

Microentrepreneurship and Market Access in Development

A dissertation submitted by

Carolyn E. Pelnik

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Economics and Public Policy

Tufts University

May 2025

© 2025, Carolyn E. Pelnik

Adviser: Cynthia Kinnan

Abstract

This dissertation examines entrepreneurship, financial frictions, and market access in low-income countries. In the first paper, I study the spatial misallocation of microentrepreneurs within a large city. I show that observationally similar business owners earn vastly different profits across city locations within Kampala, Uganda. When entrepreneurs receive a subsidy and information together, they move locations and increase their own profits by 45%. But impacts fade after subsidies cease, consistent with risk-averse, utility-maximizing entrepreneurs facing substantially greater profit variation in locations with higher average profits. Using data from the experiment to estimate a model where entrepreneurs choose locations and realize profit in equilibrium, I show that aggregate income increases by 37% when financial and information constraints are relieved at scale. The results demonstrate spatial misallocation of entrepreneurs within the city and simultaneous financial and information frictions that inhibit the mobility of entrepreneurs.

In the second paper (with Joe Kaboski, Molly Lipscomb, and Virgiliu Midrigan), I study the extent to which income growth depends on financial services access in the presence of indivisible investment opportunities. We offered peri-urban Ugandans a choice between a safer, lower payoff and a riskier, larger payoff lottery; 27% of participants choose the riskier lottery, with winners investing more in land and durable business assets, eventually increasing income. In contrast, winning the small lottery has only transitory impacts on business inventory and livestock. Our quantitative model shows that the aggregate effects of financial deepening are sizable if the indivisible investment can be accumulated but not if it is in fixed supply.

In the third paper (with Viva Bartkus, Wyatt Brooks, and Joe Kaboski), I show how reliance on monopsonistic middlemen reduces income in remote communities where financial frictions inhibit investment in transportation infrastructure. We study a cooperative intervention which organizes together fishing communities in the Amazon to purchase large boats and deliver their fish directly to market. We find that the intervention increases income by 27%, largely through an increase in price received. We formalize a model in which the market power of middlemen itself can create a poverty trap, which can be eliminated with cooperative investment.

Acknowledgments

I am deeply grateful to my advisors, Cynthia Kinnan, Kyle Emerick, Adam Storeygard, and Doug Gollin, who made my work better and imparted skills that I will carry with me for my entire career, all while being some of the kindest people I know. I was constantly blown away by their generosity of time and advice. I am indebted to Cynthia's thoughtfulness, unparalleled knowledge of methods and the literature, and ability to identify and diagnose problems in the gentlest way possible. Kyle's willingness to troubleshoot issues in the field whenever they arose both greatly improved my work and saved me immeasurable stress. None of this would have been possible without Maggie McMillan, who started the Economics and Public Policy (EPP) PhD program at Tufts and whose vision has led to an incredibly vibrant and supportive place to do research in development economics.

Thank you to my parents, Cathy and Tom, whose support has always been a constant and remained unwavering throughout their heroic battle to beat my dad's cancer diagnosis this past year. Thank you to my partner, Brennan, brother, Eric, and sister-in-law, Nica, for their support, levity, and perspective through it all.

I am grateful for the wonderful friends who I had the chance to meet during and because of graduate school: Ernesto Tiburcio, Mai Mahmoud, and Billy Babis; my adopted cohort, Leticia Donoso Peña, Manali Sovani, Sarah Shaukat, and Anne Krahn; my office (and occasional house) mates, Josué Awonon, Siddhi Doshi, Aja Kennedy, and Axel Eizmendi Larrinaga; Connie Chamberlain, who made Cambridge home; and last but certainly not least, my housemates during fieldwork, Tamara Kerzhner, Rebecca Wai, and Yusra Abdullahi, who know the ups and downs of the first paper in this dissertation better than anyone. I wish everyone friends who will perform minor surgery on your toe, sludge through mud (and altitude sickness) for 8 days while hiking to the highest peak in Uganda, and drop everything to meet for curry and gin at the end of a tough day.

