of cows one's household owns, accessing medical care, finding new business opportunities, protecting businesses from harm/loss by someone else, resolving an interpersonal dispute, and managing one's own business. Respondents could answer: 'No, cannot', 'Maybe' or 'Definitely'. We re-coded the variables so that 'No' = 0, and 'Maybe' or 'Definitely' = 1. All outcome variables are Winsorized at the 99 percent level. Standard errors, in parentheses, are clustered at the branch level. All regressions control for date of interview and include district fixed effects.

Some direct measures of behaviour relevant to the skills transfer hypothesis are presented in Table 3.6.2 below. Here, we ask those women who own livestock at endline whether or not they practise a range of animal husbandry best practices, using survey instruments recommended by the FAO and World Bank (Pica-Ciamarra et al., 2011). While these measures have the advantage of being more directly relevant to behaviour, they have their own problems: these results are vulnerable to social desirability bias (treatment households know better what they *should* be doing, whether or not they are actually doing it), and to selection issues (treatment households are more likely to answer the module in the first place since they are more likely to own livestock).

With these caveats in mind, we do find that treatment women are more likely than those who own livestock in the control group to perform a range of animal husbandry best practices, such as vaccinate livestock, practise deworming, or use selective breeding practices. However, we fail to detect any significant differences across treatment groups – except for the counter-intuitive finding that those in the fortnightly coaching treatment group may be more likely to practise deworming than those in the weekly coaching treatment group. It does not appear that the *amount* of coaching received has any marginal effect on the likelihood of a recipient engaging in animal husbandry best practices. It is difficult to discern whether the overall effect of the UPGP intervention on animal husbandry best practices is due to coaching or is driven by the access provided by BRAC through the intervention to vaccinations and veterinary medicines – though the effects on controlled breeding practices suggests some role for coaching.

Table 3.6.2: Animal husbandry best practices

	(1) Practises vaccination	(2) Practises tick treatment	(3) Practises deworming	(4) Practises controlled breeding
All treatment groups	0.098**	-0.045	0.179***	0.109***
	(0.041)	(0.055)	(0.037)	(0.034)
Weekly coaching (T1)	0.098**	-0.085	0.152***	0.114***
	(0.045)	(0.061)	(0.040)	(0.037)
Fortnightly coaching (T2)	0.101**	-0.055	0.214***	0.104**
	(0.048)	(0.060)	(0.040)	(0.044)
Fortnightly + group coaching (T3)	0.097**	-0.005	0.184***	0.108***
	(0.041)	(0.056)	(0.040)	(0.037)
Control mean at endline	0.38	0.29	0.62	0.47
% change, all T v C	26	-15	29	23
% change, T1 v C	26	-29	25	24
% change, T2 v C	27	-19	35	22
% change, T3 v C	26	-2	30	23
T1-T2 p-value	0.92	0.45	0.03	0.79
T1-T3 p-value	0.96	0.09	0.22	0.84
T2-T3 p-value	0.89	0.26	0.24	0.91
Adjusted R-squared	0.03	0.05	0.04	0.01
Number of observations	2745	2745	2745	2745
Number of clusters	88	88	88	88

Notes: Outcome variables are only measured for those who currently own large ruminants – goats, sheep or cows/buffalo. Outcome variables are dummy variables indicating respectively whether or not a respondent practises vaccination, tick-treatments, de-worming, or selective breedings on their livestock. Standard errors, in parentheses, are clustered at the branch level. All regressions control for date of interview and include district fixed effects.

In sum, we see no evidence that, on the margin, the amount of coaching delivered within the UPGP has any effects on measured hard skills – either self-reported or in terms of reported behaviours. We find mixed evidence that coaching – or the UPGP bundle itself – shifts recipients' hard skills at all.

3.6.2 Coaching as a commitment device

In Section 3.2 we hypothesised that coaching may also operate by increasing programme engagement. This effect could be driven by the fact that the coach functions as a compliance check: anticipating that a coach will visit you somewhat regularly may create an incentive not to liquidate the asset prematurely, and to follow best asset management practices (for instance, to stay up-to-date with vaccinations in the case of livestock). Blattman et al. (2016) seem to suggest this 'supervision' mechanism is driving the effect of coaching in their study. In addition, at least in the Bangladeshi context, BRAC operates an extremely broad range of social, financial, health and education services in rural areas. The BRAC coach may simply function as the link between the individual recipient and this broader ecosystem – helping individuals access services they may otherwise have struggled to access or not known about. In BRAC's own motivation for the coaching component of the UPGP, based on their own field work and field observations, they stress role of coaching as a means of encouraging compliance.

It is difficult to test directly for the presence of these causal mechanisms in our study design. Partly, this is because we would not expect that this mechanism would be highly sensitive to the intensity of coaching (fortnightly or weekly) or the mode of delivery (individual or individual and group-based). For these 'supervisory' and 'intermediary' mechanisms to operate, it would

be more important that visits are regular and expected than for them to be highly frequent.

We do show, in Figure 3.6.1, that the proportion of respondents who reported liquidating the asset received through the programme is highest in the two years *after* the cessation of coaching and consumption support and the soft requirement to retain the asset. Respondents were asked at endline whether they still owned the main asset[18] that was given to them by BRAC in 2016, and if not, until what year they owned it.[19] Only about 10 percent of respondents still owned the asset in 2022. About 25 percent of respondents reported liquidating their asset while coaching was being delivered in the first two years of the programme. However, in the following two years, when home visits by BRAC programme officers had ceased, an additional 50 percent of respondents reported liquidating their asset.

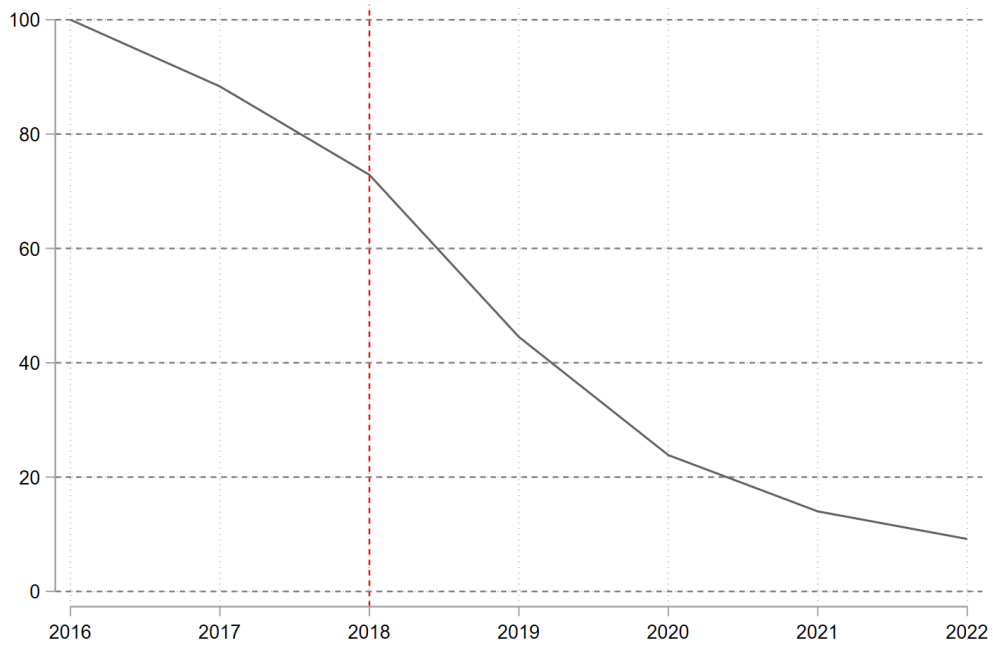
While it is impossible to attribute causation to the relationship between the end of coaching and the acceleration of asset liquidation, the pattern we observe in Figure 3.6.1 is at least consistent with the hypothesis that coaching served to delay asset liquidation among recipients. It is also worth noting that asset retention / liquidation did not differ across treatment groups – suggesting that if coaching serves to delay asset liquidation, this is not sensitive to the amount of coaching received.[20]

[18] 80 percent of the time this was a single cow, in the remaining 20 percent this was four goats)

[19] For respondents who no longer owned their asset, in 88 percent of cases they had sold the asset and only in 12 percent of cases had the asset died or been lost.

[20] If anything, premature asset liquidation was higher for T1 than for T2 and T3, though differences are not significant.

Figure 3.6.1: Proportion of recipients retaining their asset by a given year



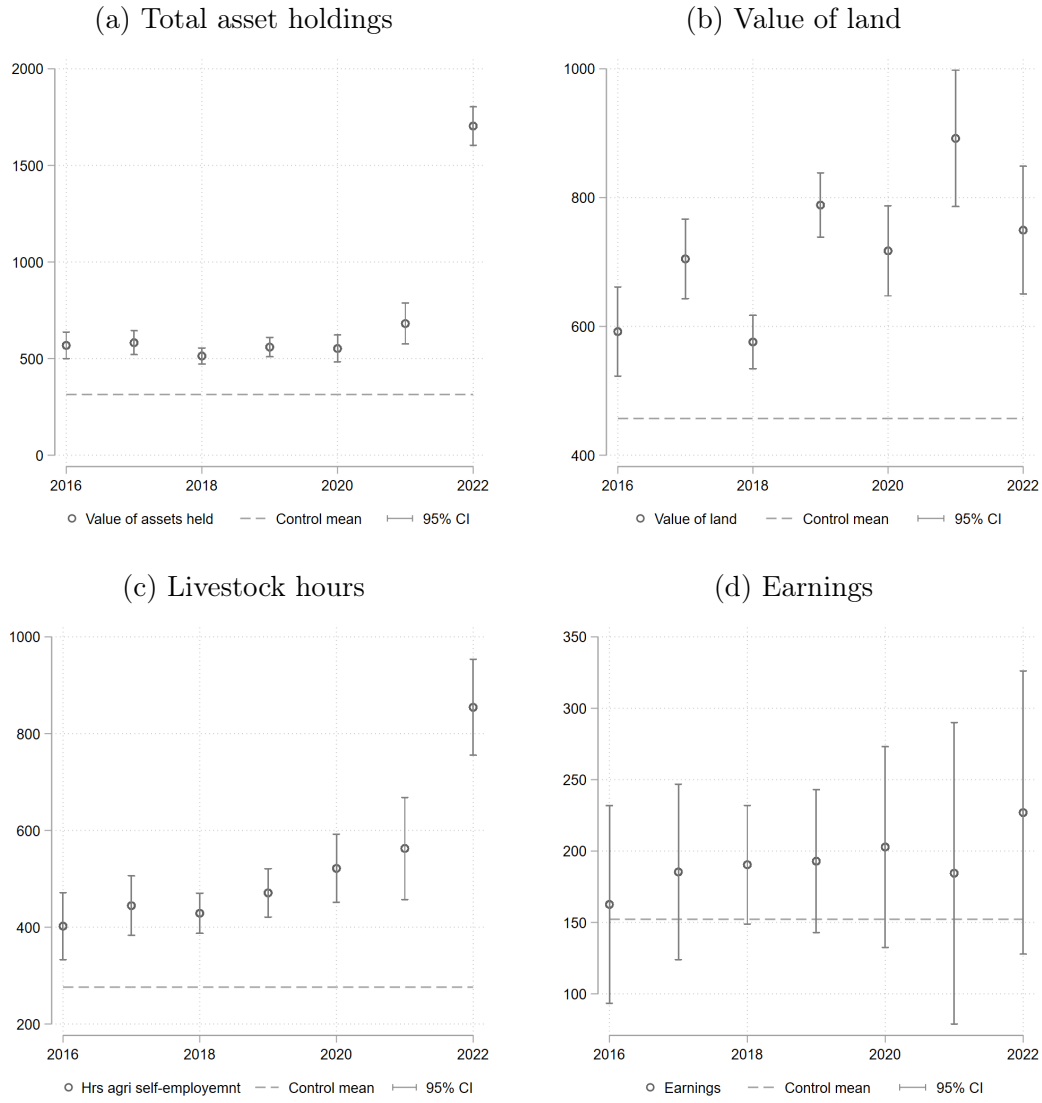
Notes: The red dotted vertical line approximately represents the programme end – the point at which the soft requirement not to liquidate the asset lapsed. The figure reports the proportion of those in T1, T2 and T3 who, by a given year, reported retaining the asset that was transferred to them through the UPGP intervention. Respondents were asked, if they reported having received an asset, the year until which the asset was held.

In Figure 3.6.2 we show that the point at which a recipient liquidates their asset is correlated with a number of the main outcomes considered in this paper. In Figure 3.6.2 we plot the relationship between the duration for which the BRAC asset was held and total household asset holdings (3.6.2a), the value of non-homestead land held by the household (3.6.2b), the total yearly hours dedicated by the respondent to livestock rearing (3.6.2c), and total monthly household earnings from paid economic activity (3.6.2d).

Those in the treatment group are in most cases better-off than those in the control group regardless of when they liquidated their asset. However, there is a strong relationship between delaying liquidation (or not liquidating

assets) and the outcomes considered. Since the value of household assets is heavily influenced by whether or not a household owns livestock, the asset holdings of those who retained the original asset by 2022 was approximately three times higher than those who had liquidated the asset. Since labour supply is shaped by access to livestock in this context (Bandiera et al., 2017), it is unsurprising that those who retained the asset by 2022 also dedicated significantly more hours to livestock rearing than those who liquidated their assets. While the relationship is less clear between asset retention and earnings, we also observe the expected upward slope in earnings. The relationship between asset retention and land holdings is interesting – it appears that recipients are diversifying their asset base by accumulating land (as noted in the discussion of Table 3.5.3), but that the value of the land that is accumulated is greater among those who delay (or do not) liquidate their assets. While some of the asset diversification may be funded by the sale of the BRAC asset, it appears that on average this is not the dominant driver of asset accumulation. It is thus likely that asset liquidation is primarily geared towards consumption.

Figure 3.6.2: Relationship between BRAC asset retention and livelihood outcomes



Notes: Total asset holdings is defined as the total value of all durable *household* assets. Value of land is defined as the total value of land owned by the *household*, excluding rented land and the value of the land on which the household's homestead is built. Earnings is defined as total monthly *household* earnings from paid economic activity. Livestock hours is defined as the total yearly hours dedicated by the respondent to livestock rearing and to other forms of agricultural self-employment. All outcome variables are Winsorized at the 99 percent level. Error bars represent 95% confidence intervals.

In summary, there is suggestive evidence that the coaching delivered through home visits during the first two years of the UPGP programme acted as a break on premature asset liquidation. When we also consider the close association between welfare outcomes and the length of time over which the asset is retained, it appears plausible that coaching does play an important role in the UPGP by encouraging programme compliance. The *amount* of coaching received does not seem to determine differences in asset retention.

3.7 Discussion and conclusion

In this paper we set out to investigate the role that coaching plays within ultra-poor graduation programmes. Specifically, we seek to understand how sensitive the welfare benefits of graduation interventions are to the intensity and the type of coaching inputs. We do so through a multi-arm field experiment implemented in Bangladesh in 2016, in which experimental variation was built into the amount of coaching (weekly vs fortnightly) and type of coaching (individual vs individual and group-based) that recipients received. We measure six-year impacts through an endline survey conducted in 2022.

Variations in the intensity and type of coaching do not seem to explain any meaningful variation in the main welfare outcomes we consider. While the programme itself does lead to large and sustained impacts on household asset holdings, earnings, land holdings, and hours worked in high-productivity activities, variation in coaching plays no discernible role in explaining the variation we observe in these outcomes.

However, if we believe the growing body of field experimental research on the role of coaching within ultra-poor graduation-style programmes, it appears that including some regular coaching within the intervention bundle

does seem to matter: When coaching has been removed entirely, treatment effects of these programmes often disappear (Blattman et al., 2016; Burchi and Strupat, 2018; Sedlmayr et al., 2020; Banerjee, Duflo, and Sharma, 2021).

If the treatment effects of graduation-style programmes are sensitive to receiving some vs *no* coaching, but are not sensitive to receiving some vs *more* coaching, this suggests that the effects of coaching on programme outcomes is strongly non-linear.

We expect coaching to have an impact on our key outcome variable via an intermediate effect on some other outcome. For instance, we might imagine coaching to increase a recipient's business skills, which allows her to make productive use of the asset which is transferred to her through the programme. In an effort to tease out some of the possible mechanisms through which coaching may be operating, we explore whether we can detect effects of the intervention and of the coaching variations on working through skills transfers or by encouraging programme compliance.

We find no evidence that either coaching variations or the programme as a whole has any effects on measures of business or life skills. That is, when comparing treatment to control, treatment women (who received a bundle of inputs including coaching) are statistically indistinguishable from control women (who received nothing) in terms of their confidence in their ability to perform a range of basic tasks. We interpret this as evidence against a 'human capital transfer' hypothesis of coaching effectiveness. We also find no clear evidence that coaching increases social inclusion within treatment villages, or that it has positive effects on a range of psycho-social characteristics.

Given the lack of evidence that coaching has any effect on these inter-

mediate candidates, we propose that, if coaching does play an important role within graduation interventions, this is most likely to be by encouraging programme compliance. That is, by expecting regular visits, programme recipients comply with the expectations of the programme – for instance, by following best practices in caring for their asset and by resisting the temptation to liquidate the asset prematurely. This explanation remains highly speculative, and may be an avenue for further research.

Ultimately, this paper offers a piece of good news that has direct implications for the scaling of graduation programmes. Coaching has been identified as a major bottleneck to bringing graduation-style progammes to scale – both because offering regular one-on-one coaching is expensive, and because scarce human resources are more difficult to scale than other inputs. While providing some form of regular coaching is likely to be an important element of the graduation model, we show that benefits can be sustained with significantly lower coaching inputs. A successful big push intervention need not be quite so big to stay just as good.