In Uganda and Colombia, Victoria Twinomuhangi and Claudia Paola Salcedo Vasquez were the best implementing partners, turned friends, imaginable. Vicky's ingenuity and dedication made possible massive data collection efforts in Uganda. Claudia Paola's unflappable determination to improve livelihoods in Colombia made possible an ambitious study of market access in rural communities. Spending time with them both was a highlight of my years as a graduate student.

I am also grateful for the lifelong friends whose love has been one of the greatest joys of my life since before the PhD. Lacey, thank you for the travel and research adventures that started in college (and friendship since then) and for nudging us towards Tanzania in 2011, which in many ways kickstarted this whole journey.

Katie, thank you for always being game for a ridiculous weekend (or week) in DC or New York or Marrakech or Lisbon or Cape Town or Berlin. Saha, thank you for remembering every major life event no matter how geographically far apart we may be. Blair, thank you for being my original fieldwork housemate, engaging in all the things that you (mostly) only do when you're 25 ("Oh, it's a 5-day hike? Obviously we can do it in 2"), and for continuing to be a dear friend in Boston while we both navigate our grad school and post-grad school lives.

Thank you to coauthors who are also mentors and first brought me into the world of field experiments (and development economics more broadly). Molly Lipscomb is the reason I am an economist. Somehow she thought a meandering engineering undergrad would be a reasonable economics TA and RA. Joe Kaboski taught me about the complementarities between field experiments and macro models and together with Molly gave me the chance to learn about running field studies in the best way possible – by spending tons of time in the field. Wyatt Brooks was an everwilling and overly generous sounding board for my job market paper. Viva Bartkus had the vision to bring development economists to the table with practitioners to solve real-world business problems, and it led to some of my most exciting collaborations. I am grateful to all for such insanely generous investment in and encouragement of my career.

Finally, thanks to Rory and Chester for the walks and goofiness that only a dog and cat can provide. My dissertation would have suffered without you.

Contents

Abstract	ii
Acknowledgments	iv
List of Tables	xi
List of Figures	xii
1 Moving to Profitability? Alleviating Constraints on Microentrepreneur Location	1
1.1 Introduction	1
1.2 Baseline Setting and Motivating Facts	6
1.2.1 Sampling Procedure	6
1.2.2 Sample Characteristics	7
1.3 Model	8
1.3.1 Entrepreneur’s Preferences	9
1.3.2 Determination of Profits in General Equilibrium	9
1.3.3 Model Predictions	10
1.4 Experimental Design and Data Collection	11
1.4.1 Experimental Sample and Randomized Treatment Assignment	11
1.4.2 Implementation of the Interventions	13
1.4.3 Data Collection	14
1.5 Empirical Analysis and Results	17
1.5.1 Take-Up and Description of Moves	17
1.5.2 Estimation Strategy	18
1.5.3 Effect on Business Location During the Intervention	21
1.5.4 Effect on Financial Outcomes During the Intervention	24

1.5.5	Persistence: Effect of Treatment After the Intervention Ends	26
1.6	Mechanisms	27
1.6.1	Moving During the Intervention: Profitable or Preferable?	27
1.6.2	Role of Information: Inducing High-Return Entrepreneurs to Move	29
1.6.3	Role of Liquidity: Insurance for Risky Moves	30
1.7	Model Estimation and Counterfactual Simulations	30
1.7.1	Assigned Parameters	31
1.7.2	Estimated Parameters	32
1.7.3	Model Simulations	33
1.8	Conclusion	36
1.9	Figures	38
1.10	Tables	47
	Appendices	58
1.A	Derivation of Profit Response to Firm Entry in Equilibrium	58
1.B	Data Appendix	61
1.B.1	The Conditional Cash Transfer	61
1.B.2	Estimating Treatment-on-Treated Effects	62
1.C	Mechanisms Appendix	66
2	How Important are Investment Indivisibilities for Development? Experimental Evi-	
	dence from Uganda	67
2.1	Introduction	67
2.2	Model and Motivation	71
2.2.1	Set Up	71
2.2.2	Model Results: Poverty Traps and Risk Loving Behavior	73
2.3	Field Experiment	78
2.3.1	Experimental Design	78
2.3.2	Data	81
2.4	Who selects the large lottery?	85
2.4.1	Methods of Predicting Lottery Choice	85
2.4.2	Correlates of Lottery Choice	86
2.5	How are winnings spent and what are the household impacts?	88
2.5.1	How Lottery Winnings are Spent: Impacts on Investment	89