Appendix

3.A Details on the asset package

Table 3.A.1: Details on the asset transfer package

Asset package	% of recipients receiving package	Value in 2016 $PPP
Cow(s) and goat(s)	28	534
Cow(s) and poultry	48	540
Goats and poultry	21	379
Other	3	325
Weighted average		494

3.B Robustness checks

3.B.1 Double-post selection LASSO controls

Table 3.B.1: Overall treatment effects of the BRAC UPGP with weekly (T1), fortnightly (T2), or fortnightly + group (T3) coaching

	(1) Assets	(2) Consumption	(3) Earnings	(4) Value of land	(5) Livestock hrs	(6) Casual hrs	(7) All hrs
Weekly coaching (T1)	280.944***	8.572***	32.372***	454.994***	152.768***	3.291	224.117***
	(41.145)	(1.964)	(10.118)	(154.040)	(26.771)	(2.352)	(60.471)
Fortnightly coaching (T2)	288.078***	4.895***	37.437***	284.509**	178.938***	-2.876	163.350***
	(41.299)	(1.798)	(10.905)	(113.041)	(32.505)	(2.372)	(62.160)
Fortnightly + group coaching (T3)	295.256***	7.190***	29.452***	364.284***	194.952***	-0.231	208.956***
	(43.119)	(1.731)	(8.807)	(122.957)	(36.895)	(1.781)	(66.556)
Control mean at endline	314.23	44.35	152.25	457.00	276.20	546.85	769.82
% change, T1 v C	89	19	21	100	55	1	29
% change, T2 v C	92	11	25	62	65	-1	21
% change, T3 v C	94	16	19	80	71	-0	27
T1-T2 p-value	0.87	0.05	0.64	0.21	0.39	0.01	0.28
T1-T3 p-value	0.75	0.46	0.79	0.53	0.24	0.07	0.79
T2-T3 p-value	0.85	0.20	0.52	0.47	0.65	0.18	0.38
Number of observations	7356	7351	7380	7356	7380	7356	7380
Number of clusters	88	88	88	88	88	88	88

Notes: *p or q < .1; ** p or q < .05; *** p or q < .01. Total asset holdings is defined as the total value of all durable *household* assets. Consumption is defined as total monthly *per capita* food and non-food consumption. Earnings is defined as total monthly *household* earnings from paid economic activity. Value of land is defined as the total value of land owned by the *household*, excluding rented land and the value of the land on which the household's homestead is built. Livestock hours is defined as the total yearly hours dedicated by the respondent to livestock rearing and to other forms of agricultural self-employment. Casual hours is defined as the total yearly hours dedicated by the respondent to casual agricultural day labour and to casual maid labour. Total hours is defined as the total yearly hours dedicated to all remunerative work, including 'casual hours', 'livestock hours', and 'other hours', meaning it is not the simple sum of casual and livestock hours. All outcome variables are Winsorized at the 99% level. Standard errors, in parentheses, are clustered at the branch level. All regressions control for date of interview and include district fixed effects. Following Benjamini et al. (2006), we use false discovery rate (FDR) corrections and report sharpened q-values to account for multiple hypothesis testing across our primary outcome variables. While asset holdings and consumption were pre-specified as the two primary outcome variables (see Appendix 3.G), we conservatively apply FDR corrections to all seven outcome variables. Sample sizes in the four treatment and control groups are as follows: C = 1723 ; T1 = 2115 ; T2 = 2389 ; T3 = 2328

3.B.2 Accounting for differences in Village Action Committee formation within Treatment Arm 3

Table 3.B.2: Overall treatment effects of the BRAC UPGP with weekly (T1), fortnightly (T2), fortnightly + group coaching with VAC (T3a) vs without VAC (T3b)

	(1) Assets	(2) Consumption	(3) Earnings	(4) Value of land	(5) Livestock hrs	(6) Casual hrs	(7) All hrs
Weekly coaching (with VAC) (T1)	298.903***	8.364***	39.061***	502.092***	158.449***	-4.683	227.952***
	(44.499)	(1.970)	(9.619)	(156.064)	(28.595)	(52.698)	(60.445)
Fortnightly coaching (with VAC) (T2)	308.109***	4.141**	47.381***	287.119**	191.725***	72.163	193.405***
	(42.758)	(1.895)	(10.952)	(121.577)	(35.057)	(49.206)	(59.757)
Fortnightly + group (with VAC) (T3a)	253.055***	5.023**	36.698***	220.734	206.799***	45.182	240.059***
	(55.269)	(1.983)	(12.359)	(166.594)	(42.718)	(56.206)	(63.064)
Fortnightly + group (without VAC) (T3b)	394.701***	6.833***	51.727***	494.591***	223.474***	-39.065	236.205***
	(44.752)	(2.137)	(11.482)	(138.320)	(47.121)	(57.723)	(76.130)
Control mean at endline	314.23	44.35	152.25	457.00	276.20	546.85	769.82
% change, T1 v C	95	19	26	110	57	-1	30
% change, T2 v C	98	9	31	63	69	13	25
% change, T3a v C	81	11	24	48	75	8	31
% change, T3b v C	126	15	34	108	81	-7	31
T1-T2 p-value	0.82	0.03	0.43	0.13	0.29	0.08	0.54
T1-T3a p-value	0.43	0.12	0.84	0.13	0.24	0.37	0.83
T1-T3b p-value	0.02	0.50	0.35	0.96	0.15	0.47	0.91
T2-T3a p-value	0.28	0.66	0.46	0.65	0.70	0.62	0.41
T2-T3b p-value	0.03	0.23	0.76	0.12	0.50	0.02	0.47
T3a-T3b p-value	0.00	0.42	0.29	0.08	0.73	0.13	0.95
Adjusted R-squared	0.04	0.08	0.04	0.03	0.05	0.16	0.08
Number of observations	7421	7416	7445	7421	7445	7421	7445
Number of clusters	88	88	88	88	88	88	88

Notes: *p < .1; ** p < .05; *** p < .01. Total asset holdings is defined as the total value of all durable *household* assets. Consumption is defined as total monthly *per capita* food and non-food consumption. Earnings is defined as total monthly *household* earnings from paid economic activity. Value of land is defined as the total value of land owned by the *household*, excluding rented land and the value of the land on which the household's homestead is built. Livestock hours is defined as the total yearly hours dedicated by the respondent to livestock rearing and to other forms of agricultural self-employment. Casual hours is defined as the total yearly hours dedicated by the respondent to casual agricultural day labour and to casual maid labour. Total hours is defined as the total yearly hours dedicated to all remunerative work, including 'casual hours', 'livestock hours', and 'other hours', meaning it is not the simple sum of casual and livestock hours. All outcome variables are Winsorized at the 99% level. Standard errors, in parentheses, are clustered at the branch level. All regressions control for date of interview and include district fixed effects. C = 1723 ; T1 = 2115 ; T2 = 2389 ; T3a = 1029 ; T3b = 1299

3.B.3 Log transformation of outcome variables

Table 3.B.3: Overall treatment effects on log-transformed outcome variables

	(1) Assets	(2) Consumption	(3) Earnings	(4) Value of land	(5) Livestock hrs	(6) Casual hrs	(7) All hrs
Weekly coaching (T1)	0.792***	0.165***	0.447***	0.756***	158.281***	-4.015	227.991***
	(0.107)	(0.039)	(0.124)	(0.197)	(28.551)	(53.292)	(60.473)
Fortnightly coaching (T2)	0.878***	0.079**	0.485***	0.560***	191.611***	72.561	193.432***
	(0.114)	(0.037)	(0.133)	(0.191)	(34.964)	(49.486)	(59.691)
Fortnightly + group coaching (T3)	0.945***	0.117***	0.451***	0.736***	216.538***	-4.278	237.808***
	(0.114)	(0.035)	(0.136)	(0.199)	(38.357)	(51.391)	(64.560)
Control mean at endline	4.24	3.68	3.82	0.74	276.20	546.85	769.82
% change, T1 v C	19	4	12	102	57	-1	30
% change, T2 v C	21	2	13	75	69	13	25
% change, T3 v C	22	3	12	99	78	-1	31
T1-T2 p-value	0.36	0.02	0.74	0.28	0.29	0.08	0.54
T1-T3 p-value	0.11	0.17	0.97	0.91	0.11	1.00	0.87
T2-T3 p-value	0.49	0.28	0.79	0.26	0.50	0.07	0.38
Adjusted R-squared	0.06	0.09	0.06	0.05	0.05	0.16	0.08
Number of observations	7421	7416	7445	7421	7445	7421	7445
Number of clusters	88	88	88	88	88	88	88

Notes: *p < .1; ** p < .05; *** p < .01. Total asset holdings is defined as the total value of all durable *household* assets. Consumption is defined as total monthly *per capita* food and non-food consumption. Earnings is defined as total monthly *household* earnings from paid economic activity. Value of land is defined as the total value of land owned by the *household*, excluding rented land and the value of the land on which the household's homestead is built. Livestock hours is defined as the total yearly hours dedicated by the respondent to livestock rearing and to other forms of agricultural self-employment. Casual hours is defined as the total yearly hours dedicated by the respondent to casual agricultural day labour and to casual maid labour. Total hours is defined as the total yearly hours dedicated to all remunerative work, including 'casual hours', 'livestock hours', and 'other hours', meaning it is not the simple sum of casual and livestock hours. All outcome variables are Winsorized at the 99% level. Standard errors, in parentheses, are clustered at the branch level. All regressions control for date of interview and include district fixed effects. The outcome variables are transformed using the formula $ln(X + 1)$. C = 1723 ; T1 = 2115 ; T2 = 2389 ; T3 = 2328

3.B.4 Winsorization at 95%

Table 3.B.4: Overall treatment effects with Winsorization at the 95% level

	(1) Assets	(2) Consumption	(3) Earnings	(4) Value of land	(5) Livestock hrs	(6) Casual hrs	(7) All hrs
Weekly coaching (T1)	265.180***	7.385***	33.351***	231.312***	147.785***	-5.216	217.530***
	(36.889)	(1.715)	(8.234)	(62.591)	(27.319)	(50.090)	(56.990)
Fortnightly coaching (T2)	282.535***	3.340**	40.258***	165.478***	176.369***	67.844	179.794***
	(36.529)	(1.631)	(9.348)	(58.201)	(33.661)	(46.163)	(56.998)
Fortnightly + group coaching (T3)	298.450***	5.307***	39.456***	196.227***	202.797***	0.012	224.741***
	(37.805)	(1.484)	(8.954)	(59.534)	(36.261)	(48.366)	(61.507)
Control mean at endline	298.02	43.38	146.40	242.96	267.95	529.78	751.95
% change, T1 v C	89	17	23	95	55	-1	29
% change, T2 v C	95	8	27	68	66	13	24
% change, T3 v C	100	12	27	81	76	0	30
T1-T2 p-value	0.63	0.02	0.43	0.26	0.33	0.08	0.47
T1-T3 p-value	0.36	0.19	0.52	0.55	0.10	0.90	0.90
T2-T3 p-value	0.61	0.21	0.94	0.53	0.45	0.09	0.34
Adjusted R-squared	0.05	0.08	0.05	0.04	0.05	0.17	0.07
Number of observations	7421	7534	7445	7421	7445	7421	7445
Number of clusters	88	88	88	88	88	88	88

Notes: *p < .1; ** p < .05; *** p < .01. Total asset holdings is defined as the total value of all durable *household* assets. Consumption is defined as total monthly *per capita* food and non-food consumption. Earnings is defined as total monthly *household* earnings from paid economic activity. Value of land is defined as the total value of land owned by the *household*, excluding rented land and the value of the land on which the household's homestead is built. Livestock hours is defined as the total yearly hours dedicated by the respondent to livestock rearing and to other forms of agricultural self-employment. Casual hours is defined as the total yearly hours dedicated by the respondent to casual agricultural day labour and to casual maid labour. Total hours is defined as the total yearly hours dedicated to all remunerative work, including 'casual hours', 'livestock hours', and 'other hours', meaning it is not the simple sum of casual and livestock hours. All outcome variables are Winsorized at the 95% level. Standard errors, in parentheses, are clustered at the branch level. All regressions control for date of interview and include district fixed effects. C = 1723 ; T1 = 2115 ; T2 = 2389 ; T3 = 2328

3.C Cost benefit analysis

In Table 3.C.1 below, we provide some rough estimates of plausible cost-benefit ratios across the three treatment arms. A few caveats are important to highlight before discussing the estimates: BRAC either did not record or has been unable to provide detailed cost data disaggregated by treatment arm. They have only been able to supply total cost data for the entire 2016 cohort of 80,000 recipients – only approximately 6,000 of which were included in the experiment. We have therefore had to make some very rough assumptions about the per-recipient cost, and about the cost differential across treatment arms. While we describe these decisions in detail, the exercise remains highly speculative and these estimates should be interpreted with caution.

In addition, estimates of the benefits of the programme are primarily driven by the estimated consumption effects at six years. As is common prac-

tice (see, for instance, Bandiera et al. (2017)), we consider the consumption benefits of the programme to be the *cumulative, household level* consumption effects over a given time horizon. Because we only observe consumption at one moment in time, this requires making strong assumptions about the consumption benefits during both past and future periods which we do not observe directly. For transparency and following common practice (again, see Bandiera et al. (2017)) we use simple linear interpolation to estimate past and future consumption gains. This has the effect of potentially amplifying what may be imprecisely estimated differences that we observe between treatment groups at one moment in time.

The first row of Table 3.C.1 reports the estimated programme costs across the three treatment arms in 2016 $PPP. Because the 'standard' 2016 UPGP offering was identical to our third treatment arm (fortnightly and group-based coaching), we use BRAC's admin data to estimate what we expect is a fairly accurate per-recipient cost for this treatment arm. We are able to break down this cost between costs which are likely to remain fixed across treatment groups (i.e. the price of the asset, etc) and coaching-related costs which are likely to vary (coaching staff salaries, staff transport, etc). We estimate that 40% ($477) of per-participant programme costs would be sensitive to the amount of coaching delivered while the remainder would be fixed across participants ($716).

We estimate that, compared to the fornightly-plus-group coaching arm, BRAC saves 15% of the coaching costs in the fortnightly-only arm, bringing coaching-related costs down to $405 for that arm. We then assume that doubling the amount of coaching delivered in the weekly treatment arm relative to the fortnightly arm increases coaching costs to BRAC by a factor of 1.7, to $690. This yields the cost figures in row 1.

Row 2 adjusts these costs by the social discount rate, set at 5% , , following World Bank guidelines and Bandiera et al. (2017).

The second panel reports estimated programme benefits, based on the treatment effect estimates reported in Table 3.5.1. Consumption benefits are reported as the cumulative gains in household consumption attributable to the programme upo to and including year six after the programme began. Estimates for past years consumption gains are unobserved, and so are imputed using linear interpolation, assuming that consumption gains (beyond those attributable to the direct consumption support) were zero in the first year.

Bandiera et al. (2017) report cost-benefit estimates based on extending effects observed over the four-year horizon to 20 years, assuming a linear increase in consumption gains. We do the same for 12 years after the programme started and report these in the second row in Panel B. These projections are deflated in future years using the 5% social discount rate.

The third row in Panel B reports the combined treatment effects on asset and land holdings.

The final two panels report the combined benefits (assets and consumption) over the six and twelve years. We conservatively assume in the 12 year estimates that there is no further asset growth, but do assume further (discounted) linear consumption benefits.

Panel C reports cost-benefit ratios for the three treatment arms over the six- and twelve-year horizons. We observe that T2 comes off least favorably: Returns to receiving the UPGP with only fortnightly coaching are not positive over the six-year horizon, though they do become strongly positive if we believe that consumption benefits will be sustained over twelve years.

Cost benefit ratios for T1 and T3 are much more similar – already positive over the six-year horizon. Estimates do suggest that the weekly coaching treatment is marginally more cost effective than the fortnightly-plus-group coaching treatment despite higher costs. These results suggest that rather than simply halving the amount of one-on-one coaching provided, it may be preferable to replace these with more cost-effective group-based coaching sessions.

However, we urge caution in placing too much weight on this conclusion: First of all, as described above, the cost data is very imprecisely estimated and rests on highly speculative assumptions. Second, the cost-benefit ratio is being driven by estimated consumption benefits at one point in time. While consumption benefits of the programme do appear to be significantly different when comparing T1 and T2, differences between T1 and T3 are not significant (see Table 3.5.1). In addition, consumption results are unique among the outcome variables we consider in the main analysis: only for consumption are outcomes for T1 significantly more positive than for T2 or T3.

Table 3.C.1: Roughly estimated cost-benefit analysis across three treatment arms

	T1: Weekly Coaching	T2: Fortnightly Coaching	T3: Fortnightly + Group
Panel A: Estimated Programme Costs			
Cost per household at year 0	\$1,406	\$1,121	\$1,193
Cost per household discounted (5% per annum) at year 6	\$1,794	\$1,502	\$1,599
Panel B: Estimated Programme Benefits			
Cumulative increase in household consumption expenditure at year 6	\$1,361	\$724	\$1,056
NPV Change in household consumption expenditure at year 12	\$4,865	\$2,589	\$3,778
Effect on household asset holdings at year 6	\$798	\$593	\$718
6 year total benefits (assets and consumption)	\$2,159	\$1,317	\$1,774
12 year total benefits (assets and consumption)	\$5,663	\$3,183	\$4,496
Panel C: Cost Benefit Ratios			
Cost/benefit ratios at year 6	1.20	0.88	1.11
Change in financial assets year 12	3.16	2.00	2.81

Notes: Household consumption is defined as in Table 3.5.1, but is reported at the household rather than at the per capita level. Assets include both productive assets and non-homestead land holdings, both also defined as in Table 3.5.1. All monetary amounts are PPP-adjusted US$ terms, set at 2016 prices and deflated using the rural CPI published by the Bangladesh Bank.

3.D Evidence of treatment effects on other outcomes

3.D.1 Does the UPGP crowd in or crowd out additional governmental/non-governmental support?

In Table 3.D.1 we examine whether the BRAC UPGP crowds in or crowds out other sources of social assistance. Specifically, we estimate the the treatment effect of the intervention on the number of instances in which a household accessed help. We have hypothesised that one of the roles of coaches may be as intermediaries between individuals and other sources of economic, health and social support. In the endline survey, we asked recipients to list all instances in which household members received support from governmental or non-governmental agencies in the last 12 months. Note that there is therefore a lag of at least three years between when coaching was last delivered through the programme and the first instance of support which respondents were asked to report. It is therefore highly unlikely that we would pick up on the direct effects of a coach acting as an intermediary, but rather woudl measure the lasting effects of an increased ability to access help.

While point estimates are positive, overall they are not significant (apart from a marginally significant effect for the weekly coaching group). Differences across treatment groups are also insignificant. While we cannot measure lasting effects of treatment on the likelihood of receiving additional support, it does not appear that the BRAC UPGP crowds out additional support, despite recipients being materially better off than the control group.

Table 3.D.1: Number of instances household received help from government / NGO in last year

	(1) Instances of govt/non-govt assistance
All treatment groups	0.104
	(0.065)
Weekly coaching (T1)	0.137*
	(0.081)
Fortnightly coaching (T2)	0.100
	(0.076)
Fortnightly + group coaching (T3)	0.075
	(0.087)
Control mean at endline	1.64
% change, all T v C	6
% change, T1 v C	8
% change, T2 v C	6
% change, T3 v C	5
T1-T2 p-value	0.66
T1-T3 p-value	0.50
T2-T3 p-value	0.77
Adjusted R-squared	0.05
Number of observations	8529
Number of clusters	88

Notes: *p or q < .1; ** p or q < .05; *** p or q < .01. The outcome variable is derived from responses to the question: "In the last year, did any member of your HH receive any governmental/non-governmental help?"? Assistance was listed for each household number. The outcome variable is defined as the total household-level instances of support from either governmental or non-governmental sources. (In 90 percent of cases support was listed as being from the government). Instances can be one-off, or regular. Controls are included for date of interview as well as district fixed effects. Sample sizes in the four treatment and control groups are as follows: C = 1723 ; T1 = 2115 ; T2 = 2389 ; T3 = 2328

3.D.2 Social inclusion

Here we test for effects of the three coaching variants of the UPGP on a range of variables related to social inclusion. To measure social inclusion we construct indices from a number of survey questions included in the endline for this purpose.

Table 3.D.2 displays the treatment effect estimates on this range of social

inclusion related outcomes. We detect no treatment effects on the social inclusion indices in the first three columns, or on the 'willingness of other villagers to share land' variable in column 4.

However, there does appear to be evidence in columns 5 that those receiving the intervention experience a negative change in one indicator of social inclusion relative to the control group. Treatment group women are 20 percent less likely than control group woman to perceive themselves as being treated equally by richer villagers. In addition, there is evidence that recipients in the weekly coaching treatment arm (T1) experience a negative change in social inclusion relative to the control group on two other indicators: Women in the weekly coaching treatment arm are also 33 percent less likely to have attended a village meeting in the last year, and 11 percent less able to participate in *shalish*.[21] However, differences between treatment groups for these effects are mostly insignificant.

On the other hand, those receiving the intervention (specifically those women in the fortnightly coaching arm (T2) and fortnightly and group coaching arm (T3)) appear significantly less likely to consider themselves among the poorest members of a village – perhaps reflecting the economic rather than the social effects of the UPGP.

Since these findings are not strongly significant or consistent across indicators of social inclusion, we want to be cautious in drawing inferences about the effect of the UPGP on social inclusion outcomes. However, it does seem that the UPGP intervention in general (rather than coaching variants in particular) may crowd out some forms of social connection within villages. Perhaps, for example, access to forms of BRAC support through the pro-

[21] *Shalish* is an informal conflict dispute resolution institution.

gramme crowd out support sought from other informal village-based institutions. To the extent that engagement in these institutions sustain cross-class social connection, this may also be compromised – indicated by the negative effects on 'equal treatment by richer villages' in column 5. However, given the questionable strength of the evidence, we advance these hypotheses with caution.

Table 3.D.2: Treatment effects on social inclusion

	(1) Social inclusion: Other villagers	(2) Social inclusion: Village leaders	(3) Social inclusion: Villager committee	(4) Villagers will share land	(5) Treated as equal by richer villagers	(6) Consider self among poorest	(7) Attended village meeting	(8) Can participate in Shalish
All treatment groups	-0.038	0.015	0.028	0.018	-0.050**	-0.084***	-0.034	-0.049
	(0.037)	(0.037)	(0.051)	(0.040)	(0.023)	(0.024)	(0.032)	(0.031)
Weekly coaching (T1)	-0.032	0.007	0.019	-0.005	-0.053**	-0.038	-0.061*	-0.067**
	(0.039)	(0.039)	(0.055)	(0.050)	(0.026)	(0.027)	(0.036)	(0.033)
Fortnightly coaching (T2)	-0.060	0.017	0.020	0.000	-0.043	-0.105***	-0.045	-0.065
	(0.051)	(0.050)	(0.064)	(0.057)	(0.030)	(0.032)	(0.042)	(0.041)
Fortnightly + group coaching (T3)	-0.028	0.023	0.043	0.054	-0.051*	-0.116***	-0.000	-0.019
	(0.041)	(0.046)	(0.058)	(0.048)	(0.026)	(0.028)	(0.040)	(0.035)
Control mean at endline	0.79	0.69	0.67	0.54	0.25	0.72	0.18	0.59
% change, all T v C	-5	2	4	3	-20	-12	-19	-8
% change, T1 v C	-4	1	3	-1	-21	-5	-33	-11
% change, T2 v C	-8	2	3	0	-18	-15	-24	-11
% change, T3 v C	-4	3	6	10	-21	-16	-0	-3
T1-T2 p-value	0.50	0.83	0.99	0.93	0.72	0.03	0.63	0.96
T1-T3 p-value	0.91	0.67	0.56	0.25	0.95	0.00	0.12	0.09
T2-T3 p-value	0.43	0.90	0.67	0.32	0.78	0.72	0.29	0.20
Adjusted R-squared	0.04	0.07	0.05	0.04	0.03	0.06	0.03	0.04
Number of observations	5172	5172	5172	7534	7534	7534	7534	7534
Number of clusters	88	88	88	88	88	88	88	88

Notes: Columns 1-3 present treatment effects on indices of social inclusion. The outcome in column 1 is an index measuring feelings of social inclusion specific to 'other villagers', column 2 is specific to 'village leaders', and column 2 is specific to 'village committee members'. Indices are constructed from answers to whether the respondent could approach other villagers/village leaders/village committee members about: managing own business, finding new business opportunities, how to resolve a dispute, accessing medical care, and furthering the education of children in one's household. Columns 4-7 are dummy variables, constructed from the following questions: 'Do you think other villagers will share/lease/mortgage land to you?' (4), 'Are you treated as an equal by other members of the village who are richer than you?' (5), 'In the last year, have you attended a village meeting?' (7). Column 6 is constructed as a dummy variable equal to 1 if a respondent considered themselves at the bottom of a ten step ladder 'where the poorest people in your village stand on the bottom (the first step) and the richest people in your village stand on the highest step (the tenth step).' Column 8 is a simple index constructed from response to the following questions: 'Do you think you will get help from shalish regarding any problem you may face?' and 'Can you participate in shalish willingly?'. Standard errors, in parentheses, are clustered at the branch level. All regressions control for date of interview and include district fixed effects.