2.5.2	Impact on the Full Household Budget	92
2.5.3	Understanding the Large Effect on Land Values	96
2.6	Calibrated Model Results	98
2.6.1	Calibration	98
2.6.2	Counterfactual Assessment of Introduction of Financial Services	100
2.7	Conclusions	103
2.8	Figures	105
2.9	Tables	108
Appendices		122
2.A	Data Appendix	122
2.A.1	Consumption	122
2.A.2	Home Durables	123
2.A.3	Savings	123
2.A.4	Income	123
2.A.5	Agricultural Assets	124
2.A.6	Business Assets	125
2.A.7	Land	125
2.A.8	Net Credit	126
2.B	Evidence of GE Effects for Land	127
2.C	Predicting Lottery Choice	131
2.D	Attrition	135
2.D.1	Balance Between Retained and Attrited Sample	135
2.D.2	Lee Bounds	136
2.E	Multiple Hypothesis Corrections	140
2.F	Description of Grant Purchases	143
2.G	Model Validation Regressions	145
3 Big Fish in Thin Markets: Competing with the Middlemen to Increase Market Access in the Amazon		149
3.1	Introduction	149
3.2	Background on the Brazil Amazon, Supply Chain, and Experiment	153
3.2.1	Setting	153
3.2.2	Existing Supply Chain	154

3.2.3	Intervention	156
3.3	Empirical Analysis	158
3.3.1	Data	158
3.3.2	Balance	160
3.3.3	Impact on Households	161
3.3.4	Impact on the Pirarucu Market	163
3.3.5	Cost-Benefit Analysis	165
3.4	Model	166
3.4.1	Fishing Communities	166
3.4.2	Wealth Dynamics	168
3.4.3	Middlemen	170
3.4.4	Interventions to Escape Poverty Trap	171
3.4.5	Is there a poverty trap?	173
3.5	Conclusion	174
3.6	Figures	176
3.7	Tables	178
Appendices		182
3.A	Main Appendix	182
3.B	Data Appendix	186

List of Tables

1.1	Baseline Sample Characteristics	47
1.2	Experimental Design	48
1.3	Balance Table	49
1.4	Lasso-selected characteristics of compliers	50
1.5	Effect on Location During the Intervention	51
1.6	Effect on Financial Outcomes During the Intervention	52
1.7	Effect on Location and Financial Outcomes After the Intervention	53
1.8	Heterogeneity in Treatment Effects on Profit During the Intervention: By Likelihood of Com- plying with the CCT+Info	54
1.9	Heterogeneity in Treatment Effects on Profit During the Intervention: By Likelihood of Com- plying with the CCT Only	55
1.10	How Pass Through of Profit to Consumption Varies by Treatment Group (Panel)	56
1.11	Simulation Results	57
1.B.1	Experimental Sample: Composition along Stratification Variables	63
1.B.2	Attrition by Treatment Arm	63
1.B.3	Take-Up of the Intervention	64
1.B.4	Baseline Characteristics of Compliers	65
2.1	Lottery choices	108
2.2	Lottery winners	108
2.3	Balance between small lottery winners and non-winners	109
2.4	Balance between large lottery winners and non-winners	110
2.5	Statistically significant differences between those choosing the large v. small lottery	111
2.6	Grant effects on investment - First endline	112
2.7	Grant effects on investment - Second endline	113