3.D.3 Psycho-social outcomes

Coaching may serve as a means of increasing the psychological wellbeing, aspirations and self-efficacy or locus of control of recipients, by exposing them to role models in the community (the coaches) as well as through life-skills which they may acquire through coaching. Here we consider a range of variables which were constructed to detect whether there were any changes in these psycho-social characteristics.

We consider four outcome variables: locus of control, depression, self-reported happiness, and a measure of hope or aspiration. Locus of control refers to an individual's belief about the extent to which they have control over the events that affect their lives, where people with an 'internal' locus of control believe that their actions and decisions have a direct impact on the outcomes they experience, while those with an 'external' locus of control believe that their experiences are largely determined by factors outside of their control, such as luck, fate, or other people's actions. We measure locus of control to test the hypothesis that coaching changes beneficiaries' beliefs about the extent to which they are empowered to make changes in their lives and control the outcomes that they care about.[22] We measure depression using the CESD-10 item depression scale (Radloff, 1977).[23] The forth col-

[22]We measure locus of control using Lumpkin's brief locus of control scale (Lumpkin, 1985). Lumpkin's brief locus of control scale is constructed as the average of six Likert-scale questions. The first three questions focus on internal locus of control: *i)* 'When I make plans I am almost certain to make them work.' *ii)* 'Getting people to do the right thing depends upon ability. Luck has little or nothing to do with it.' *iii)* 'What happens to me is my own doing.' A second set of questions focuses on external locus of control: *iv)* 'Many of the unhappy things in people's lives are partly due to bad luck.' *v)* 'Getting a good job depends mainly on being in the right place at the right time.' *vi)* 'Many times I feel that I have little influence over the things that happen to me.'

[23]We selected the CESD-10 scale over other scales (such as the PHQ-9 scale) because of its simplicity and track-record in developing country contexts, including in Bangladesh (Folb et al., 2015; Khan et al., 2020). Each item on the CESD-10 is scored on a four-point Likert scale, with response options ranging from 0 (rarely or none of the time) to 3 (most

umn report effects on a dummy variable equal to 1 if recipients considered themselves to be happy with their lives overall, and zero otherwise. The fifth column reports the difference between the step on Cantril's ladder[24] on which respondents expect to be in five years time, and the step of the ladder that perceive themselves to be on now. This variable therefore represents the expectations an individual has for an improvement in the state of their lives.

or all of the time). The total score on the scale ranges from 0 to 30, with higher scores indicating greater severity of depressive symptoms. The items are: 'I felt depressed', 'I felt that everything I did was an effort', 'My sleep was restless', 'I was happy', 'I felt lonely', 'People were unfriendly', 'I enjoyed life', 'I felt sad', 'I felt that people dislike me', 'I could not 'get going'.'The CESD-10 can be displayed as a score variable, or as a dummy where if an individual's score is above a certain threshold (in this case 10), they are considered to display depressive symptoms.

[24]Respondents were asked to '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' and were asked 'Which step of the ladder best represents the way you personally feel you stand these days?' and 'On which step of the ladder do you expect to be 5 years from now?'.

Table 3.D.3: Treatment effects on psycho-social outcomes

	(1) Depression score	(2) Presence of depressive symptoms	(3) External locus of control	(4) Overall happy	(5) Expected 5yr change on Cantril ladder
All treatment groups	0.039	-0.029	-0.048	0.063*	-0.152**
	(0.329)	(0.026)	(0.033)	(0.037)	(0.076)
Weekly coaching (T1)	-0.065	-0.032	-0.016	0.052	-0.023
	(0.362)	(0.030)	(0.047)	(0.037)	(0.098)
Fortnightly coaching (T2)	0.197	-0.015	-0.093**	0.072	-0.188**
	(0.450)	(0.032)	(0.037)	(0.043)	(0.085)
Fortnightly + group coaching (T3)	0.033	-0.037	-0.047	0.068*	-0.259***
	(0.381)	(0.029)	(0.036)	(0.040)	(0.083)
Control mean at endline	13.01	0.73	3.20	0.67	2.02
% change, all T v C	0	-4	-1	9	-8
% change, T1 v C	-1	-4	-1	8	-1
% change, T2 v C	2	-2	-3	11	-9
% change, T3 v C	0	-5	-1	10	-13
T1-T2 p-value	0.50	0.51	0.06	0.48	0.06
T1-T3 p-value	0.79	0.84	0.49	0.52	0.01
T2-T3 p-value	0.67	0.38	0.16	0.91	0.26
Adjusted R-squared	0.07	0.04	0.05	0.02	0.04
Number of observations	7421	7421	7534	7511	7235
Number of clusters	88	88	88	88	88

Notes: Column 1 presents the raw score, out of 30, of the Center for Epidemiologic Studies Depression Scale (CESD-10). Higher scores indicate more severe depressive symptoms. Column 2 presents the CESD-10 as a dummy variable, equal to 1 if an individual's score is above a threshold of 10, and 0 otherwise. Column 3 presents results on the Lumpkin (1985) locus of control scale, where higher values indicate a more external locus of control – i.e. a 'worse' outcome. Column 4 reports effects on a dummy variable in which respondents were asked whether, overall, they were happy or unhappy with their lives. This variable is also coded such that it is equal to 1 if a respondent is 'unhappy', and 0 otherwise. Column 5 reports the difference between the step on Cantril's ladder on which respondents expect to be in five years time, and the step of the ladder that perceive themselves to be on now. Specifically, respondents were asked to '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' and were asked 'Which step of the ladder best represents the way you personally feel you stand these days?' and 'On which step of the ladder do you expect to be 5 years from now?'. Standard errors, in parentheses, are clustered at the branch level. In the first three columns, higher values indicate more negative outcomes while in the last two columns higher values indicate positive outcomes. All regressions control for date of interview and include district fixed effects.

We observe few significant treatment effects. For depression and locus of control, overall effects are fairly tightly centered around zero. For locus of control, those in the fortnightly coaching treatment arm appear to have a somewhat more 'internal' locus of control. Estimates on the intervention's effects on overall happiness suggest that treatment groups are slightly more likely to report being happy than the control group, though effects are imprecisely estimated and appear to be driven by increases in the fortnightly coaching and group treatment group (T3). Surprisingly, we find that those receiving the intervention have significantly lower expectations about future improvements in the state of their lives when compared to the control group, driven by effects on those in the fortnightly coaching (T2) and fortnightly and group coaching treatment group (T3). Differences between treatment groups are largely insignificant, a lone exception being that beneficiaries in the fortnightly coaching and group treatment group (T3) have substantially lower expectations for improvements in the state of their lives than those who receive weekly coaching (T1).

The most likely explanation for the counter-intuitive negative association of treatment status with expectations of improvements in quality of life is that treatment group women on average place themselves on higher rungs of the Cantril ladder in the present – on average on rung 1.4 of the 10-rung Cantril ladder, compared to rung 1.1 for control group women. Expectations for 5 years from the present are similar – both at 3.3. The negative effects in Table 3.D.3 therefore appear to be driven by the fact that the control group women perceive themselves as better off in the present. Nevertheless, the results do provide some evidence against the hypothesis that the GDBC treatment, and the coaching element within it, increase hope and/or aspirations.

3.D.4 Financial inclusion

Table 3.D.4 presents treatment effects on a number of variables relating to financial inclusion. We find that the intervention increases the amount and the incidence of savings relative to the control group, even six years after the programme began, though effects are not significant for the fortnightly coaching group and for the weekly coaching group are only significant for the value and not for the incidence of savings. For the fortnightly and group coaching arm (T3), the number of recipients who claim to have any household savings is approximately 40 percent higher than for the control group, and on average they save approximately 70 percent more on average. Effects are smaller and less significant for the non-group coaching arms (T1 and T2). Differences in effects between T2 and T3 are also significant for both outcomes, though not between T1 and T3. This finding is perhaps somewhat counter-intuitive: If group-based sessions are opportunities for recipients to set up informal risk sharing institutions, this should reduce the demand for savings, not increase it. The differential impact between T2 and T3 remains a puzzle, but could perhaps be interpreted as suggestive evidence *against* the notion that group-meetings double-up as informal risk sharing mechanisms. Overall, it remains notable that effects on savings for some treatment arms persist several years after the benefits of the programme have ceased.

Results for receiving loans or giving loans are more ambiguous. We observe significant overall effects on the likelihood of lending to others, but not on the likelihood of receiving loans on either the intensive or extensive margin – although point estimates are positive.

All three treatment groups appear less likely to have *given* remittances (though this effect is only weakly significant for the fortnightly coaching arm

(T2)), while all three treatment groups appear more likely to have *received* remittances – with significant differences for T1 and T3. Differences in giving or receiving remittances are not significant across treatment groups. Given that treatment groups are demonstrably material better-off than treatment groups, it may be surprising that remittance sending is lower and remittance receiving is higher. These effects may be explained by the establishment of the 'Village Poverty Reduction Committees' in treatment villages – a community mobilisation effort set up by BRAC explicitly to encourage better-off villagers to support the ultra-poor.

Overall, the evidence for the impact of treatment on financial inclusion is mixed. The UPGP intervention appears to increase savings, especially among the group coaching arm, has no discernible impact on borrowing, and in some treatment groups leads to a reduction in remittance sending, and in others an increase in remittance receiving. Results have not been corrected for multiple hypotheses – doing so may well reduce the significance of some of these somewhat paradoxical results.

Table 3.D.4: Treatment effects on financial inclusion

	(1) Has savings	(2) Amount saved	(3) Has borrowed money/goods	(4) Amount borrowed	(5) Has lent money/goods	(6) Gave money to others	(7) Received money from others
All treatment groups	0.106***	24.290**	0.036	63.055	0.031**	-0.085***	0.043**
	(0.026)	(11.658)	(0.022)	(39.134)	(0.015)	(0.031)	(0.017)
Weekly coaching (T1)	0.070**	23.881*	0.015	50.066	0.032**	-0.055	0.048**
	(0.033)	(14.363)	(0.025)	(43.938)	(0.016)	(0.037)	(0.020)
Fortnightly coaching (T2)	0.075**	2.939	0.040	63.543	0.015	-0.115***	0.026
	(0.032)	(14.184)	(0.026)	(45.732)	(0.019)	(0.037)	(0.021)
Fortnightly + group coaching (T3)	0.164***	39.790***	0.055**	75.926*	0.040**	-0.094**	0.051**
	(0.033)	(14.214)	(0.025)	(42.771)	(0.017)	(0.041)	(0.023)
Control mean at endline	0.36	76.45	0.57	330.59	0.06	0.65	0.11
% change, all T v C	30	32	6	19	52	-13	38
% change, T1 v C	20	31	3	15	54	-8	43
% change, T2 v C	21	4	7	19	26	-18	23
% change, T3 v C	46	52	10	23	67	-14	45
T1-T2 p-value	0.85	0.08	0.23	0.69	0.26	0.10	0.27
T1-T3 p-value	0.01	0.24	0.05	0.44	0.62	0.34	0.91
T2-T3 p-value	0.01	0.00	0.50	0.75	0.11	0.60	0.29
Adjusted R-squared	0.05	0.02	0.05	0.13	0.01	0.07	0.02
Number of observations	7421	7421	7534	7421	7534	7534	7534
Number of clusters	88	88	88	88	88	88	88

Notes: 'Has savings' is a dummy variable equal to 1 if a respondent reports having any household savings at endline, and 0 otherwise. 'Amount saved' is the total amount of the respondent's household savings, expressed in 2016 $PPP. 'Has borrowed money/goods' is a dummy variable equal to 1 if the respondent's household has an outstanding loan whether in cash or kind and whether to a formal or informal lender. 'Amount borrowed' is the total amount of the respondent's household debt, expressed in 2016 $PPP. 'Has lent money/goods' is a dummy variable equal to 1 if respondent's household has lent money or goods to another individual or household. 'Gave money to others' is a dummy variable equal to 1 if the respondent's household transferred money or goods to another household in the last 12 months, with no expectation of repayment. 'Received money from others' is a dummy variable equal to 1 if the respondent's household received money or goods to another household in the last 12 months, with no expectation of repayment. 'Amount saved' and 'Amount borrowed' are Winsorized at the 99 percent level. Standard errors, in parentheses, are clustered at the branch level. All regressions control for date of interview and include district fixed effects.

3.E Balance

Table 3.E.1: Baseline balance across all control and treatment groups

	(1) Control		(2) Weekly		(3) Fortnightly		(4) Fortnightly+group		T-test Difference						F-test for joint
Variable	N/[Clusters]	Mean/SE	N/[Clusters]	Mean/SE	N/[Clusters]	Mean/SE	N/[Clusters]	Mean/SE	(1)-(2)	(1)-(3)	(1)-(4)	(2)-(3)	(2)-(4)	(3)-(4)	orthogonality
Age	1723	44.233	2115	44.076	2389	42.896	2328	42.862	0.158	1.337	1.372	1.179	1.214	0.035	0.399
	[22]	(1.325)	[22]	(1.176)	[22]	(0.987)	[22]	(1.286)							
Household head is a woman	1723	0.548	2115	0.519	2389	0.529	2328	0.498	0.029	0.019	0.050	-0.010	0.021	0.031	0.398
	[22]	(0.030)	[22]	(0.037)	[22]	(0.037)	[22]	(0.037)							
Household size	1723	3.128	2115	3.102	2389	3.259	2328	3.297	0.027	-0.131	-0.169	-0.157	-0.196	-0.038	0.283
	[22]	(0.158)	[22]	(0.175)	[22]	(0.183)	[22]	(0.205)							
Respondent is married	1723	0.495	2115	0.528	2389	0.524	2328	0.570	-0.033	-0.029	-0.075*	0.004	-0.042	-0.046	0.962
	[22]	(0.029)	[22]	(0.032)	[22]	(0.032)	[22]	(0.034)							
Monthly earnings (2016$PPP)	1723	1393.534	2115	1389.971	2389	1425.213	2328	1440.225	3.563	-31.679	-46.691	-35.242	-50.254	-15.011	0.030
	[22]	(157.746)	[22]	(136.777)	[22]	(130.212)	[22]	(147.824)							
Monthly earnings from livestock (2016$PPP)	1723	30.533	2115	30.203	2389	30.828	2328	24.854	0.330	-0.296	5.679	-0.625	5.349	5.974	0.206
	[22]	(10.577)	[22]	(6.032)	[22]	(9.882)	[22]	(5.519)							
Monthly household expenditure (2016$PPP)	1697	6.193	2092	6.210	2366	6.251	2298	6.456	-0.017	-0.057	-0.262**	-0.041	-0.246**	-0.205*	2.404*
	[22]	(0.097)	[22]	(0.091)	[22]	(0.073)	[22]	(0.075)							
Attrited at EL (narrow definition)	1719	0.136	2106	0.119	2381	0.124	2323	0.142	0.016	0.012	-0.006	-0.005	-0.022	-0.018	0.977
	[22]	(0.015)	[22]	(0.007)	[22]	(0.007)	[22]	(0.013)							
Attrited at EL (broad definition)	1723	0.126	2115	0.110	2389	0.105	2328	0.124	0.016	0.020	0.002	0.004	-0.014	-0.019	1.008
	[22]	(0.015)	[22]	(0.007)	[22]	(0.007)	[22]	(0.012)							

Notes: The value displayed for t-tests are the differences in the means across the groups. Standard errors are clustered at variable branch. ***, **, and * indicate significance at the 1, 5, and 10 percent critical level.

3.F Attrition

Table 3.F.1: Differential attrition

	(1) Remain in sample		(2) Attritted		(2)-(1) Pairwise t-test	
Variable	N	Mean/(SE)	N	Mean/(SE)	N	Mean difference
Age	7421	43.865 (0.161)	1108	40.690 (0.445)	8529	-3.175***
Household head is a woman	7421	0.525 (0.006)	1108	0.500 (0.015)	8529	-0.025
Household size	7421	3.187 (0.020)	1108	3.310 (0.049)	8529	0.123**
Respondent is married	7421	0.530 (0.006)	1108	0.546 (0.015)	8529	0.016
Monthly earnings (2016$PPP)	7421	50.238 (0.610)	1108	42.542 (1.344)	8529	-7.696***
Monthly earnings from livestock (2016$PPP)	7421	1.094 (0.125)	1108	0.463 (0.096)	8529	-0.631*
Monthly household expenditure (2016$PPP)	7356	26.547 (0.426)	1072	26.447 (0.893)	8428	-0.100
Number of cows	7421	0.012 (0.002)	1108	0.016 (0.006)	8529	0.005
F-test of joint significance (F-stat)						8.551***
F-test, number of observations						8428

Notes: The value displayed for t-tests are the differences in the means across the groups. The value displayed for F-tests are the F-statistics. ***, **, and * indicate significance at the 1, 5, and 10 percent critical level.

3.G Pre-analysis plan

The experiment was pre-registered on the AEA RCT registry, and can be found here.

3.H Construction of variables for comparison with Bandiera et al. (2017) and Banerjee, Duflo, Goldberg, et al. (2015)

Table 3.5.2 presents a comparison of ITT estimates across a number of outcomes between the present 6-year study, and Bandiera et al. (2017)'s first estimates of the effects of the TUP programme in Bangladesh, and Banerjee, Duflo, Goldberg, et al. (2015)'s pooled evaluation of the effects of graduation programme pilots in six countries – Ethiopia, Ghana, Honduras, India, Pakistan and Peru. To ensure that comparisons are as close to being like-for-like as possible, we closely follow the variable and index definitions provided by Banerjee, Duflo, Goldberg, et al. (2015) and Bandiera et al. (2017), and note below where and why our index construction differs.

In Table 3.5.2 we present comparisons for nine out of the ten primary outcome measures studied in Banerjee, Duflo, Goldberg, et al. (2015). We omit outcomes for women's empowerment because our baseline and endline instruments did not collect data on the necessary variables. Each outcome is an index combining outcomes for individual/household i related to outcome k, denoted Y_i^k. Following Banerjee et al. (2015a), we construct each index k by first defining every outcome within the relevant group of outcomes such that higher values correspond to better outcomes. We then standardize each outcome into a z-score by subtracting the control group mean and dividing by the control group standard deviation (SD) for the corresponding survey round. We then average all the z-scores and again standardise to the control group within each round. We convert all monetary values to 2014 USD PPP terms.

Following Banerjee, Duflo, Goldberg, et al. (2015) and Bandiera et al. (2017), we estimate the following specification:

$$Y_i^k = \alpha + \beta_1 T_i + Z_i^k + \nu_d + \epsilon_i \tag{3.2}$$

where Y_i^k is the outcome k of interest for either household or adult i, T is an indicator for having been randomly selected into the programme, Z_i^k is the household or adult's baseline value of the outcome k, and ν_d denote district fixed effects. Standard errors are clustered at the branch level, which was the unit of randomisation.

Column (1), Total per capita consumption: The dependent variable is standardized total per capita consumption per month. Following Banerjee, Duflo, Goldberg, et al. (2015) and Bandiera et al. (2017), we exclude expenditure on income-generating activity, use a per capita measure (rather than expenditure per adult equivalent, as used elsewhere), and use 2014 PPP prices.