2.8	Grant effects on components of the household budget constraint - First endline	114
2.9	SUR with household budget constraint - First endline	115
2.10	Grant effects on components of the household budget constraint - Second endline	116
2.11	SUR with household budget constraint - Second endline	117
2.12	Grant effects on components of the household budget constraint - Third endline	118
2.13	Effect of winning the grant on likelihood of purchasing, improving, and selling land	119
2.14	Calibration Parameters and Moments	120
2.15	Counterfactual Aggregate Impacts of Financial Interventions	121
2.A.1	Estimated capital gains by region	125
2.B.1	Effect of grants disbursed nearby on others' land values nearby (first stage)	128
2.B.2	Effect of others' land values nearby on own land value (second stage)	129
2.C.1	Characteristics of those choosing the large v. small lottery	131
2.C.2	Predicting those who chose the large lottery: Lasso-selected covariates	132
2.C.3	Lasso-selected predictors of midline lottery choice and baseline hypotheticals	133
2.D.1	Balance between retained and attrited, among those choosing the small lottery	135
2.D.2	Balance between retained and attrited, among those choosing the large lottery	136
2.D.3	Grant effects on components of the household budget constraint - First endline - Lee Bounds	137
2.D.4	Grant effects on components of the household budget constraint - Second endline - Lee Bounds	138
2.D.5	Grant effects on components of the household budget constraint - Third endline - Lee Bounds	139
2.E.1	Grant effects on components of the household budget constraint - First endline - Multiple hypothesis corrections	140
2.E.2	Grant effects on components of the household budget constraint - Second endline - Multiple hypothesis corrections	141
2.E.3	Grant effects on components of the household budget constraint - Third endline - Multiple hypothesis corrections	142
2.F.1	Lottery winner grant uses	143
2.G.1	Model-Simulated Monte Carlo Regressions	148
3.1	Balance on Baseline Community Characteristics	178
3.2	Effects on Households	179
3.3	Effects on Market Pricing and Channel (Community Level)	180
3.4	Effects on Quantity Sold and Revenue Earned by Communities	181
3.A.1	Sailing Boats from Mamirauá to Market Variable Transportation Costs	182

3.A.2	Summary Statistics for the 2018 Household Survey	183
3.A.3	Testing for Differential Pre-Trends in Fishing Levels 2013-2017, Mamirauá	184
3.A.4	Effects on Household Time Allocation	185

List of Figures

1.1	Between-Parish Profit Variation for Observationally Similar Microenterprises	38
1.2	Parish Profits at Baseline	39
1.3	Example of Information Treatment Allocation	40
1.4	Distribution Regressions: Effect of Treatment on Probability of Locating >X Kilometers from Pre-Intervention Business Location	41
1.5	Profit Distributions During the Intervention	42
1.6	Profit Distributions: Moving to a Higher Profit Parish Versus Not	43
1.7	Distribution of Parish Amenities	44
1.8	Model Fit	45
1.9	Correlations Between Estimated and Assigned Parameters	46
1.B.1	Distance to More Profitable Parishes	61
1.C.1	Distribution Regressions: Effect of Treatment on Probability of Locating >X Kilometers from Pre-Intervention Business Location, Among Likely CCT+Info Compliers	66
2.1	Asset-Dependent Behavior and Poverty Traps	105
2.2	Asset-Dependence of Lottery Choice	106
2.3	Distribution Regressions: Effect of Winning the Lottery on the Distributions of Income at Each Endline	107
3.1	Pirarucu Supply Chain and Market (Reais/kg) in Mamiraua, Amazonas, 2018	176
3.2	Wealth Dynamics Under Monopsonistic Middlemen Pricing	177

Chapter 1

Moving to Profitability? Alleviating Constraints on Microentrepreneur Location

Carolyn Pelnik

1.1 Introduction

The location of economic activity is associated with its value. In low- and middle-income countries, workers in urban areas earn more than three times those in rural areas (Gollin, Lagakos, and Waugh, 2014). In the United States, average earnings vary considerably between commuting zones (Card, Rothstein, and Yi, 2023). These spatial income gaps may reflect market frictions that inhibit a worker’s ability to move between regions. In these cases, relieving constraints on mobility, such as by paying moving costs or providing information about income prospects in other regions, increases workers’ earnings (Bryan, Chowdhury, and Mobarak, 2014; Baseler, 2023). Much less is known about spatial variation in earnings *within* cities, especially among the self-employed.

In this paper, I quantify the extent of spatial misallocation of microentrepreneurs within a large city. First, I use original field surveys to show descriptively that similar microentrepreneurs earn vastly different profits across city locations. Second, I use a field experiment (n=2,883) to empirically test whether spatial profit gaps reflect misallocation. I randomly allocate a moving subsidy, cross-randomized with an information

intervention, and estimate the effect of moving on microentrepreneurs' profits.¹ On average, entrepreneurs increase profits by 45% as the result of moving, but only when they receive both the subsidy and information. Third, I rationalize these results in a structural model where entrepreneurs choose locations and realize profit in equilibrium. Using data from the experiment to estimate the model, I find that aggregate income increases by 37% when financial and information frictions are relieved at scale, net of changes in profit due to firm entry and exit from locations across the city. Relieving financial or information constraints alone is not sufficient, indicating the presence of simultaneous market failures that lead to spatial misallocation.²

I focus on seemingly mobile microentrepreneurs—street vendors and motorcycle taxi drivers—in Kampala, the capital of Uganda and home to approximately 2 million people.³ I document that microenterprise profits associated with the 90th percentile most profitable city parish are three times those of the 10th percentile parish, even after controlling for individual entrepreneur characteristics. Absent market frictions, we expect microentrepreneurs to move from lower to higher profit places and compete down differences in profit between locations. Thus, if spatial profit gaps reflect market frictions that inhibit moving between business locations, i.e., misallocation, then entrepreneurs should realize higher profits by relocating. If, on the other hand, profit gaps reflect sorting, i.e., selection, then moving will not increase profits (e.g., Roy, 1951; Borjas, 1987). If there are large differences in nonmonetary amenities between locations and entrepreneurs trade off profit against amenity value, then moving may not increase welfare (e.g., Rosen, 1979; Roback, 1982). The cause of spatial variation in income therefore informs which policies will increase business income and entrepreneur welfare.

The experiment cross-randomizes an information intervention with two liquidity interventions, testing directly whether spatial profit gaps reflect misallocation and whether misallocation results from information and financial frictions. The information intervention identifies locations in the city that have higher average profits within an entrepreneur's sector. The first liquidity intervention is a moving subsidy, which pays microentrepreneurs who relocate their businesses within the city, and the second is an unconditional transfer of equivalent value to the subsidy.⁴ The moving subsidy induces exogenous variation in business location,

¹I use microentrepreneur profits synonymously with net earnings, or the entrepreneur's revenue less costs, *not* accounting for the value of their own time. This treatment of profits is standard in the development literature and reflects the well-documented lack of separation between production and consumption decisions in owner-occupied firms, i.e., where the owner is the only employee and daily business earnings are the primary source of household income.

²Here, misallocation refers to *profit* misallocation. Below, I use the model to quantify profit and utility misallocation.

³Microentrepreneurs in this sample operate with mobile capital, such as motorcycles and wheelbarrows, instead of from fixed structures, and they operate informally, without paying rent, taxes or licensing fees. While this setting differs from the typical business environment in high-income countries, the entrepreneurs in this study operate businesses that represent a frequent source of income for urban Ugandans and those living in low-income countries more broadly. Uganda's National Labor Force Survey indicates that 35% of Kampala's labor force is self-employed with no employees other than the owner, and more than 95% of these businesses operate informally without registering with government authorities (Uganda Bureau of Statistics, 2021), like the businesses in this study.