Column (2), Food security index: We use the same survey instruments as Bandiera et al. (2017) to construct this index. Banerjee, Duflo, Goldberg, et al. (2015)'s survey instruments differ slightly (see Bandiera et al. (2017) for a description). We use three indicators: (i) We define a variable equal to 1 if the respondent reported that her household's status in terms of food availability was "neither deficit nor surplus" or "food surplus" and 0 if she said it was "always deficit" or "deficit sometimes". (ii) We define a variable equal to 1 if the respondent reported that in the month preceding the survey, her household never had less than enough food to eat and 0 otherwise. (iii) We define a dummy variable equal to 1 if the respondent reported that her

household could afford to have two meals per day most of the time during the last year and 0 otherwise.

Column (3), Asset index: We follow the approach taken by Banerjee, Duflo, Goldberg, et al. (2015) in constructing this index of household and productive assets. As described by Bandiera et al. (2017), we 'calculate the median unit value for each type of asset; (ii) calculate the value of each asset in terms of goats (the numeraire asset) by dividing the unit value of each asset by the median unit value of goats; (iii) calculate total asset value by multiplying the unit value of each asset (expressed in terms of goats) by the number of each asset owned; (iv) standardize the total asset value.'

Column (4), Total time spent working, main woman: We construct a standardized measure of the total time spent by the main woman of the household in productive activities on a typical day during the past year. Banerjee, Duflo, Goldberg, et al. (2015) construct a variable representing the total minutes spent on all productive activities in the *day prior to survey day*. Liek Bandiera et al. (2017), we do not have this information, so instead convert information we collected on hours of labour supplied in the past 3 months to one year to a daily figure, multiplying this by 60 to get minutes per day. Banerjee, Duflo, Goldberg, et al. (2015) present results for all adults in the household, while Bandiera et al. (2017) present results for only the main female respondent. We follow Bandiera et al. (2017)'s convention, and so omit Banerjee, Duflo, Goldberg, et al. (2015)'s results.

Column (5), Income and revenue index: Like Bandiera et al. (2017), we include the following in our income and revenues index: (i) household livestock revenue per month; (ii) household agricultural income per month; (iii) household non-farm micro-enterprise income per month; (iv) household

income from paid labor per month. We only include the earnings of the female main respondent. Banerjee, Duflo, Goldberg, et al. (2015) also include a self-reported economic status indicator. Like Bandiera et al. (2017), we do not have the data necessary to include this.

Column (6), Physical health index: We construct a physical health index using a similar but shorter list of variables as Bandiera et al. (2017). We construct the index using two indicators: (i): We construct a composite index based on responses to the whether the respondent would be able to perform five physical activities, rated on a scale from 1 to 3 where 1=easily, 2=with trouble and 3=unable. Responses are re-coded so that higher values imply 'better' outcomes, averaging these rescaled responses to construct an index. The physical activities were: walking one mile at a normal speed, drawing a pail of water from a tube-well, and standing up from a sitting position on the floor without help. (ii) we use information on the respondent's self-perception of her current health on a 3-point scale (1=good, 2=average, 3=bad), re-coded so that higher values imply better health. Bandiera et al. (2017)'s index also included a variable on whether any household member had an illness in the 15 days before the survey, which we did not collect. They also included two additional activities in indicator (i): carrying a heavy load (e.g. 10 seer rice) for 20 yards, and using a ladder to climb to a storage place at least 5 feet high. Banerjee, Duflo, Goldberg, et al. (2015) use different, but closely related indicators, with differences detailed in Bandiera et al. (2017).

Column (7), Mental health index: We use a simple measure derived from a variable describing how the respondent considers her life in terms of happiness on a scale from 1 to 3 where 3=very happy, 2=happy and 1=unhappy. While we also collected a full CESD depression scale, we only have these variables at endline, so cannot control for baseline values. For this

reason we opt for a simple uni-dimensional outcome. In contrast, Banerjee, Duflo, Goldberg, et al. (2015) and Bandiera et al. (2017) construct an index using, in addition to the simple indicator we use, another one and two indicators respectively (see Bandiera et al. (2017) for details).

Column (8), Political awareness index: We use a simple measure derived from a variable reporting whether the respondent voted in the last election. While we have more information on political awareness and participation, these were not included in the baseline survey instrument, and so we have opted to use a simpler indicator in which we can control for baseline values. In contrast, Banerjee, Duflo, Goldberg, et al. (2015) and Bandiera et al. (2017) construct an index using four or five indicators respectively (see Bandiera et al. (2017) for details). Banerjee, Duflo, Goldberg, et al. (2015)'s index includes our voting variable, but Bandiera et al. (2017)'s does not.

References

[1] Lenore Sawyer Radloff. "The CES-D scale: A self-report depression scale for research in the general population". In: *Applied psychological measurement* 1.3 (1977), pp. 385–401.

[2] James R Lumpkin. "Validity of a brief locus of control scale for survey research". In: *Psychological reports* 57.2 (1985), pp. 655–659.

[3] Yoav Benjamini, Abba M Krieger, and Daniel Yekutieli. "Adaptive linear step-up procedures that control the false discovery rate". In: *Biometrika* 93.3 (2006), pp. 491–507.

[4] Naomi Ahmad, Faizuddin Ahmed, Dean Mitchell Jolliffe, Mehar Akhter Khan, Mehrin Ahmed Mahbub, lftath Sharif, Vinaya Swaroop, Nobuo Yoshida, Salman Zaidi, and Johannes. Zutt. "Poverty maps of Bangladesh". In: 90487 (2010). URL: http://documents.worldbank.org/curated/en/160611468014459434/Technical-report.

[5] Ugo Pica-Ciamarra, Derek Baker, Nancy Morgan, and Alberto Zezza. "Measuring the contribution of livestock to household livelihoods: a livestock module for multi-topic household surveys". In: (2011).

[6] Vivi Alatas, Abhijit Banerjee, Rema Hanna, Benjamin A. Olken, and Julia Tobias. "Targeting the Poor: Evidence from a Field Experiment in Indonesia". In: *American Economic Review* 102.4 (June 2012), pp. 1206–40. DOI: 10.1257/aer.102.4.1206.

[7] Alexandre Belloni, Victor Chernozhukov, and Christian Hansen. "Inference on treatment effects after selection among high-dimensional controls". In: *The Review of Economic Studies* 81.2 (2014), pp. 608–650.

[8] Aart Kraay and David McKenzie. “Do poverty traps exist? Assessing the evidence”. In: *Journal of Economic Perspectives* 28.3 (2014), pp. 127–148.

[9] Abhijit Banerjee, Esther Duflo, Nathanael Goldberg, Dean Karlan, Robert Osei, William Parienté, Jeremy Shapiro, Bram Thuysbaert, and Christopher Udry. “A multifaceted program causes lasting progress for the very poor: Evidence from six countries”. In: *Science* 348.6236 (2015), p. 1260799. DOI: 10.1126/science.1260799.

[10] Naomi Folb, Crick Lund, Lara R Fairall, Venessa Timmerman, Naomi S Levitt, Krisela Steyn, and Max O Bachmann. “Socioeconomic predictors and consequences of depression among primary care attenders with non-communicable diseases in the Western Cape, South Africa: cohort study within a randomised trial”. In: *BMC Public Health* 15.1 (2015), pp. 1–12.

[11] Christopher Blattman, Eric P. Green, Julian Jamison, M. Christian Lehmann, and Jeannie Annan. “The Returns to Microenterprise Support among the Ultrapoor: A Field Experiment in Postwar Uganda”. In: *American Economic Journal: Applied Economics* 8.2 (Apr. 2016), pp. 35–64. DOI: 10.1257/app.20150023. URL: https://www.aeaweb.org/articles?id=10.1257/app.20150023.

[12] Johannes Haushofer and Jeremy Shapiro. “The short-term impact of unconditional cash transfers to the poor: experimental evidence from Kenya”. In: *The Quarterly Journal of Economics* 131.4 (2016), pp. 1973–2042.

[13] Oriana Bandiera, Robin Burgess, Narayan Das, Selim Gulesci, Imran Rasul, and Munshi Sulaiman. “Labor markets and poverty in vil-

lage economies". In: *The Quarterly Journal of Economics* 132.2 (2017), pp. 811–870.

[14] Karthik Muralidharan and Paul Niehaus. "Experimentation at scale". In: *Journal of Economic Perspectives* 31.4 (2017), pp. 103–124.

[15] Omar Al-Ubaydli, John A List, and Dana L Suskind. "What can we learn from experiments? Understanding the threats to the scalability of experimental results". In: *American Economic Review* 107.5 (2017), pp. 282–286.

[16] Tessa Bold, Mwangi Kimenyi, Germano Mwabu, Justin Sandefur, et al. "Experimental evidence on scaling up education reforms in Kenya". In: *Journal of Public Economics* 168 (2018), pp. 1–20.

[17] Francesco Burchi and Christoph Strupat. "Unbundling the impacts of economic empowerment programmes: evidence from Malawi". In: (2018).

[18] Munshi Sulaiman. "Livelihood, Cash Transfer, and Graduation Approaches: How Do They Compare in Terms of Cost, Impact, and Targeting?" In: *Boosting Growth to End Hunger by 2025: The Role of Social Protection* (2018), pp. 102–20.

[19] Guadalupe Bedoya, Aidan Coville, Johannes Haushofer, Mohammad Isaqzadeh, and Jeremy P Shapiro. *No Household Left Behind: Afghanistan Targeting the Ultra Poor Impact Evaluation*. Working Paper 25981. National Bureau of Economic Research, June 2019. DOI: 10.3386/w25981.

[20] Omar Al-Ubaydli, John A List, and Dana Suskind. *The Science of Using Science: Towards an Understanding of the Threats to Scaling Experiments*. Working Paper 25848. National Bureau of Economic Research, May 2019. DOI: 10.3386/w25848. URL: http://www.nber.org/papers/w25848.

[21] Marc F Bellemare and Casey J Wichman. "Elasticities and the inverse hyperbolic sine transformation". In: *Oxford Bulletin of Economics and Statistics* 82.1 (2020), pp. 50–61.

[22] Asaduzzaman Khan, Rushdia Ahmed, and Nicola W Burton. "Prevalence and correlates of depressive symptoms in secondary school children in Dhaka city, Bangladesh". In: *Ethnicity & health* 25.1 (2020), pp. 34–46.

[23] Richard Sedlmayr, Anuj Shah, and Munshi Sulaiman. "Cash-plus: Poverty impacts of alternative transfer-based approaches". In: *Journal of Development Economics* 144 (2020), p. 102418.

[24] Colin Andrews, Aude de Montesquiou, Inés Arévalo Sánchez, Puja Vasudeva Dutta, Sadna Samaranayake, Janet Heisey, Timothy Clay, and Sarang Chaudhary. *The state of economic inclusion report 2021: The potential to scale*. World Bank Publications, 2021.

[25] Abhijit Banerjee, Esther Duflo, and Garima Sharma. "Long-Term Effects of the Targeting the Ultra Poor Program". In: *American Economic Review: Insights* 3.4 (Dec. 2021), pp. 471–86. DOI: 10.1257/aeri.20200667. URL: https://www.aeaweb.org/articles?id=10.1257/aeri.20200667.

[26] BRAC. *Overview: Ultra-Poor Graduation Programme, Bangladesh*. 2021. URL: http://www.brac.net/program/wp-content/uploads/2021/09/UPG-programme-Bangladesh-overview.pdf.

[27] Clare Balboni, Oriana Bandiera, Robin Burgess, Maitreesh Ghatak, and Anton Heil. "Why do people stay poor?" In: *The Quarterly Journal of Economics* 137.2 (2022), pp. 785–844.

[28] Abhijit Banerjee, Dean Karlan, Robert Osei, Hannah Trachtman, and Christopher Udry. "Unpacking a multi-faceted program to build sus-

tainable income for the very poor". In: *Journal of Development Economics* 155 (2022), p. 102781. ISSN: 0304-3878. DOI: https://doi.org/10.1016/j.jdeveco.2021.102781.

[29] Thomas Bossuroy, Markus Goldstein, Bassirou Karimou, Dean Karlan, Harounan Kazianga, William Parienté, Patrick Premand, Catherine C Thomas, Christopher Udry, Julia Vaillant, et al. "Tackling psychosocial and capital constraints to alleviate poverty". In: *Nature* 605.7909 (2022), pp. 291–297.

[30] Scott Macmillan. *Hope Over Fate: Fazle Hasan Abed and the Science of Ending GLobal Poverty*. Rowman & Littlefield, 2022.

[31] John Mullahy and Edward C Norton. *Why Transform Y? A Critical Assessment of Dependent-Variable Transformations in Regression Models for Skewed and Sometimes-Zero Outcomes*. Working Paper 30735. National Bureau of Economic Research, Dec. 2022. DOI: 10.3386/w30735. URL: http://www.nber.org/papers/w30735.

4 | Is employment a panacea for poverty? A mixed-methods investigation of employment decisions in South Africa

This paper was developed from my MPhil thesis, for which I was awarded the Papiya Ghosh Prize for the best MPhil thesis. It was published in World Development in 2020.

Abstract

Unemployment is a key determinant of poverty in South Africa and labour market inequalities reflect deep-rooted socio-economic inequalities. In a context of high rates of poverty and unemployment, we would expect a job loss to be associated with a decline in wellbeing. Using nationally representative panel data and original qualitative data collected in Cape Town, I find that, on average, this hypothesis holds. However, this aggregate effect conceals heterogeneities in the relationship between labour market transitions and wellbeing which are of special analytic interest. In particular, this study focuses on those cases which go against the grain of the overall labour market-wellbeing nexus – that is, cases in which black, urban youths turn down or quit wage work. An analysis of these examples helps illuminate how disadvantaged workers face non-negligible disincentives to certain forms of low-skill employment and reveals the circumstances under which these disincentives may outweigh the disincentives to unemployment. To aid this investigation, I develop a model which analyses the welfare effect of job losses as being jointly determined by the strength of outside options and disincentives to work. Using qualitative data, I provide evidence in support of this model and show that, under certain circumstances, transitioning out of employment will be the welfare optimising choice for workers: Younger workers with no dependants and with alternative sources of support can be said to have stronger outside options, and are especially likely to turn down or quit low-quality jobs. Older workers, with dependants and without alternative sources of support, are more likely to accept and persist in low-quality jobs. This study argues that understanding the complexity of the incentives that workers face and which inform labour market choices will be indispensable in effectively designing policies which aim to reduce inequalities in labour markets – in South Africa and beyond.

Acknowledgements: I am grateful to two anonymous referees, Pramila Krishnan, Hannah Dawson, Vimal Ranchhod, Murray Leibbrandt, Joshua Budlender, David Neves, Debbie Budlender, participants at talks and conferences including the 2017 HDCA Conference in Cape Town, the 2018 HDCA Conference in Buenos Aires, the 2018 CSAE Conference in Oxford, the 2019 UNU-WIDER conference in Bangkok, and at SALDRU, ReSEP and Oxford African Studies Centre sem¬inars, as well as several anonymous reviewers at the Oxford Department for International Development for helpful suggestions on earlier drafts. Nina Zizzamia provided excellent assistance with proofreading. Arden Finn provided helpful Stata code for cleaning and deflating National Income Dynamics Study data. I am especially grateful to Simone Schotte, who has been an inimitable partner in both the design of this research project and in running field research in Khayelitsha. I am also indebted to an exceptional team of fieldworkers, among whom Mzulungile Cabanga, Sibongile Mthini, Andiswa Mtini and Amanda Moocha deserve special mention. Responsibility for any errors remains my own.

4.1 Introduction

In high-unemployment developing countries, employment creation is almost universally held up as the key to achieving meaningful poverty reduction. For example, in South Africa – the country study which is the focus of this article – it is seldom questioned that inclusion in the labour market is a 'good thing'. This study sets out to investigate whether economic inclusion through employment is the panacea for the poor that it is so often claimed to be. Because of the relatively greater dependence of households on wage employment in urban areas and the greater dynamism of urban labour markets, this study focuses on answering this question specifically in an urban context.

If workers often choose to turn down or leave wage employment, this is an important potential determinant of both unemployment and employment dynamics. Consequently, this will have profound implications for the effectiveness of strategies which hope to address unemployment through job creation. This study attempts to understand the conditions under which disadvantaged South Africans choose to turn down economic 'inclusion' in the form of wage employment. These decisions are assessed as being determined by two distinct but interrelated factors. First, a decision to leave or turn down wage employment will be more likely if the benefits are not significantly greater than the costs of employment. Second, those with stronger outside options may be less tolerant of 'bad' forms of work relative to those with weaker outside options. This model of behaviour is consistent with the standard neoclassical model of labour choice, where agents make a utility maximising decision by weighing the costs and benefits of an employment option against the costs and benefits of an outside option. However, both

the costs of employment and the strength of outside options are often not amenable to quantitative measurement. For instance, costs of employment might include psychological burdens while outside options may be determined by the strength of complex and unquantifiable social networks. One consequence of this is that, if workers value their psychological wellbeing over their wage employment, they may make welfare-maximising choices which lead to decreases in observed income or consumption. This may lead to misleading results in a standard quantitative analysis. Furthermore, the relatively high levels of social protection through state grants and the high rates of formal employment mean that the outside options of the unemployed are often stronger than would be the case in other African contexts, making this an interesting case for studying this question.

To ensure sensitivity to these considerations, this study employs a mixed-methods approach. National Income Dynamics Study (NIDS) panel data is used in an initial quantitative analysis and a subsequent qualitative analysis is undertaken on original data from in-depth life history interviews (LHIs) and focus group discussions (FGDs) which were collected in a research project implemented in Khayelitsha, a large black township in Cape Town, in 2017.

I find that urban, unskilled, black workers can occupy one of two stylised states: On the one hand, those with weak outside options will be likely to accept 'bad' forms of wage employment. Older workers, those with dependents, those with less support in the household, and those with weaker social networks can be said to have weaker outside options, and are hence less empowered to turn down even the lowest paid, most degrading and precarious forms of work. On the other hand, those with stronger outside options are more likely to avoid wage employment. Younger workers, those with few dependants, those with alternative sources of support and those with strong

social networks appear to have stronger outside options and are consequently more likely to turn down 'bad' jobs.

However, these differences are complicated in a dynamic perspective, since one of the features most strongly associated with 'bad' forms of employment is precariousness. Precariousness in this context is understood to mean a high exogenous likelihood of losing employment. A core hypothesis of this study is that the welfare impact of this precariousness is fundamentally different for those with strong versus those with weak outside options. For those with relatively weak outside options, the loss of a job will have a severe negative impact on wellbeing, while for those with relatively strong outside options, wage employment is often deliberately temporary and might be taken up to smooth consumption, so that precariousness will have a less severe negative impact. The analysis of qualitative data provides strong support for this hypothesis in the South African urban context.

This study draws on two strands of literature: On the one hand, I draw on literature which evaluates labour market behaviour using neoclassical models of labour-leisure choice, such as work done on labour market responses to social grants in South Africa (Bertrand et al., 2003; Posel et al., 2006; Klasen and Woolard, 2009; Abel, 2019). On the other hand, this paper also situates itself within a literature which adopts a livelihoods approach (Chambers and Conway, 1992) to understand the role of the labour market within a 'portfolio' of livelihood strategies adopted by the poor (Du Toit and Neves, 2009; Ferguson, 2015).

I draw on each strand of literature as I bring novel qualitative evidence to bear on the hypothesis that when an individual's expected utility from unemployment exceeds their expected utility from employment, the agent

will choose not to work. In doing so, this study contributes towards the reconceptualization of the social and economic value attributed to work in high inequality and unemployment contexts. At the heart of this reconceptualization is the argument that, for the unskilled urban poor, certain forms of wage work can be considered as 'survivalist', in a similar sense that certain forms of unprofitable informal self-employment are conventionally considered to be survivalist. That is, like other activities where the benefits are only marginally greater than the costs, workers consider wage employment as a 'second best' option, to be avoided when possible. Understanding that low-skill employment may be seen by many as an aspirational compromise is imperative if efforts to combat unemployment through job creation are to be effective.

4.2 Contested terrain: the meaning of work in South Africa

In 1994, the ANC-led government inherited an economy which did not have the capacity to absorb the millions of unskilled, unemployed black South Africans who were rapidly entering the labour market following the abolition of influx control in 1986. The emergence of high rates of unemployment in the 1980s can be explained by a combination of this surge in supply and a stagnation in demand due to the shift to an increasingly capital-intensive growth path in the 1980s (Seekings and Nattrass, 2005). This shift crowded out the labour-intensive, low-wage industries which had greater absorptive capacity for low-skill labour. In keeping with this trajectory, the ANC government which won power in 1994 strove to promote an industrial growth path that was internationally competitive, export-led, high-wage and high

value-added. By reproducing a growth path with a limited capacity to absorb unskilled workers, post-apartheid economic policy effectively sustained the structural basis of unemployment.