⁴Specifically, the moving subsidy pays \$2 per day to entrepreneurs who move their businesses three or more kilometers from their pre-intervention business location.

testing whether entrepreneurs can increase profits by moving business locations. In doing so, it identifies the extent to which there is misallocation (i.e., location-specific profit premia have not been competed away) as opposed to selection by entrepreneurs into particular places (i.e., individual entrepreneur characteristics drive profits).⁵ The unconditional cash transfer relieves liquidity constraints in general. If entrepreneurs prefer to allocate the marginal dollar towards moving but face liquidity constraints that prevent them from doing so (due to, for example, incomplete credit, savings, or insurance markets), then responses to the subsidy and unconditional cash transfer should be the same. Alternatively, if entrepreneurs prefer not to allocate funds towards moving, but this preference is of lower value than the subsidy, then the subsidy alone will induce entrepreneurs to move business locations.⁶ Cross-randomizing both liquidity interventions with the information intervention allows me to test whether liquidity, information, or both restrict firm location decisions.

I find that only entrepreneurs who take up the moving subsidy when it is provided with information are more likely to move into higher profit parishes. Those taking up the moving subsidy and also receiving information are 37% more likely than the control group to move into parishes with higher average profits.⁷ They also realize statistically significant revenue and income gains during the intervention. Their revenue is 26% higher and profit is 45% higher than the control group, net of the moving subsidy itself. The large profit effect is driven by increases in revenue without an increase in costs, underscoring the nature of the spatial arbitrage opportunity: entrepreneurs taking up the subsidy with information increase profit solely by moving, without expanding their businesses on any other dimension. Those receiving information alone, the moving subsidy alone, the unconditional cash alone, or the unconditional cash with information experience no statistically significant effect on the likelihood of moving into a higher profit parish, on revenue, or on profit.

Information serves to induce a high-return subset of entrepreneurs to take up the moving subsidy. Those taking up the moving subsidy when it is offered with information differ on baseline characteristics from those taking up the moving subsidy when it is offered without information. They have higher baseline knowledge of profits throughout the city, longer commutes, and locate their businesses in parishes with lower average profits than those taking up the subsidy without information. The subgroup of entrepreneurs with these baseline characteristics realizes statistically significantly higher profit when offered either the moving

⁵A hybrid case might involve a location-entrepreneur match, where market frictions inhibit entrepreneurs from accessing their individual highest profit location, leaving untapped location-entrepreneur profit premia. In this case, I still expect a profit effect of relieving frictions that restrict entrepreneurs' access to particular locations. Absent market frictions, idiosyncratic location-entrepreneur matches would show up as sorting, where there are common entrepreneur characteristics that lead to sorting by particular types of entrepreneurs into particular places.

⁶The subsidy will induce moves by those for whom it exceeds the net pecuniary and nonpecuniary costs of moving.

⁷Here, I provide treatment-on-treated (ToT) effects, which reflect the effect of moving business locations for those taking up the subsidy. The corresponding intent-to-treat (ITT) effects are below.

subsidy with information or the unconditional cash with information. I find no effect—of the subsidy or the unconditional transfer, with or without information—on profit for the group with baseline characteristics resembling those who are likely to take up the moving subsidy without information. This suggests that information serves a self-selection function by inducing a group that is able to realize differentially higher returns of moving to in fact take up the moving subsidy, in contrast to when the moving subsidy is offered without information. When offered with information, the unconditional transfer serves in a similar capacity as the subsidy for this subgroup of high-return movers.

The role of the liquidity interventions is consistent with an insurance-like mechanism. In an omnibus test for consumption smoothing ([Townsend, 1994](#)), I find that receipt of either the subsidy or the cash transfer reduces the correlation between the entrepreneur’s income and consumption. Moreover, entrepreneurs who locate in places with higher average profits also experience much greater variance in profits. Thus, entrepreneurs may desire insurance to locate in places that have higher average profits but are also risky. Consistent with a desire for insurance, the moving, revenue and profit effects realized by those receiving liquidity and information fade throughout the month after the intervention. Given the much larger variance associated with high-profit places, a plausible degree of risk aversion—a coefficient of relative risk aversion of 1.16—can rationalize a decision to move while the liquidity interventions pay out (and effectively function as a consumption floor) but not after they cease. Incomplete insurance, thus, may explain both the lack of persistence and the role of liquidity in inducing entrepreneurs to move business locations.