However, the state did commit to deracialising and vastly expanding the provision of social grants – most important among them the state old age pension, disability grant, and child support grant. Today, one third of all South Africans receive a state grant, over fifty percent live in households with grant income, and grants comprise the most important source of income for half of these households (SASSA, 2018).

These two facts – high unemployment combined with high rates of formal employment (relative to other counties at similar levels of development), and relatively robust social protection - characterise much of economic life in South Africa. Du Toit and Neves (2009), Ferguson (2015), and Ferguson and Li (2018) argue that formal social protection (in the form of grants) and formal employment function to facilitate and lubricate systems of informal social protection through social networks. Through informal distributional channels, many of those without direct access to wage employment or grant income are nevertheless able to benefit from systems of redistribution which are themselves sustained by social grants and income from the formal labour market.

Perhaps the most widely accepted view on the effect that South Africa's system of social protection has on employment is that access to social grants within households has unintended negative impacts on labour market participation amongst prime-aged adults (Bertrand et al., 2003; Abel, 2019). Part of this effect occurs through migration and household responses, which function as a safety net to the unemployed, who often in-migrate into households

with grant access (often leading to the impoverishment of these households) (Klasen and Woolard, 2009). However, this narrative remains contentions, with conflicting evidence suggesting that social grant access can also lead to increased labour market participation through job-search driven migration amongst prime aged adults, especially women (Posel et al., 2006; Eyal and Woolard, 2011). South Africa's relatively robust systems of informal social protection may also help explain why, unlike in other African contexts, informal employment does not absorb the majority of economically active adults who are not in formal wage employment. Only 17 percent of the employed population is employed in informal work, which is very low by developing country standards and contributes to South Africa's outlier status with regard to unemployment (Rogan and Skinner, 2018).

A dominant school of thought holds that the challenge of unemployment will be resolved by generating wage employment – even if at poverty wages. This argument assumes that the unemployed will be willing to take up the sorts of jobs being created. However, to the extent that job creation for the poor has occurred, this has largely consisted in the creation of precarious forms of work in low-skill service and retail sectors (Theron, 2010; Tregenna, 2010; Finn, 2015). Even in the ostensibly formal sector, 'informalised' employment arrangements – in which work is out-sourced through recruitment agencies and labour brokers – are becoming increasingly common. Bhorat et al. (2016) find that 13.7 percent of total employment growth between 1995 and 2016 has occurred through outsourced and subcontracted jobs, and that poorer, younger and more vulnerable workers are disproportionately likely to find employment in these jobs. If the kinds of jobs being created are of poor quality, it is not a foregone conclusion that the unemployed masses will flock to these jobs.

In South Africa and beyond, an emerging body of literature in both economics and economic anthropology has begun to investigate why poor workers sometimes shun wage jobs, and what these decisions reveal about the characteristics of these workers and the jobs they can access. For instance, Blattman and Dercon (2018), using experimental evidence in Ethiopia, find that industrial jobs had no positive welfare effect on those who were offered these jobs and that most had quit after 11 months. They suggest that low-paid and physically taxing industrial work is used primarily as a welfare smoothing strategy by those who suffer temporary idiosyncratic shocks. Since wage employment essentially functions for these workers as a stop-gap measure, as soon as the immediate economic stressors which forced them to enter wage employment in the first place are no longer pressing, employees quit their formal sector jobs – often even before finding alternative employment options. Those who did not quit are typically older, less able and less skilled – suggesting that they have inferior outside options relative to those who quit. Teal (2018), using census and panel survey data, finds that the lack of expansion of wage employment in Ghana has largely been due to the superior employment growth of own account workers, since those who can access capital to start their own businesses tend to prefer self-employment to wage employment. Like Blattman and Dercon (2018), Teal (2018) finds that 'there is no reason to think firm wage employment is the preferred outcome for most workers' (p. 21). Dawson (2018) lends further support to this argument. Drawing on ethnographic research in urban Gauteng, she finds that many urban South African youth may 'choose' unemployment: opting for informal, township-based livelihood portfolios, rather than wage jobs with their attendant stresses and indignities. That wage work for the unskilled might be seen as a 'second best' option should not be particularly surprising

in a context in which low-skill workers are disproportionately likely to perceive their jobs to be insecure, unfairly paid, uninteresting, and lacking in opportunities for upward mobility (Mncwango, 2016).

The existence of informal social protection, lubricated by income earned through formal wage work and through social grants, creates systems of support to unemployed South Africans which are not available in contexts in which social grants and formal employment is less common (Bertrand et al., 2003; Klasen and Woolard, 2009; Abel, 2019). Effectively, these forms of informal social protection can be understood as strengthening the outside option to formal employment and raising reservation wages, making it less likely that job-seekers will accept 'bad' jobs.

4.3 Methods and data

4.3.1 Quantitative data

The quantitative element of this study uses NIDS panel data (SALDRU, 2018a; SALDRU, 2018b; SALDRU, 2018c; SALDRU, 2018d; SALDRU, 2018e). NIDS began collecting data in 2008 with a nationally representative sample of over 28,000 individuals in 7,300 households. I use the full five waves of data, which are each spaced approximately two years apart. In some of the analysis, data from pairs of consecutive waves t - 1and t are pooled, such that the analysis of changes over time represent changes between 2008 and 2010/11, 2010/11 to 2012, 2012 to 2014/15 and 2014/15 to 2017 respectively, controlling for period-specific fixed effects. The analysis in Section 4 uses poverty status based on consumption expenditure as a proxy for wellbeing. Poverty is defined using the StatsSA upper-bound poverty line

(UBPL) or the StatsSA food poverty line (FPL) (Stats SA, 2015). Expressed in March 2017 Rands, these poverty lines are respec-tively R1136 and R515 per capita, per month.

Panel weights are used to correct for the presence of attrition in NIDS. Of the 26,775 sample members who were successfully interviewed in 2008, 15,673 were re-interviewed in all four subsequent waves, giving an attrition rate for the balanced panel of 41.47 percent. However, between-wave attrition, which is more relevant when pooled wave-to-wave transitions are used, is lower, ranging from 9.3 to 21.1 percent.

4.3.2 Qualitative data and methods

Field research was undertaken in Khayelitsha, a young, large, and rapidly growing township about 25 kilometers from Cape Town's city centre. Khayelitsha was established in 1985 as a site for relocations from other overcrowded black townships in Cape Town. Khayelitsha has a high share of young people and is largely made up of rural to urban migrants – almost all of whom are from the Eastern Cape. Khayelitsha has high levels of income poverty and is facing serious socio-economic challenges. According to 2011 Census data, only 36 percent of residents above the age of 20 have completed high school, only 62 percent of the labour force is employed, and 55 percent of households live in informal dwellings.

To select study areas within Khayelitsha I developed a sampling frame which would ensure that my sample represented the local socio-economic diversity. Using this sampling frame, 15 study areas were selected and from within these 300 households were identified and surveyed for basic demographic and contact information. These 300 households constituted the sam-

ple which was used in further stages. Details on sampling and surveying methods are covered in detail in the Appendix.

In the second stage of field research, four FGDs, ranging between 3 and 10 participants, were held with groups which were selected from the base sample. The purpose of these group discussions was two-fold: First, using visual aids in the form of four boxes arranged in ascending order from 4 (lowest box) to 1 (highest box), participants were prompted to discuss what it meant for people to be in different 'levels' of society (see the Appendix for more details on this exercise). Second, participants were prompted to discuss what determines movements between boxes – both upward and downward mobility. In both elements of the discussion – defining boxes and discussing dynamics – participants drew on their own experiences and the experiences of people known personally to them. This exercise was designed to elicit insights into the determinants of movements between welfare states, the nature of risk faced by the urban poor, and the role of social net-works in determining economic mobility.

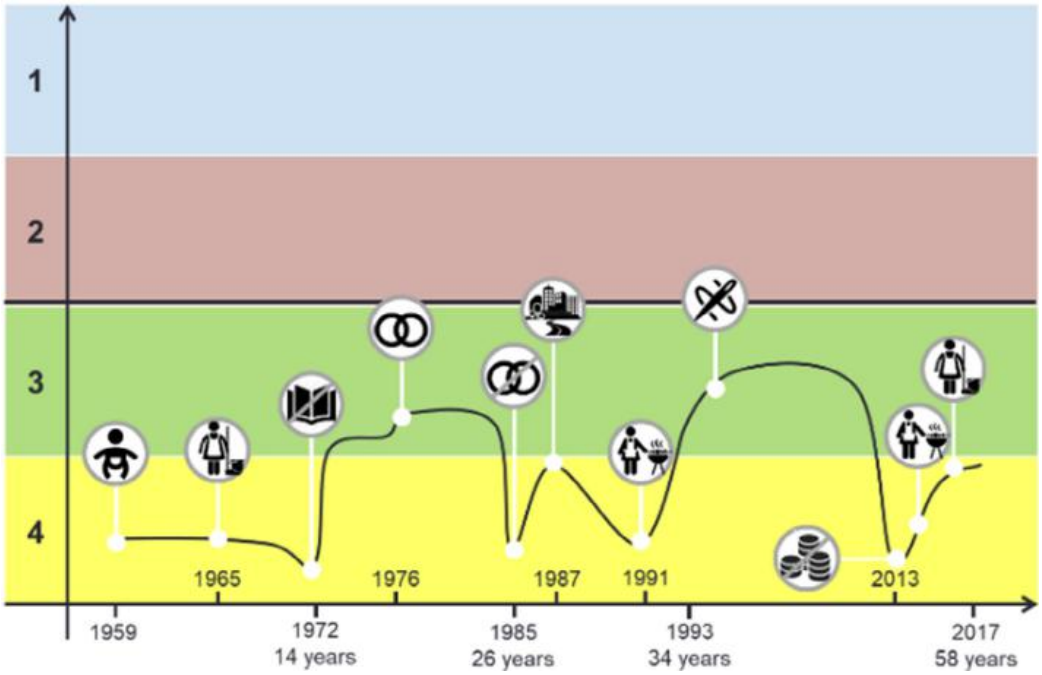
These FGDs were also intended to help develop a means of measuring social and material wellbeing which could be used in LHIs. This measure of wellbeing was intended to be more subjectively meaningful than narrow money-metric proxies of income and expenditure, while at the same time facilitating a degree of comparability between cases which would allow an analysis to draw con-clusions based on comparisons. The box exercise proved to be well-suited to this purpose.

Having arrived at a definition of welfare as described above, I undertook the primary element of the field research, consisting of semi-structured LHIs. These methods drew on the LHI techniques used by Davis (2011) and Davis

and Baulch (2011) in rural Bangladesh and on Adato et al. (2007) 'household events mapping' technique used in rural KwaZulu-Natal. In both methodologies, the respondent/household recounts the individual's/household's history, focusing on events and processes which determined individual/household welfare. Making use of a large sheet of blank paper, major events, processes and coping strategies are mapped chronologically. In addition to the x-axis representing time, Davis (2011) also incorporated a y-axis rep-resenting wellbeing. This allows the interviewer and interviewees to trace together the fluctuations of wellbeing over time and to link these fluctuations to their determinants.

LHIs were structured in two stages. In the first stage, the LHI was conducted in isiXhosa by a fieldwork assistant while either I or a second fieldwork assistant used coloured flashcards to make notes of major life events. In the second stage, the participatory welfare mapping exercise was introduced to the interviewee, in which a large sheet of paper was placed on the floor and an x-axis for time and a y-axis representing the four 'wellbeing boxes' was drawn (Fig. 1). The coloured flashcards were then placed on the graph in chronological order, which the interviewee was asked to use to trace their life-course welfare trajectory. Photographs were taken of this welfare map for use in the analysis of determinants of welfare changes. A stylised representation is shown in Fig. 1.

Figure 4.3.1: Mapping welfare trends over time



Notes: Post-analysis representation by the author of a life-course wellbeing map sketched during life-history interviews with the respondent's participation.

Table 4.3.1: Sample means - Qualitative sample compared to NIDS for African adults in Cape Town (2017).

	Mean for sample	African mean for Cape Town
Age	45	38
Household size	5	4
Educational attainment (highest grade)	7	10
Male	57%	43%
Rural-to-urban migrant	73%	74%
Working for pay	13%	43%
Self-employed	17%	4%
Unemployed	60%	12%
Not economically active	23%	34%
Informal dwelling	43%	37%
Household receives pension	27%	20%
Household receives disability grant	30%	13%
Household receives CSG	47%	49%
Household receives remittances	23%	11%
Household expenditure	R3,069	R6,024
Expecting upward mobility within 2 years	67%	75%

Notes: Calculations using data from author's sample of 30 respondents surveyed prior to qualitative interviews, and from NIDS Wave 5 (2017). NIDS sample limited to urban African respondents resident in the City of Cape Town district council and aged over 18 years. Averages from NIDS are weighted using Wave 5 post-stratified weights.

In addition, each of the 30 individuals who participated in life-history interviews were also administered a detailed survey, designed for comparability with NIDS. Drawing on this survey data, Table 1 presents summary statistics on my qualitative sample and compares these to summary statistics from NIDS. The statistics from NIDS are from the 2017 survey wave and are limited to black adults resident in Cape Town. While neither sample is representative at any level, this comparison nevertheless helps identify the particularities of the qualitative sample used in this analysis.

The overarching finding based on this comparison is that the qualitative sample used in this analysis is substantially more dis-advantaged than the average for Cape Town. Total household consumption is half that of the NIDS sample. Household size is larger, educational attainment is lower, and reliance on grant income is higher. Most importantly, however, is the fact that unemployment is dramatically higher in my sample. On one level this is appropriate, since this is a study on unemployment. However, my findings will also need to be interpreted with an acknowledgement that the qualitative sample has a weaker than average attachment to the labour market.

4.4 Poverty dynamics and employment volatility - a quantitative perspective

Stats SA (2017) estimated that, in 2015, 55.5 percent of the South African population was poor. Behind cross-sectional poverty figures such as these, a great deal of economic mobility is hidden. Fortunately, NIDS follows individuals over time, allowing us to observe changes in their economic wellbeing. Table 2 reports a poverty matrix for between-wave poverty transitions. This

distinguishes between three gradations of poverty – the non-poor, the poor, and those in extreme poverty[1] – and illustrates the extent of movements between these gradations. The most stable categories are the non-poor and those in extreme poverty. 64 percent of those in extreme poverty and 77 percent of the non-poor remained in the same state, while 41 percent of the poor remained poor, with 31 percent moving into extreme poverty, and 29 percent moving out of poverty.

Table 4.4.1: Poverty transition matrices for South Africa, 2008–2017

		t			
		Extreme poor	Poor	Non-poor	Total
$t-1$	Extreme poor	63.4	26.1	10.5	100
	Poor	30.6	40.7	28.7	100
	Non-poor	7.2	16.0	76.9	100

Notes: Poverty is defined using the StatsSA upper-bound poverty line (UBPL) or the StatsSA food poverty line (FPL) (StatsSA, 2015b). Extreme poverty refers to those whose monthly per capita consumption is below the food poverty line. Expressed in March 2017 Rands, these poverty lines are respectively R1136 and R515 per capita, per month. Author's calculations using NIDS balanced panel for waves 1–5 (weights corrected for panel attrition).

Complementing this, Table 3 presents results of an investigation into employment volatility which is summarised by generating six dynamic employment categories based off the number of NIDS waves in which an individual is observed to be employed. Using the balanced panel and considering the sample of 3591 adults with valid employment responses for each wave, we find that 30 percent were observed to be employed in all five waves, with an additional 16 percent employed in four out of five waves. 13 percent of this sample had never been employed, and 14 percent had not been em-

[1]Extreme poverty refers to those whose monthly per capita consumption is below the food poverty line (R515), meaning that meeting basic caloric requirements will not be possible even if all expenditure is dedicated to food. For more details, see Stats SA (2015). Those in extreme poverty are a subset of those in poverty such that all those in extreme poverty are classified as poor, but not vice versa.

ployed in four out of five periods. 27 percent of the sample transitioned into and out of employment more frequently – being employed in either two or three waves. There is thus evidence of substantial employment volatility in the working-age population, as well as high levels of chronic unemployment. This aggregate perspective is useful as a benchmark against which to compare the labour market patterns of the specific sub-populations: Table 3 shows that black South Africans, those living in rural areas, those in female headed households and those with low levels of education are substantially more likely to experience chronic exclusion from the labour market or volatile employment compared to those in urban or white households, those with secondary or tertiary education or those in male headed households.

Table 4.4.2: Number of periods employed

	Always employed	4	3	2	1	Never employed	Obs.
Total	29.7	16.0	14.4	12.6	13.9	13.3	*3595*
African	27.6	16.7	14.6	13.3	14.2	13.6	*2970*
Non-African	40.2	12.5	14.4	10.4	12.6	9.9	*594*
<Matric	21.6	15.7	15.5	14.2	16.5	16.7	*2728*
Matric	37.9	16.9	13.1	12.5	11.4	8.2	*517*
Tertiary	60.4	16.8	10.8	5.2	3.6	3.1	*345*
Youth (16-24)	5.7	11.4	15.0	22.8	24.0	21.1	*2089*
Prime (25-50)	29.7	16.0	14.4	12.6	13.9	13.3	*3591*
Older (51-64)	8.0	7.7	12.6	13.2	24.2	34.3	*1214*
Female	20.8	14.0	14.7	14.4	18.2	17.9	*2428*
Male	44.3	19.4	13.9	9.8	6.9	5.8	*1167*
Urban	36.8	16.8	15.1	10.9	10.3	10.0	*1666*
Rural	15.1	13.1	11.8	16.7	22.3	21.0	*1313*
Poor	17.7	16.2	16.4	16.1	16.4	17.12	*2748*
Non-poor	51.2	15.6	10.9	6.4	9.5	6.41	*843*

Notes: All cell proportions are weighted using Wave 5 panel weights. Except for results presented for different age categories, all other results apply only to adults aged between 25 and 50 years in 2008. "Non-African" identifies Whites and Coloureds. The Indian sample is small, and has been omitted. "Poor" and "Non-poor" categories are defined using the StatsSA Upper Bound Poverty Line (R1136 in March 2017 Rands) and per capita household consumption. All restrictions applied (Column 1) are defined using Wave 1 variables (2008).

As expected, there is a close association between job losses and declines in wellbeing. Table 4 reports initial values of three welfare indicators – poverty,

per capita consumption, and life-satisfaction – and the changes in these indicators associated with job losses.[2] Table 4 disaggregates changes in wellbeing associated with job loss across four demographic groups: black men between the ages of 16 and 24, black women between the ages of 16 and 24, black men between the ages of 25 and 50, and black women between the ages of 25 and 50. These groups are selected for their relevance to the qualitative analysis in later sections, where the link between labour market events and wellbeing is shown to be closely linked to gender and life-course dynamics.

The first column of Table 4 reports initial values of poverty, per capita consumption, and subjective life-satisfaction for each demo-graphic group.[3] By definition, these values are calculated during a period of employment. Poverty is higher and per capita consumption is generally lower among young adults compared to prime-aged adults, and among women compared to men. Mean life-satisfaction, ranging from one (least satisfied) to ten (most satisfied), is concentrated around 5 for all groups, and is highest for young men.

[2]A job loss is defined as a transition from a state of being employed to a state of either being unemployed or being economically inactive.

[3]Note that, because individuals in this sample were employed in the initial period, the chronically unemployed are by definition excluded – which explains why poverty rate is slightly lower among this sample than the population as a whole.

Table 4.4.3: Changes in indicators of wellbeing co-occurring with job loss

		Initial Value	Change associated with job loss	t	n
Poverty rate	Men (16-24)	58.4%	1.6 pp	0.7	507
	Women (16-24)	61.8%	7.3 pp	2.8	435
	Men (25-50)	52.4%	7.2 pp	4	1112
	Women (25-50)	59.1%	4.5 pp	2.7	1779
P.c. expenditure (ZAR)	Men (16-24)	1775.7	-145.7	-1.7	507
	Women (16-24)	1802.1	-194.9	-2.6	435
	Men (25-50)	2389.6	-534.1	-2.5	1112
	Women (25-50)	2228.2	-557.1	-4	1779
Life satisfaction	Men (16-24)	5.5	-0.9	-2.8	507
	Women (16-24)	5.3	-0.1	-2.4	435
	Men (25-50)	5	-0.6	-4.6	1112
	Women (25-50)	5.2	-0.2	-4.2	1779

Notes: Author's calculations using NIDS waves 1-5 pooled panel of wave-to-wave transitions. Limited to the African sub-sample. Life-satisfaction reports the response to the question: 'Using a scale of 1-10 where 1 means 'Very dissatisfied' and 10 means 'Very satisfied', how do you feel about your life as a whole right now?' t-statistics report the results of a t-test against the null of no change between periods.

Unsurprisingly, poverty rates increase for all groups following the loss of a job. Prime aged men and young women experience the largest increases in poverty rates, of approximately seven percentage points. Young men, however, experience only a small increase in their poverty rate, of 1.6 percentage points. Changes in per capita consumption associated with a job loss reveal that prime-aged women and prime-aged men experience the largest decreases in consumption, although off the base of higher initial average consumption. Young men are again less severely affected. These descriptive statistics suggest that young men do not, on average, experience job loss as a major negative economic event. Older workers, on the other hand, experience job loss as a more intensely negative economic event, precipitating a substantial rise in the poverty rate for these groups.

These findings are not mirrored in the descriptive results focusing on life-satisfaction. Following the loss of a job, the decline in life-satisfaction for men is significantly higher than for women in both age groups. Young men, while

experiencing on average only mild declines in economic wellbeing following the loss of a job, experience large declines in life-satisfaction – falling almost a full step on the ten-step ladder. On the other hand, women in both age groups experience only small changes in life-satisfaction. This suggests that the decline in life-satisfaction for young men in particular is not being driven by a large deterioration in economic circumstances, but by something else.

To understand the roots of this puzzle we need to recognise that wage work has a social value which extends beyond its narrow economic significance. For South African men, this social value is tied to norms of masculinity and power which were historically forged during the apartheid system of migrant labour, in which men returned from work in the mines and industries with the economic power to support themselves and their households. As Ferguson (2015) argues, for South African men, 'social membership and adult masculinity alike have long been linked to a highly valorised capacity for wage labour' (p. 235). On this reading, unemployment denies men access not only to resources, but also to labour-based social membership. At the same time, however, it is highly significant that transitioning out of employment is not associated, on average, with a significant change in material well-being for young black men. How these tensions resolve themselves through labour market behaviour, especially in a context in which work itself is often of a low quality, will be further explored in subsequent sections.

4.5 A model for analysing worker employment decisions

The evidence in the preceding section shows that job losses are associated with declines in wellbeing. However, there appears to be substantial heterogeneity in this association across demographic groups and across welfare metrics. Adding further complication, evidence from LHIs collected towards this project suggests that poor, young, urban, black workers choose to leave work surprisingly often. In a high unemployment context, this finding puts strain on the naïve assumption that job creation and poverty reduction go hand-in-hand. This section proposes a more textured understanding of the relationship between employment and wellbeing, and in doing so develops an account of the conditions under which jobs are good for the poor, and under which conditions they might not be.

At the outset, it ought to be acknowledged that evidence from LHIs, in the most general sense, confirm the standard view that, for the majority of households, labour market events dominate as determinants of major trajectory-shifts in economic wellbeing. In graphs tracing life-course welfare trajectories, most major troughs are associated with negative labour market events, just as major peaks are associated with positive labour market events. Both the pecuniary and non-pecuniary benefits of wage work are generally affirmed in interviews, making it unsurprising that a narrative has emerged in which economic mobility – both upward and downward – appears to be largely determined by changes in the labour market.[4]

[4]Of the non-pecuniary benefits of employment discussed by participants in LHIs and FGDs, the economic stability and security which respondents reported benefitting from during periods of employment featured most prominently. In several interviews (both FGDs and LHIs), this economic security was leveraged to take out loans to set up busi-

However, both the quantitative and qualitative data analysed in this study provide support for findings in a recent empirical literature which argues that there is little reason to suggest that workers always prefer wage employment to other livelihood strategies, even when wage employment is an option. This finding is compatible with standard neoclassical labour-leisure choice theory, simply affirming that, as long as an agent's expected utility from unemployment exceeds their expected utility from employment, the agent will choose not to work (Cahuc et al., 2014)[5]. Drawing primarily on my qualitative data, part of this paper's contribution consists in bringing novel empirical evidence to bear on this hypothesis, specifically with regard to the often-overlooked psychological costs associated with employment in high inequality contexts.

Implicit in these arguments is the suggestion that those who are employed in wage jobs fall into two categories of workers. On the one hand, workers with weak outside options tend to depend heavily on wage employment when they have access to it. On the other hand, those workers with stronger outside options are unlikely to rely heavily on wage labour, except when a shock leads them to seek wage work as a means of dealing with temporary material adversity. This second group are more likely to leave their jobs as soon as the temporary constraint which drove them into employment is no longer binding. Blattman and Dercon (2018) summarise this distinction,

nesses. At least five respondents reported that the severance pay upon being dismissed helped buffer the negative effect that job-loss had on their wellbeing. The value of gaining experience and on-the-job training features less frequently in life-histories. This finding is likely to be a feature of my sample, which have historically been employed in low-skill jobs where experience may be less valuable in terms of improving employment prospects.

[5]In this framework, the expected utility of employment encompasses the pecuniary and non-pecuniary benefits of working minus the material and psychological costs associated with working, while the expected utility of unemployment encompasses leisure and alternative income generating activities, minus the costs associated with unemployment.

arguing that 'workers with the poorest outside options remain [employed]', while those with stronger outside options 'use industrial jobs as temporary employment to cope with adverse shocks and unemployment spells' (p. 45).

One can consider the strength of this outside option as being jointly determined by human, social and financial capital. Human and social capital increase the likelihood of finding employment in the future, thereby diminishing the risk of not working in the present. Social capital may also be leveraged to provide informal social protection through the redistribution of resources. Individuals may also have direct access to other forms of non-wage income – being a direct recipient of a social grant or engaging in criminal activities, for instance.

For simplicity, I label the combination of these endowments as 'resources'. I also assume for the moment that jobs on offer are identical and of low quality. In this set-up, we might imagine that a resource threshold exists, where those whose resources fall below the threshold do not possess an outside option strong enough to permit them to turn down or leave low-quality jobs. On the other hand, those whose resources place them above the threshold may be inclined to refuse these jobs. Thus, workers can occupy two states depending on which side of the resource threshold they fall onto: Workers with resources which place them below the threshold (for simplicity of exposition referred to as 'resource poor') will engage in wage labour as a core livelihood strategy, while workers with resources at a level above this threshold (the 'resource non-poor') may avoid low-quality jobs and either wait for better work or invest in forms of self-employment.

This simple model predicts static states but fails to explain transitions between employment states for either the asset poor or asset non-poor. For-

tunately, drivers of volatility can easily be incorporated within the model. Workers who belong to the resource non-poor category may transition into unfavourable forms of wage labour if they suffer a shock and temporarily fall below the resource threshold, such that they engage in wage work as a temporary coping mechanism. In this case, employment is likely to be short-lived.

On the other hand, workers who are counted among the resource poor may transition out of employment as temporary jobs end or as they are fired. In other words, this labour market volatility is not being driven by the decision to optimise wellbeing under certain constraints, but is being exogenously imposed on them by the fact that their employment arrangements are typically precarious. However, because wage labour for these workers represents the core element of their livelihood portfolios, the loss of a job will have a severe impact on wellbeing.

From an aggregate perspective, workers in both states are observed to transition frequently into and out of employment, driving aggregate labour market volatility and poverty dynamics discussed in Section 4. However, the consequences of this volatility on individual wellbeing may be fundamentally different.

The welfare effects of job loss will also be determined by the cost/benefit balance of being employed. Benefits will be offset or outweighed by the costs associated with employment to a varying extent across different jobs. Because of this, distinguishing between the resource poor and resource non-poor does not adequately capture the welfare consequences of job losses, since welfare costs are jointly determined by outside options and employment quality. Taking account of both the strength of outside options and the costs

of employment will be necessary to determine how job losses may affect wellbeing.[6]

4.6 Is employment a panacea to poverty? Qualitative evidence

4.6.1 Aggregate patterns: labour market events and wellbeing

Table 5 summarises the association between labour market events and welfare transitions for the 30 individuals whose life-histories provide the material for this section. Table 5 reports the frequency of various labour market events and their associated impact on wellbeing, taken from life-history interviews and the wellbeing trend maps constructed with respondents' participation (such as that represented in Fig. 1 above). To do this, I first identify all labour market events discussed in each interview and then, using both the interview transcription as well as photographs of the map drawn with the participant, a judgement is made on whether the participant felt that each event in question had a strongly negative, negative, neutral, positive, or strongly positive effect on life-course wellbeing. Each of the 30 respondents discuss several labour market events, which, when added up, equal 295 distinctly identified events.[7]

[6]It is important to note that both outside options and costs of employment are dynamic and may change rapidly for any particular individual. Whether costs of employment and outside options determine a worker to be averse to certain forms of employment is descriptive of a state of being which is fluid, rather than descriptive of a static category of person.

[7]Because these judgements are subject to the researcher's interpretation, to prevent my own hypotheses about labour market behaviour from influencing my judgements, another researcher, also involved in the qualitative fieldwork of this project but without knowledge of my hypothesis, independently classified the welfare effects of the various labour market

Table 4.6.1: Incidence of labour market events and associated wellbeing effects

	Total	V. neg.	Neg.	Neutral	Pos.	V. pos.	Key event	
Find (new) job	109.5	0	2	35	55	17.5	18	16%
Lose job	48	9	21	18	0	0	10	21%
Quit job	28.5	3	5.5	19	0.5	0.5	4	14%
Business start	20	0	0	6	9	5.5	6	30%
Business collapse	10.5	7.5	1	2	0	0	8	76%
Fam. starts/switches work	52	0	0	14	27	11	11	21%
Fam. quits/loses work	26	6	11	9	0	0.5	5	19%

Notes: : Calculated using life-course welfare maps and transcriptions taken from the life-history interviews of the author's sample of 30 respondents. Because of concerns of introducing bias from the author's priors, a second researcher, also involved in the qualitative fieldwork of the project but without knowledge of the author's hypotheses, independently classified the welfare effects of the various labour market events. These two estimates where similar. Each cell is the average of these two independently calculated frequencies and independently classified welfare effects. Strongly negative (positive) events are defined as events leading to a major negative (positive) change in wellbeing, while negative (positive) events are defined as events leading to a moderate or small negative (positive) change in well-being. Key events are defined as events causing major trajectory shifts in life-course wellbeing.

As expected, Table 5 shows that, on average, job gains were associated with positive changes in wellbeing and job losses with negative changes in wellbeing. A few findings stand out. First, while most job gains were associated with increases in wellbeing and most job-losses associated with declines in wellbeing, about one third of job gains and job losses were classified as having a neutral impact on welfare. Second, about 28 cases of voluntary job-quits were reported, most of which had a neutral effect on wellbeing. Finally, job losses or gains affecting family members appear to have strong indirect effects on participants, illustrating how informal social protection is sustained by the labour market participation of other household members.

With these aggregate patterns in mind, in the subsections that follow I draw more deeply on evidence from LHIs and FGDs, analysing relevant individual cases to better understand how the benefits to employment may be offset or outweighed by the costs of certain forms of work. In those cases

events. Her and my calculations were then averaged, which is why there are some cells which do not contain whole numbers.

where the degree by which the benefits are greater than the costs is marginal, I investigate how individual resources determine whether workers remain employed or decide to quit. In LHIs, respondents who had chosen to leave work or who had lost their jobs were prompted to explain how they had motivated their decision to quit or what course of events led to the termination of their employment. They were also asked to explain what welfare effects these events had resulted in. From these interviews, it became clear that, especially among the poor, black and urban demographic, the many disadvantages of employment had, for several respondents and in several instances, been perceived as greater than the anticipated benefits of alternative non-wage livelihood activities, thereby precipitating either a voluntary quit or a reduction of effort leading to a job loss. In what follows I identify and analyse some of these costs of employment.

4.6.2 Low, variable wages and expensive commutes

Wages in the kinds of low-skill work in which the participants in the study were employed are generally very low. Approximately one third of those in the life-history sample reported having left jobs because their pay was considered too low.[8] Generally, those who quit were mostly young men who (at the time they quit) had few dependants and had strong sources of support within their own households. One young man, for example, quit his job at a fast-food restaurant because he made more money stealing cars. Other young men left their jobs without having yet found alternative jobs, but could afford to do so because they did not yet have dependants and were being supported by their households. As Table 5 shows, in most cases these job-quits were

[8]Included among these are those workers who were fired because of theft or withdrawal of effort, if the respondent justified this in terms of "getting even" for unacceptably low pay.

not considered negative events.

Two individuals who, during their early working life had left their jobs and classified these events as having a neutral effect on wellbeing, at later points in both their lives experienced the loss of similar jobs, which were also considered unfairly paid. However, in later life these job losses were perceived as negative, sometimes catastrophic, events. The difference between the former and latter job losses was that during the latter job losses, both had children and were supporting wives/girlfriends and had limited alternative sources of support. These examples illustrate clearly how the interaction between job loss and contextual factors jointly determine the welfare impact of a job loss, even when a job is poorly paid. In their more recent job losses, both had much poorer outside options relative to their earlier job losses. These examples are also consistent with the quantitative evidence presented in Table 4 – in which older men experience significantly larger declines in material wellbeing compared to younger men.

Not only are wages low, they are typically highly variable. This volatility imposes a psychological and material burden on poor households, for whom it adds to the stress of getting by on extremely limited resources. A minority of respondents reported having been employed in jobs where they were expected to work regular days and/or hours. When employees are paid an hourly rate, volatility in hours worked translates into earnings volatility.

Flexible employment arrangements also interact with the spatial form of cities to compound the disadvantage of marginal workers. South African cities are characterised by spatial mismatch, meaning that the concentration of the economically active population is highest in parts of the city which are furthest from the highest concentration of jobs and economic opportunity

(Ivan. Turok, 2001; Budlender, 2016; Ivan Turok et al., 2017). Since the poor are most negatively affected by spatial mismatch, the same workers are disproportionately burdened by high transportation costs in terms of commute time and transportation expenses.

The burden of transportation costs increases when working days and hours are irregular. This can happen through two mechanisms: First, it is common for workers to purchase monthly train or bus tickets to travel between home and work. This purchase is a sunk cost, with the consequence that transport expenditure relative to income increases if a worker ends up working fewer hours than anticipated. Second, train and bus travel is usually the cheapest form of transportation available to workers living in townships on the peripheries of cities. However, trains and busses only run reliably during daylight hours, meaning that those working irregular hours and night shifts need to travel on minibus taxis, which can be significantly more expensive. In sum, transport costs for township dwellers can account for a disproportionately high share of their total expenditure. Kerr (2017) calculates that those primarily using minibus taxis experience an effective reduction in their median hourly wage of over 26 percent, while those using multiple forms of transportation (disproportionately including those who work irregular hours) experience a median reduction in hourly wages of 39 percent. Manning (2003) shows that, if workers do not on average receive full compensation for longer commutes, then utility from employment falls in proportion to the length of the commute. An implication of this is that 'if there is, on average, less than full compensation, we would expect to see higher separation rates for those with long commutes' (p.117).

One woman's story is illustrative of the broader narrative which emerged from interviews on how unpredictable and variable hours erode earnings while

at the same time inflating transportation costs. In 2017, she was employed as an on-hold worker by a recruitment agency. She would only work when another of the agency's employees could not work on a particular day. She was also required to travel to different, distant locations across the city depending on where she was needed. Unpredictable working days and varying workplace location made it especially difficult to economise on transportation, as well as imposing a psychological bur-den of uncertainty. Not only were her expenses high, her income was also lower than she had expected. In her words, this was an 'expensive job'. Sharing this sentiment, another respondent dis-cussed how a combination between variable wages and high (and variable) transportation costs 'drove [him] crazy'.

4.6.3 Burden of dependents

The relationship between the accumulation of dependants and attachment to employment is complex. In the data collected toward this study, there is evidence of this relationship working in two directions. On the one hand, the accumulation of dependents, when occurring as an exogenous demographic change, often leads to an increased attachment to work. However, it is not always clear that the accumulation of dependants is exogenous to an individual's labour market status. That is, it may be precisely because of having a job that the household accumulates dependents. Becoming responsible for additional dependants, in turn, increases the welfare cost of losing a job. This potential causal circularity makes identifying the welfare effects of this interaction especially difficult.

One might imagine three stylised scenarios: In the first scenario, an individual is employed and is responsible for a number of dependants, such that the welfare cost of losing a job would be sev-ere and would be shared by

the household members. Evidence of the negative effects that job losses can have on other household and family members is illustrated in the bottom rows of Table 5. If an individual's capacity to support dependants is saturated given their income, then there will be limited scope for the further accumulation of dependants. In this case, the wage earner faces strong incentives not to leave employment.

In the second scenario, an individual is employed, but has no dependants or has fewer dependants than they could potentially support. In this case, the 'risk' of accumulating dependants by remaining employed is non-negligible.

In the third scenario, an individual is part of a household which has other wage earners. The other wage earners in the household only make distributional demands on the individual to contribute their share to the household when that individual is employed. Another way of expressing this is to say that this individual's non-labour income may be higher during a period of unemployment than it is during a period of employment.

Because in the first scenario household absorptive capacity is saturated, it is only in the latter two scenarios that the employed individual faces the risk of being 'taxed' on their income. When combined with some of the other costs of employment discussed here (low pay and high transportation costs, among others), this risk may act as a disincentive to entry into the labour market or to remaining employed.

This hypothesis finds support in the study's qualitative data, where a general pattern emerges in which job losses tend to have more moderately negative effects on wellbeing if another house-hold member holds a job (Klasen and Woolard, 2009). In addition, gaining a job has a more subdued positive impact on wellbeing if an individual's household contains several

dependants who could make claims on a wage-earner's income. Often, as reflected in Table 5, the loss of employment of another household member has a severely negative effect on the wellbeing of a particular participant. All these patterns illustrate that, in the presence of obligations of redistribution within the household, the risk to losing a job is lowered, and the reward to gaining a job is moderated.

The following examples are illustrative: Two individuals report that the overly-demanding distributional claims made by other household members factored strongly into their decision to quit their jobs. A third individual reported that she intended on leaving her job in the near future, and made it clear that she would be sup-ported by other household members during the spell of unemployment she intended to enter, but that while employed she was expected to support herself. All three of these workers were young, had few dependents, and lived in households with other older working members. In addition, these workers were also doing unsatisfying jobs, working irregular hours and for low pay. A combination of low job quality and relatively strong outside options can explain these job-quit decisions.

However, it would be wrong to overlook the significant psycho-logical and social costs of unemployment, even for those with rel-atively strong outside options. As an able-bodied adult, depending entirely on others in a household imposes a heavy psychological burden. For men in particular, in a society where ideas of masculinity are inextricably tied up with the notion of being able to provide for others through work, finding oneself at the bottom of a distributional hierarchy as a young man can be deeply humiliating (Dawson, 2018). This perceived humiliation is worsened when these young men are dependent on elderly female grant recipients – a post-apartheid inversion of distributional hierarchies to which Ferguson (2013) and Ferguson (2015)

and Du Toit and Neves (2009) have devoted careful attention. One young man, for instance, spoke frankly about his psychological state during spells of unemployment when he was dependent on his family: 'When you are a man and you are not responsible, people look at you funny, even your family. They treat you funny, look at you funny, look at you as a no-body. Even your mother will say things that she wouldn't say to you if you were working'.

That unemployment is a particularly heavy psychological bur-den for men to bear is given further support by the quantitative analysis in Section 4. While the material impact of a job loss is actually smaller for young men than it is for young and prime aged women, men experience a significantly higher decline in psycho-logical wellbeing following the loss of a job.

This example serves to illustrate the fraught and complex rela-tionship between labour market events and household dynamics and their effects on welfare. Young workers in particular are sub-ject to conflicting incentives: On the one hand, the risk of becoming the subject of a 'social tax' creates a disincentive to wage work in a similar way that commuting costs can be considered a disincentive by leveraging a tax on income. On the other hand, the perceived shame of unemployment and dependency disincentivises unemployment. How exactly individual workers navigate the horns of this dilemma is often unclear, even to workers themselves.

4.6.4 Perception of exclusion

Dawson, 2018, drawing on ethnographic research with young men in urban Gauteng, finds that, paradoxically, inclusion in the labour market can be experienced as a powerful affirmation of a reality of structural exclusion. It is through wage work that unskilled workers are inserted into South Africa's

so-called 'first world economy'. However, rather than this being experienced as meaningful 'inclusion', it is through this 'inclusion' that disadvantaged workers become acutely aware of the reality of their structural exclusion, the apparent impossibility of meaningful upward mobility, and the rigidly embedded inequalities in the labour market. In contrast, when remaining unemployed and thus 'excluded' from the 'first world' economy, the unemployed draw on complex and improvisatory livelihood strategies within the township economy in order to survive. By engaging in these township-based hybrid livelihood strategies, unemployed young men often feel empowered by being able to exercise agency, feel 'included', and aspire to upward mobility. In my own fieldwork, 'zero-to-hero' stories of those who had 'made it' in the township economy were pervasive.

An unexpected outcome of my fieldwork research may help illuminate whether, for black workers, inclusion in the formal labour market is experienced as an affirmation of exclusion in the way that Dawson suggests it may be. In the data collected in LHIs and FGDs, a puzzling and unexpected tension emerged regarding the importance assigned to wage labour within livelihood portfolios: In FGDs, participants focussed on self-employment as a determinant of upward mobility, while in LHIs the labour market was almost universally seen to drive upward mobility. A confirmation of Dawson's hypothesis may depend on how this apparent contradiction is interpreted.

Based on preliminary quantitative analyses and on intuitive notions of the importance of jobs in determining welfare, I expected the FGDs to focus on the labour market as a core determinant of wellbeing. However, I found that participants spoke sur-prisingly little about the role of jobs in determining wellbeing, and instead centred the discussion on various forms of self-employment. For example, in all four focus groups which were held,

minibus taxi owners, tavern owners and other successful entrepreneurs were considered the emblematic township 'success stories'.[9] Those in stable, formal employment were placed on a rung below the successful entrepreneurial class. When participants were pressed to justify this stratification, they stressed the fact that successful entrepreneurs were autonomous and did not depend on others.[10] Stable employees, on the other hand, were relegated to a lower rung in society by virtue of their dependence on a 'boss'.

Reflecting this occupational stratification, when discussing determinants of upward mobility, participants focused almost exclusively on entrepreneurial pathways out of poverty. Those who could move up were ambitious, business-minded and did not have dependents who soaked up resources which could otherwise be profitably invested. Formal jobs were mentioned as routes out of poverty when they were perceived to be a means of accumulating capital or experience to start businesses (see Teal (2018)). The male-only group emphasised that access to credit to start businesses was only possible if you could demonstrate your reliability as a borrower by 'showing your payslip' to a lender. The younger mixed-gender group stressed that the experience gained on a job was valuable for starting one's own business, citing the example of an entrepreneur whose experience working as a private security guard helped him establish his own successful private security firm. It seems from these discussions that, to the extent that the labour market is seen as a means of achieving upward mobility, this is limited to acting as a causal pathway to success through entrepreneurship.

[9]This perception was common across male-only, female-only, young mixed-gender and elderly mixed-gender groups, suggesting it is not merely a product of masculine valorisation of a 'big man' figure.

[10]Note that a clear distinction was maintained between successful entrepreneurs and those engaged in survivalist self-employment.

However, in sharp contrast to the FGDs, self-employment featured infrequently as a determinant of economic mobility in LHIs (see Table 5). Rather, the type of wage employment available to an individual appeared to be an especially important determinant of welfare. Welfare transitions – especially major, trajectory-shifting changes – appeared to be determined most often by labour market events. At face value, the quasi-obsession with self-employment as a determinant of mobility in FGDs is inconsistent with the near total neglect of self-employment as a determinant of mobility in the LHIs.[11]

An important difference between LHIs and FGDs may explain this apparent tension – while in LHIs respondents discussed what had actually occurred in their lives, in focus groups, discussions were more often hypothetical. That is, while FGD participants were encouraged to speak about other people in the community and about hypothetical scenarios, LHI participants only spoke about their own past – discussing events which had actually occurred to them. It is important to note that most participants in the FGDs did not identify as being high up the welfare distribution. This implies that, when discussing success stories and upward mobility, most participants were looking up at others rather than discussing their own situations. Further, while FGDs took place in a loud and social environment where it is possible that participants fed off the affirmation of other participants, LHIs were typically much more intimate, allowing a calm and detailed recollection of lift events. These differences are relevant in terms of the way that the evidence ought to

[11]It is important to note that informal self-employment generally falls into two categories for the sample: On the one hand, some participants engage in entrepreneurial self-employment and aim to make a profit. In these cases, for instance, capital may be invested from saving from a previous wage job. On the other hand, informal self-employment can be a survivalist option of last resort, undertaken with minimal capital investment and without any intention to expand or advance the business for profit.

be interpreted in the two cases.

For the above reasons, I argue that a plausible explanation for this FGD-LHI inconsistency is that the FGDs were more expressive of aspirations and perceptions of social and economic stratification, while LHIs allow a more accurate description of actual events and a realistic account of the causal mechanisms which prevent or facilitate economic mobility.

With this understanding, those instances where LHIs and FGDs provide evidence which is apparently contradictory are of particular analytic value: For the poor, the most realistic hope of achieving upward economic mobility is through the labour market and, conversely, transitions out of employment typically determine down-ward economic mobility. This is consistent with the quantitative analysis in Section 4 and evidence from LHIs summarised in Table 5.

At the same time, however, for some, it appears that getting such a job is perceived as an aspirational compromise, where there is a preference for the autonomy and socio-economic power that comes with being a successful entrepreneur. However, constraints to capital, spatial mismatch and other barriers mean that, for most workers, this is not a viable option, with the result that many settle for wage work.

This narrative corroborates Dawson (2018)'s argument that, while wage labour is the most feasible means for achieving economic security and mobility, this very same mechanism serves as a confirmation of the more basic fact that the urban poor are unlikely to ever be meaningfully integrated into the urban society in which they sell their labour. Looking back on their lives, most of the LHI participants recognised that their only spells of upward mobility occurred through employment. However, the FGD participants appear

to aspire to something else – the possibility of inclusion within their own society in a way that leaves their autonomy and dignity intact.

Without underplaying the psychological costs to unemployment which are discussed elsewhere, this interpretation suggests that there are also serious psychological costs to employment. While these may operate subconsciously, they may nevertheless be consequential in determining an aversion to wage labour, which is only overcome by acute material need.

4.6.5 The role of weak outside options in job-quit decisions

Of special interest are those participants who, despite being employed in jobs which are characterised by the burdens of employment described above, nevertheless persist in 'bad' jobs. An unambiguous finding is that these decisions to persist are determined by a compromised individual and household resource base. These individuals are older, less educated, less experienced, responsible for larger numbers of dependants, may not have access to grant income, have no or little other employment income in the household, and may have weak social networks. Further, because of the relatively greater importance of labour market income in the livelihood portfolios of the poorest workers during periods of employment, the exogenous loss of a job is observed to have a severe welfare impact.

Most low-skilled jobs for urban workers are temporary, precarious and/or unprotected. The consequence is that, while the most disadvantaged workers are unwilling to leave their jobs, their jobs are nevertheless often of limited duration and hence they transition into and out of employment frequently. For firms, which value labour market flexibility, these workers are often the

last to join the work force, and the first to be laid off. Because these jobs are unskilled and require minimal training, and because unskilled labour is in plentiful supply, these workers are essentially disposable from the firm's perspective. For workers with a limited resource base, however, when job losses occur, they occur despite the lack of an outside option, and can therefore cause large welfare shocks.

Evidence from LHIs support the hypothesis that those who persist in 'bad' forms of wage work have lower resource bases. Of the eight participants who report having worked but never having chosen to leave work, all could be classified as falling among the chronic poor. That is, in their self-reported welfare trajectories, at no point in their life-courses had they risen out of poverty. None of these respondents have completed their schooling, all appear to have relatively weak social networks, and most were over 50 years old at the time of the interview.

Since for these workers involuntary job-losses will have significant negative consequences, they are more likely to attempt to smooth consumption by engaging in survivalist livelihood strategies. While several respondents who voluntarily quit 'bad' wage jobs (as described in the previous subsection) describe subsequent periods of passive unemployment, those who involuntarily lost jobs and had weak outside options report engaging in survivalist self-employment to get through periods of unemployment. One woman, for example, describes selling 'smileys' – roasted sheep heads – on the street. When asked how much profit she made from this business, she appeared to be confused by the question, explaining that 'we must be clear that this business was not profitable – it was a plan to survive'.

4.7 Conclusion

The intimate relationship between unemployment and poverty in South Africa is undeniable, from both a static and dynamic perspective. Just as unemployment is a cause of poverty, so transitions into unemployment are closely associated with descents into poverty. However, it remains possible that some of the unemployed and some of those who leave employment actively choose not to work. In this study, I set out to show that low-skill wage employment is not always the panacea for poverty among unemployed workers on the margins of South Africa's urban economy that many perceive it to be. This study aims to add nuance to our understanding of this relationship not merely as an empirical investigation, but also as a way of revealing what this empirical problem suggests about the way we conceptualise work and unemployment in high poverty, high inequality settings.

This paper situates itself alongside research which has begun to challenge the usefulness of prevailing linear views of economic development in which full, formal wage employment is seen as the ultimate end-goal of economic development. As Ferguson and Li (2018) put it, 'stable waged and salaried employment is losing its plausibility as the universal solution, the obvious telos of a worldwide developmental process' (p. 2). If this is the case in South Africa, then the hierarchical assumptions regarding wage labour's position vis-a-vis other livelihoods strategies may need to be revisited.

The basic finding of this study is that the effects of employment on individual wellbeing will depend jointly on two factors: first, the strength of outside options to employment, and second, the margin by which the benefits of work outweigh (or fail to outweigh) costs associated with low-skilled work.

Younger workers with few dependents, alternative sources of support in the household, and strong social networks through which they can find jobs and make distributional claims for support during spells of unemployment, can be said to have relatively strong outside options. These workers are more likely to turn down or quit 'bad' jobs – even if unemployment remains a heavy psychological burden to bear, especially for young men. On the other hand, older workers with dependants, less support within the household and weaker social networks often have no choice but to accept even the worst forms of work.

For workers on the peripheries of South Africa's cities, both those with strong and those with weak outside options compete for relevantly similar jobs. I have shown that these jobs, primarily in the unskilled service sector, impose on workers a host of burdens which offset, and sometimes outweigh, the benefits of employment. Since these jobs are inherently precarious, both groups of workers are observed to enter and exit employment frequently. However, I show that the welfare effects of this volatility differ for workers with strong outside options compared to workers with weak outside options. For workers with weak outside options, wage work is more likely to represent a core element of their livelihood portfolios when employed, with the result that the loss of a job will have a severe impact on wellbeing. For workers with stronger outside options, on the other hand, wage employment occupies a more peripheral and temporary place in their livelihood portfolios, with the consequence that transitions out of employment have a less severe impact on material welfare.

These arguments have several implications for policy in South Africa and other high unemployment countries. First, minimum wages are likely to benefit both workers with stronger and weaker outside options, but for differ-

ent reasons: For those with stronger outside options, a minimum wage will strengthen their attachment to the labour market, thereby reducing spells of unemployment. For those with weaker outside options, I argue that their relatively strong attachment to the labour market should not necessarily be cause for celebration, but often simply reflects their poverty of choice. By clarifying the supply side explanation for the origins of this attachment, this paper makes a case for improving job quality also for those who are strongly attached to the labour market.

Second, this paper provides reason to expect that policies which reduce spatial mismatch are likely to increase the attachment of workers to the labour market by increasing effective wages without imposing a burden on firms. These policies could include targeted transport subsidies, improving urban transport infrastructure, and increasing urban densification.

Third, low wage employment imposes a psychological burden on workers when it is experienced as an affirmation that meaningful upward mobility is out of reach in a structurally unequal society. Reducing structural inequalities, for instance through progressive taxation and redistributive social spending, may help reduce this psychological burden.

The argument put forward in this paper is more than merely descriptive, but is intended to prompt a deeper reflection into the social value of low-skill work in a high unemployment context. By understanding that the material and psychological burdens of low-skill employment lead many poor workers to consider wage employment as a 'second best' livelihood option, to be avoided if possible, we can begin to see that wage work is little more than a survival strategy for the poor, where the benefits are often only marginally greater than the costs. The fact that millions of workers settle for low-skill wage

work should not be encouraging, but ought to be interpreted for what it is – reflective of the fact that most of these workers are not empowered to choose otherwise.

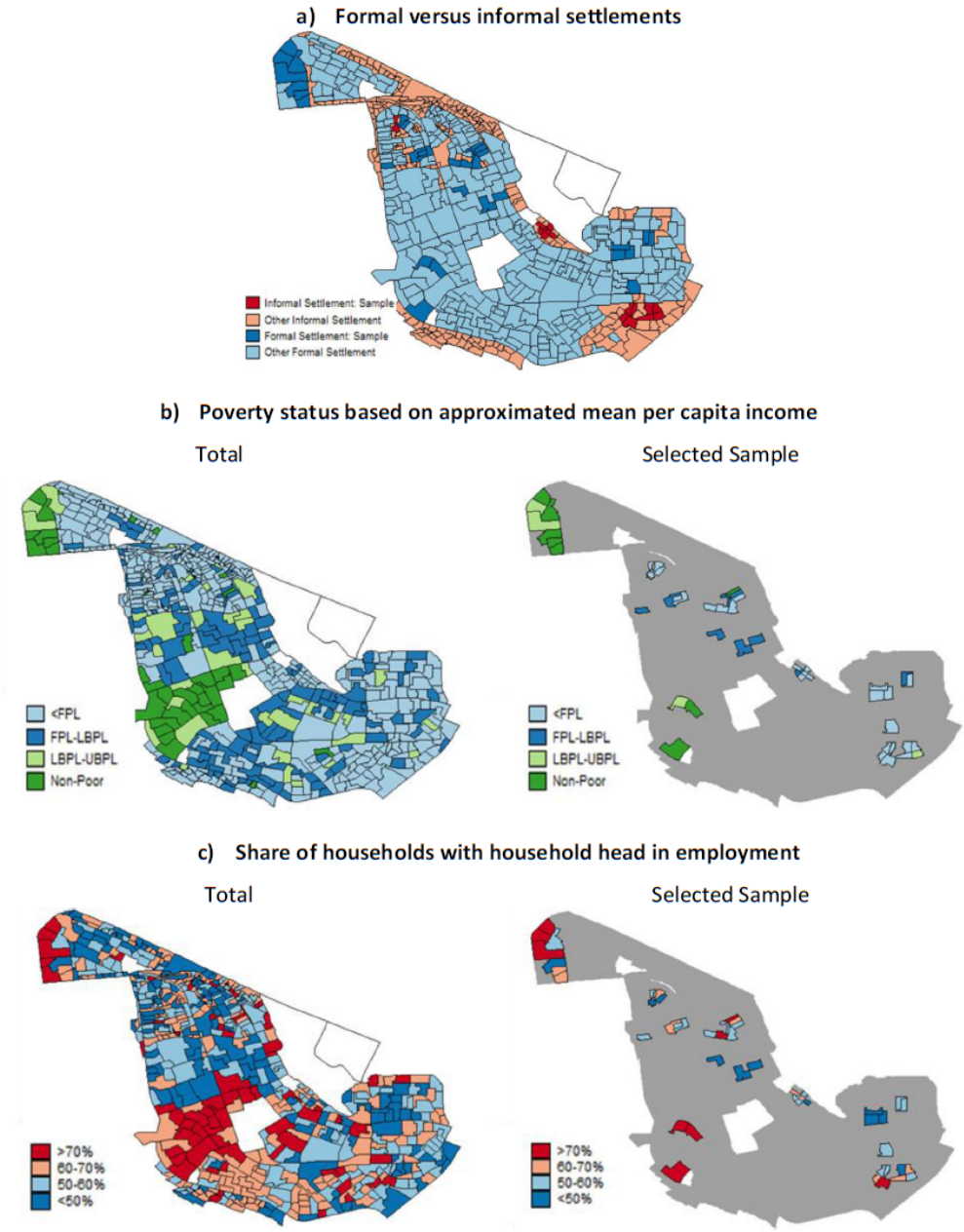
Appendix

4.A Site selection, field research methods and post-collection analysis techniques

In selecting areas within Khayelitsha to conduct interviews, I took care to develop a sampling frame which would ensure that my sample reflected the socio-economic diversity within Khayelitsha and could thus allow me, while falling short of making claims of representativeness, to meaningfully interpret the study results within the broader quantitative analysis.

The maps in Fig. A1 illustrate how the 2011 Census data was used, at the small area, sub-ward council level to select areas within Khayelitsha to sample households. Using this data and maps such as those reproduced below, data was collected in areas which ensured a basic degree of diversity – in terms of educational attainment, poverty status, concentration of informal housing, concentration of rural to urban migrants, household size, and labour market insertion. While this sampling strategy provided a guideline to the selection of areas, the input of field researchers, who live in Khayelitsha, was also invaluable in identifying appropriate areas for sample selection. Fig. A1 shows which areas within Khayelitsha where selected as study sites, and how these were chosen to guarantee a basic degree of socio-economic diversity.

Figure 4.A.1: Sample Selection



Source: Calculations by the author and Simone Schotte, using 2011 South African Census data. Note: FPL abbreviates Stats SA Food Poverty Line, UBPL abbreviates Stats SA Upper Bound Poverty Line. Similar maps were produced for various other variables used to inform sample selection. These maps are available from the author upon request.

Using the sampling frame discussed above, in a first stage, 15 study areas were selected and within these study areas a random walk technique was used to select approximately 300 households to whom short surveys were administered. These short surveys consisted of collecting basic demographic and household information, as well as administering a quick wealth assessment module. These 300 households became the sample which was used in further stages, and the surveys administered in this stage became the basis for the selection of individuals for participation in subsequent fieldwork stages.

In the second stage of field research, four FGDs were held with groups which were selected from the sample collected in the first survey round. FGDs ranged from between 3 and 10 participants. Some groups were single gender, and others were mixed. Discussions lasted approximately two hours each. These were conducted in isiXhosa by a moderator who was intensively trained during preparation for fieldwork. As is common practice among qualitative researchers in South Africa, FGDs were simultaneously translated by a second research assistant, while I transcribed this translation (Du Toit and Neves, 2007; Du Toit and Neves, 2009; Hey and Beinart, 2017). Transcriptions and audio recordings of FGDs were then analysed and discussed with two research assistants in a post-FGD debrief to ensure the accuracy of the translation and transcription and to analyse preliminary findings.

As discussed in the main text, these FGDs were designed to develop a means of measuring social and material wellbeing which could be used in LHIs. The key characteristics of the various boxes which were converged upon in the FGDs are described below:

Box 1:

Those in box one are the economic success stories of the township. They are usually successful entrepreneurs in the community – taxi and tavern owners are especially prominent. Some among those stably employed in permanent and well-paid work may also be considered to belong to this box. The children of those in Box 1 are educated, employed and/or married and live on their own.

Box 2:

Those in Box 2 are neither poor nor wealthy. They are stably employed, usually in a lower-level white collar occupation. Compared to those in box one, they tend to face the need to support a larger number of dependents, which may also include members of the extended family.

Box 3:

Those in Box 3 are considered poor. They have their most elementary needs such as food and electricity satisfied, but still need to economise a lot. They cannot put any money to the side, since all income is consumed. They may match most closely with those households falling below the UBPL. They tend to derive income from the labour market, but are typically in jobs that are poorly paid, of limited duration, unregulated by labour legislation, or subject to seasonal fluctuations. This box may also include smaller households that have access to an old age pension or disability grant.

Box 4:

Those in Box 4 were repeatedly characterised as 'going to sleep on an empty stomach', which most closely matches with the conditions of those households falling below the FPL. They typically have no access to income from the labour market but survive off child support grants and/or support from others in the community, including food donations by neighbours. Some eke out a living from survivalist forms of self-employment and/or do ad hoc piece jobs. These people live hand-to-mouth. They have few assets, live in informal dwellings, and might not be able to afford electricity.

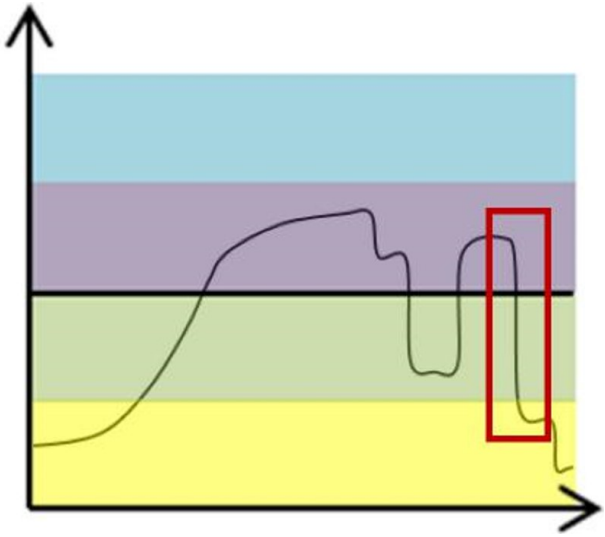
The final stage of research involved in-depth life history inter-views with 30 individuals from the original sample of 300. Each of these 30 life-history interview respondents were also administered a more extensive survey, covering the following modules: Demographic information, a household roster, labour market participation, household living standards, personal well-being and perceived mobility and economic shocks. Summary statistics from this survey are presented in Table 1.

Following the completion of fieldwork data collection, post-collection data analysis was undertaken by analysing transcriptions of FGDs and LHIs, and returning to audio recordings when necessary. Photographs of participatory life-course welfare trend graphs were especially important for gauging the magnitude of the impact of particular events in determining mobility patterns.

Mobility determinants are analysed separately in cases of upward versus downward mobility. A finding which emerged strongly from qualitative fieldwork was that the effect of particular drivers of mobility hinged crucially on how these interacted with other events, processes or contextual factors. This finding informs the development of a schema which was used to analyse economic mobility and welfare trajectories in LHIs. This schema approaches the analysis of any particular event by implicitly assuming a counterfactual world in which the event/process determining mobility does not occur. For example, a driver of downward mobility is understood as an event or process which results in downward mobility being greater than that which would have been expected if the event/process had not occurred. However, this event might occur in conjunction with another event/process/contextual factor, which might compound the negative effect on welfare. Equally, the effect of the same event may be offset by an event/process/contextual factor which

moderates the impact of the driver of downward mobility. For example, if a member of a household loses a job (a potential driver of downward mobility), but this event is linked to a decrease in household size, the moderating effect of this household event may cancel out the downward pressure of the labour market event. Fig. A2 demonstrates how this schema was applied to the analysis of events in LHIs by drawing on an illustrative example.

Figure 4.A.2: Illustration of the analysis of LHI events – Sicelo's life history.



Biographical summary: : Sicelo grew up in a village in the Eastern Cape in a large, poor household with a single parent. After failing to complete high school and struggling to find work in the Eastern Cape, he moved to Cape Town to find a job. He worked for a short time as a construction worker and as a security guard. He soon quit to attend a training course to become a petrol attendant. For the next six years he was formally employed at various petrol stations in Cape Town. During this time, he married a woman who was working as a cashier at a petrol station. Their combined incomes allowed them to escape poverty while at the same time supporting their two children. This poverty escape is represented by the movement into Box 2 in the graph above. This period of economic wellbeing and stability ended when Sicelo was fired from a string ofjobs and suffered a severe accident at the workplace. After several years, he once again found a position as a petrol attendant in Cape Town. However, in 2016 he was fired once again. The effect of this more recent job loss is represented in the red box in the figure above, and analysed below: Analysis of job loss event: Sicelo was fired from his job as a petrol attendant for unintentionally damaging a client's car. Cause: Poor performance due to stress caused by family disputes surrounding his grandfather's funeral, and marital strife concerning his infidelity. Immediate effect: Lost his wage job as a petrol attendant. Compounding factor(s): Sicelo's wife left him shortly after he lost the job. She left with their household possessions. Moderating factor(s): Meagre support from his girlfriend and from his cousin. Consequences: Severely compromised mental health. Material poverty. Social humiliation from unemployment and need to rely on a girlfriend to survive. Estrangement from children. Notes: Post-analysis representation by the author (right). Participant names are anonymised, as throughout the study.

This example captures how demographic and labour market events interact in complex ways to determine major shifts in life-course welfare trajectories. In this case, Sicelo's job loss seems to have been partly due to bad luck and partly the consequence of domestic and family issues which compromised his performance at work. Marital strife was both a cause and a consequence of the job loss, and ultimately compounded the effect that job loss had on Sicelo's wellbeing. At the same time, his relationship with another woman was also implicated in causing the job loss, but in this case also functioned as a way of coping with the impact of the event. The compound effect of this event has pulled Sicelo into a highly precarious economic situation. If Sicelo were to suffer another shock he would have no way of coping. Despite this, Sicelo may well move back out of poverty if he is able to find work. His past experience indicates that, while losing work is not infrequent, he is usually able to find work after a spell of unemployment. If this occurs, Sicelo may, once again, move out of poverty. This example illustrates how important it is to take these interactions into account when analysing events, and how this analysis was undertaken using the data collected in LHIs.

Finally, it is worth reflecting on the role that local fieldworkers played in both the collection and analysis of qualitative data in LHIs and FGDs. I worked with a group of between two and eight fieldworkers and interpreters, depending on the stage of data collection. I benefited from an especially productive relationship with one fieldworker who moderated most of the LHIs. He was able to pick up on many subtleties and undercurrents in interviews which I could not have noticed. These insights made post-interview debriefs, in which we reviewed transcriptions and analysed interviews at length, an essential step in the data-collection process.

References

[1] Robert Chambers and Gordon Conway. *Sustainable rural livelihoods: Practical concepts for the 21st century.* Brighton: Institute of Development Studies, 1992.

[2] Ivan. Turok. "Persistent polarisation post-apartheid? Progress towards urban integration in Cape Town". In: *Urban Studies* 38.13 (2001), pp. 2349–2377.

[3] Marianne Bertrand, Sendhil Mullainathan, and David Miller. "Public policy and extended families: Evidence from pensions in South Africa". In: *The World Bank Economic Review* 17.1 (2003), pp. 27–50.

[4] Alan Manning. "The real thin theory: Monopsony in modern labour markets". In: *Labour Economics* 10.2 (2003), pp. 105–131.

[5] J. Seekings and N. Nattrass. *Class, race and inequality in South Africa.* New Haven and London: Yale University Press, 2005.

[6] Dorrit Posel, James A Fairburn, and Frances Lund. "Labour migration and households: A reconsideration of the effects of the social pension on labour supply in South Africa". In: *Economic modelling* 23.5 (2006), pp. 836–853.

[7] Michelle Adato, Francie Lund, and Phakama Mhlongo. "Methodological innovations in research on the dynamics of poverty: A longitudinal study in KwaZulu-Natal, South Africa". In: *World Development* 35.2 (2007), pp. 247–263.

[8] Andries Du Toit and David Neves. "In search of South Africa's second economy: Chronic poverty, economic marginalisation and adverse incorporation in Mt. Frere and Khayelitsha". In: *Frere and Khayelit-*

sha (November 2007). Chronic Poverty Research Centre Working Paper 102 (2007).

[9] Andries Du Toit and David Neves. *Trading on a grant: Integrating formal and informal social protection in post-apartheid migrant networks*. Working Paper 75. Cape Town, South Africa: Programme for Land and Agrarian Studies, 2009.

[10] Stephan Klasen and Ingrid Woolard. "Surviving unemployment without state support: Unemployment and household formation in South Africa". In: *Journal of African Economies* 18.1 (2009), pp. 1–51.

[11] Jan. Theron. "Informalization from above, informalization from below: The options for organization". In: *African Studies Quarterly* 11.2/3 (2010), p. 87.

[12] Fiona. Tregenna. "How significant is intersectoral outsourcing of employment in South Africa?" In: *Industrial and Corporate Change* 19.5 (2010), pp. 1427–1457.

[13] Peter Davis. *Escaping poverty: Patterns and causes of poverty exits in rural Bangladesh Working Paper 194*. Tech. rep. Bath, U.K: Chronic Poverty Research Centre, 2011.

[14] Peter Davis and Bob Baulch. "Parallel realities: Exploring poverty dynamics using mixed methods in rural Bangladesh". In: *The Journal of Development Studies* 47.1 (2011), pp. 118–142.

[15] Katherine Eyal and Ingrid Woolard. "Female labour force participation and South Africa's child support grant". In: *CSAE 25th Anniversary Conference*. Centre for the Study of African Economies. Oxford, 2011.

[16] James Ferguson. "Declarations of dependence: Labour, personhood, and welfare in southern Africa". In: *Journal of the Royal Anthropological Institute* 19.2 (2013), pp. 223–242.

[17] Pierre Cahuc, Stéphane Carcillo, and André Zylberberg. *Labor economics*. Cambridge: MIT Press, 2014.

[18] James Ferguson. *Give a man a fish: Reflections on the new politics of distribution*. Durham and London, U.K.: Duke University Press, 2015.

[19] Arden Finn. *A national minimum wage in the context of the South African labour market*. Working Paper 2015/1. South Africa: National minimum wage research initiative at the University of the Witwatersrand, 2015.

[20] Stats SA. *Methodological report on rebasing of national poverty lines and development of pilot provincial poverty lines*. Tech. rep. Technical Report No. 03-10-11. Pretoria, 2015.

[21] Haroon Bhorat, Aalia Cassim, and Derek Yu. *Temporary employment services in South Africa: Assessing the industry's economic contribution LMIP Report 28*. Tech. rep. Pretoria: Labour Marker Intelligence Partnership, 2016.

[22] Jessica Budlender. *Edged Out: Spatial Mismatch and Spatial Justice in South Africa's Main Urban Areas – Technical Report*. Tech. rep. Johannesburg: Socio-Economic Rights Institute of South Africa, 2016.

[23] Bongani Mncwango. *Public attitudes to work in South Africa LMIP Report 16*. Pretoria: Labour Marker Intelligence Partnership. 2016.

[24] Matthew de la Hey and William Beinart. "Why have South African smallholders largely abandoned arable production in fields? A case study". In: *Journal of Southern African Studies* 43.4 (2017), pp. 753–770.

[25] Andrew Kerr. "Tax (i) ing the Poor? Commuting costs in South African cities". In: *South African Journal of Economics* 85.3 (2017), pp. 321–340.

[26] Stats SA. *Poverty Trends in South Africa: An examination of absolute poverty between 2006 and 2015*. Tech. rep. Report no. 03-10-06. Pretoria, 2017. URL: https://www.statssa.gov.za/publications/Report-03-10-06/Report-03-10-062015.pdf.

[27] Ivan Turok, Josh Budlender, and Justin Visagie. "The role of informal urban settlements in upward mobility". In: (2017).

[28] Christopher Blattman and Stefan Dercon. "The impacts of industrial and entrepreneurial work on income and health: Experimental evidence from Ethiopia". In: *American Economic Journal: Applied Economics* 10.3 (2018), pp. 1–38.

[29] Hope Dawson. "The productivity of unemployment in urban South Africa". Paper presented at the African Studies Centre's Southern African Discussion Group, Oxford, UK. 2018.

[30] James Ferguson and Tania Murray Li. *Beyond the "Proper Job:" Political-economic analysis after the century of labouring man*. Working Paper 51. Cape Town: PLAAS, UWC, 2018.

[31] Michael Rogan and Caroline Skinner. "The size and structure of the South African informal sector 2008–2014: A labour-force analysis". In: *The South African informal sector: Providing jobs, reducing poverty*. Ed. by Francois Fourie. Pretoria: Human Sciences Research Council (HSRC) Press, 2018, pp. 77–102.

[32] SALDRU. *National Income Dynamics Study 2008, Wave 1. Version 5.3*. Cape Town: SALDRU [producer], Cape Town: DataFirst [distributor]. 2018.

[33] SALDRU. *National Income Dynamics Study 2010-2011, Wave 2. Version 2.3*. Cape Town: SALDRU [producer], Cape Town: DataFirst [distributor]. 2018.

[34] SALDRU. *National Income Dynamics Study 2012, Wave 3. Version 1.3*. Cape Town: SALDRU [producer], Cape Town: DataFirst [distributor]. 2018.

[35] SALDRU. *National Income Dynamics Study 2014-2015, Wave 4. Version 1.0*. Cape Town: SALDRU [producer], Cape Town: DataFirst [distributor], Pretoria: Department of Planning Monitoring and Evaluation [commissioner]. 2018.

[36] SALDRU. *National Income Dynamics Study 2017, Wave 5. Version 1.0.0*. Pretoria: Department of Planning, Monitoring, and Evaluation [funding agency]. Cape Town: Southern Africa Labour and Development Research Unit [implementer], Cape Town: DataFirst [distributor]. 2018.

[37] South African Social Security Agency SASSA. *Annual Report, 2017/2018*. Pretoria, 2018.

[38] Francis Teal. "What constrains the demand for labour in firms in Sub-Saharan Africa? Some evidence from Ghana". In: *CSAE Conference*. Centre for the Study of African Economies. Oxford, 2018.

[39] Martin Abel. "Unintended labor supply effects of cash transfer programs: New evidence from South Africa's Pension". In: *Journal of African Economies* 28.5 (2019), pp. 558–581.

5 | Conclusion

This thesis is made up of three stand-alone papers. While contextually distinct, each paper is organised around an attempt to understand barriers that individuals in developing countries face in making productive use of their labour.

South Africa's labour market is characterised not only by high rates of unemployment, but also very high levels of labour market turnover (Kerr, 2018). Especially at the bottom of the earnings distribution, jobs are insecure and often of limited duration, with the result that employment-to-unemployment transitions are common (Schotte et al., 2018). Inadequate search while on-the-job may represent a consequential missed opportunity, since evidence from various contexts suggests that on-the-job search is substantially more effective than job search while unemployed (Blau and Robins, 1990; Faberman et al., 2017). Workers who move from employment back into unemployment often struggle to find work in a slack labour market, and lengthy spells of unemployment may lead to hysteresis (Rosholm and Toomet, 2005; Eriksson and Rooth, 2014; Mueller et al., 2021).

Paper 1 explores whether search behaviour while on-the-job may explain some of the high turnover we observe. Specifically, it is an investigation into the role that rejection during job search has on individuals' search behaviour. If there are forces that create disincentives to job-search, these may represent important barriers to labour market entry for certain workers. Using a lab-in-the-field experiment implemented with young workers in South Africa, I repeatedly ask subjects to choose between a high-return activity with frequent exposure to rejection signals and a lower-return activity with

less frequent exposure to rejection signals. I experimentally vary both the rewards and the amount of rejection that subjects experience when choosing between the two tasks, holding other salient drivers of search behaviour constant. I find that when exposed to (more frequent) rejection, subjects choose to trade-off expected earnings in order to avoid receiving rejection signals. I interpret these results as an example of active information avoidance. If this behaviour carries over into the real labour market, this would imply that frequent rejection has the effect of depressing search rates, and that workers implicitly trade-off earnings and employment opportunities to minimise exposure to rejection.

While I observe that the experience of rejection-qua-rejection reduces overall search rates in a laboratory setting, I cannot make any certain claims about the external validity of this finding for behaviour in the real labour market. The paper does, however, provide some support for the hypothesis that rejection may itself depress search rates, which at least raises the possibility that inadequate search while on-the-job may contribute to high observed employment-to-unemployment transitions. Some estimates suggest that job-finding rates in the real labour market are significantly lower than the success rates I set in the game (Garlick et al., 2023), and qualitative evidence from my own fieldwork also suggests that workers do not search as much as they 'should' if they were only acting to maximise their chances of landing a job.

Finding a job is not the only barrier to meaningful labour market integration in South Africa. In a world in which workers and firms have imperfect information about each other (Carranza et al., 2022), and in which monopsony concentrates market power in the hands of firms (Bassier, 2021; Bassier, 2023), the terms on which employment is offered is not always as favourable

as workers hope and expect. Paper 3 combines original qualitative data and nationally representative survey data to explore those paradoxical cases in which black, urban youths turn down or quit wage work. I use this paradox to explore the disincentives to certain forms of low-skill employment that South African workers face, and what this reveals about the labour market more broadly. I find that young male workers in particular have relatively strong outside options to the (often poor quality) wage work they are able to access, and so their attachment to the wage-labour is often weak. Like Paper 1, which suggests that psychological disincentives to job search while on-the-job may help explain high labour turnover, Paper 3 also contributes to our understanding of high turnover and short tenure in certain sectors of the South African labour market.

The root cause of unemployment and labour market precariousness in South Africa is no doubt structural – with inadequate demand for the kind of labour which is in most abundant supply. Papers 1 and 3 of this thesis deal with how the overall slackness of the labour market affects how workers engage with the labour market: focusing on psychological barriers to searching for work and psychological and other disincentives to keeping jobs. To the extent that these phenomena lead to low worker motivation, high turnover, sub-optimal worker-firm matches, and under-investment in on-the-job training, they may in turn conspire to reduce demand for labour, contributing to a vicious cycle. The bidirectional link between the experience of the labour market on the supply side, and the low demand for labour, is for now highly speculative and will need further work to understand.

South Africa is a particular case – where the correlation between 'good' jobs and the formal sector and 'bad' jobs and the informal sector remains somewhat tight. In most developing countries, the barriers to meaningful

work are not about moving from the informal to the formal sector, but about moving from survivalist and precarious forms of self-employment to forms of informal self-employment which offer more meaning, economic security and opportunities for wealth accumulation. Bangladesh – the context of Paper 2 – is an example.

It is widely accepted that one of the barriers to upward occupational mobility in this context is determined by access to productive capital. Evaluations of graduation models – in which ultra-poor households receive large productive assets to break a poverty trap – have been important in establishing this consensus. However, these interventions consist not only of capital inputs, but of a large bundle of other inputs – including intensive, regular, one-on-one coaching. The question which emerges, then, is whether the only binding constraint to upward occupational mobility is productive capital, or whether alleviating human capital constraints also plays a fundamental role. This question is of considerable policy interest: regular coaching – the means through which human capital transfers are intended to occur – are expensive and difficult to scale. If the size of programme effects can be maintained at lower cost, by reducing the amount of coaching recipients receive, this has the potential to significantly increase impact on the extensive margin.

Paper 2 studies this question through an experimental unbundling of a 'big push intervention' in rural Bangladesh, using a large cluster randomised control trial, I find that coaching variation in coaching plays no discernible role in explaining the considerable positive welfare effects of the programme. This is good news for scale – suggesting that there is scope for improvements on the extensive margin without compromising effects on the intensive margin.

In the paper I also seek to say something about the mechanisms through which coaching does work – for instance, whether coaching works by transferring hard skills. The project was not designed to be able to identify causal effects on these mechanisms, since there is no treatment arm that received no coaching but which did receive other elements of the programme. However, by exploiting variation I observe between treatment groups and between treatment and control, I am able to tentatively rule out several mechanisms which might be driving the impact of coaching: I find no evidence that coaching affects hard skills, and weak and inconsistent evidence on the effects of coaching on a range other other possible intermediate outcomes – psychosocial mechanisms, social inclusion and financial inclusion. Rather, I suggest two other potential mechanisms that may be driving the impact of coaching, but which our study cannot directly speak to: If coaching is effective, it may be through 'supervisory' and 'intermediary' channels. That is, by expecting regular visits, programme recipients comply with the expectations of the programme – for instance, by following best practices in caring for their asset and by resisting the temptation to liquidate the asset prematurely. In addition, coaches may act as intermediaries to help individuals access services they may otherwise have struggled to access or may not have known about. These hypotheses remain somewhat speculative, and may be a fruitful avenue for further research.

References

[1] David M Blau and Philip K Robins. "Job search outcomes for the employed and unemployed". In: *Journal of political Economy* 98.3 (1990), pp. 637–655.

[2] Michael Rosholm and Ott Toomet. "A search model of discouragement". In: *Available at SSRN 742331* (2005).

[3] Stefan Eriksson and Dan-Olof Rooth. "Do employers use unemployment as a sorting criterion when hiring? Evidence from a field experiment". In: *American economic review* 104.3 (2014), pp. 1014–39.

[4] Jason Faberman, Andreas Mueller, Ayşegül Şahin, and Giorgio Topa. *Job Search Behavior among the Employed and Non-Employed*. Working Paper 23731. National Bureau of Economic Research, Aug. 2017. DOI: 10.3386/w23731.

[5] Andrew Kerr. "Job flows, worker flows and churning in South Africa". In: *South African Journal of Economics* 86 (2018), pp. 141–166.

[6] Simone Schotte, Rocco Zizzamia, and Murray Leibbrandt. "A poverty dynamics approach to social stratification: The South African case". In: *World Development* 110 (2018), pp. 88–103.

[7] Ihsaan Bassier. *The impact of centralized bargaining on spillovers and the wage structure in monopsonistic labour markets*. Tech. rep. WIDER Working Paper, 2021.

[8] Andreas Mueller, Johannes Spinnewijn, and Giorgio Topa. "Job seekers' perceptions and employment prospects: Heterogeneity, duration dependence, and bias". In: *American Economic Review* 111.1 (2021), pp. 324–63.

[9] Eliana Carranza, Robert Garlick, Kate Orkin, and Neil Rankin. "Job Search and Hiring with Limited Information about Workseekers' Skills". In: *American Economic Review* 112.11 (2022), pp. 3547–83.

[10] Ihsaan Bassier. "Firms and inequality when unemployment is high". In: *Journal of Development Economics* 161 (2023), p. 103029.

[11] Robert Garlick, Lukas Hensel, Andrea Kiss, and Kate Orkin. "Jobseekers' Beliefs about Comparative Advantage and (Mis)Directed Search". Unpublished manuscript. Mar. 2023.