I rationalize the empirical results in a structural model where entrepreneurs choose locations and realize profit in equilibrium. I adapt a spatial equilibrium model (e.g., [Ahlfeldt et al., 2015](#); [Monte, Redding, and Rossi-Hansberg, 2018](#)) to the business location decisions of self-employed microentrepreneurs who maximize utility, not profit, reflecting a lack of separation between consumption and production decisions (e.g., [Jayachandran, 2006](#); [LaFave and Thomas, 2016](#)). Entrepreneurs realize profit in equilibrium, as new entrants compete down profits. I use survey data and exogenous variation from the experiment to estimate the model. I find that a model with fixed costs of moving, limited knowledge of profits across the city, and heterogeneous returns to moving can replicate the empirical results.

I use the model to simulate the aggregate impacts of permanently alleviating information and financial constraints on entrepreneurs. The model simulations show that aggregate income increases by 37% when information and liquidity constraints are relieved in general equilibrium, *net* of changes in the profitability of a place as new firms enter and compete down profits. Aggregate welfare gains, while positive, are not as large as aggregate income gains, reflecting a steep tradeoff between the monetary and non-monetary aspects of an entrepreneur’s objective. A relatively small set of high-profit, high-amenity places become congested as entrepreneurs enter and compete down profits. As a result, entrepreneurs move into less congested places

that have higher profits but lower amenities. The large income gains in equilibrium demonstrate sizable underlying frictions that constrain firm location decisions and suppress aggregate income.

I contribute evidence on spatial misallocation of entrepreneurs to the literature on spatial frictions in low-income countries. Prior work finds large income gains from the expansion of transportation infrastructure (Donaldson, 2018; Balboni et al., 2020; Brooks and Donovan, 2020) and from encouraging rural-urban migration (Bryan, Chowdhury, and Mobarak, 2014; Bryan and Morten, 2019). Prior work has not considered the role of the within-city mobility of entrepreneurs, as opposed to workers, despite the prevalence of self-employment in low-income countries. Other work on within-city firm location shows that consumer search frictions can explain the spatial clustering of tailoring firms (Vitali, 2024) and that higher firm profits compensate for greater pollution exposure in the city (Bassi et al., 2024). This paper provides complementary evidence on the ability of liquidity and information together to enable small-scale microentrepreneurs to overcome market frictions that otherwise prohibit moving and constrain firm profits, even in the face of urban disamenities.

I contribute evidence on the complementarity between liquidity and information to existing research on the barriers to enterprise growth, and I show that underinvestment in firm location can partially explain low microenterprise profits. Prior work documents large returns to capital for microentrepreneurs (De Mel, McKenzie, and Woodruff, 2008), but researchers find low average effects of policies that relieve liquidity constraints for microentrepreneurs, such as the expansion of microcredit (Angelucci, Karlan, and Zinman, 2015; Attanasio et al., 2015; Banerjee et al., 2015; Tarozzi, Desai, and Johnson, 2015). I show that adding information to a liquidity intervention leads high-return entrepreneurs to take it up, providing a potential policy lever to induce take up by the optimal compliers in settings where there is both a take up decision and heterogeneity in returns, as in Beaman et al. (2023).

This paper expands prior work on liquidity and information frictions that inhibit the mobility of workers by showing how these frictions also constrain the mobility of firms. Bryan, Chowdhury, and Mobarak (2014) and Abebe et al. (2021) show that paying transport costs directly can increase incomes via access to jobs, while Doornik et al. (2021) shows similar impacts of extending credit for personal vehicle purchase. Larcom, Rauch, and Willems (2017) show that information about transit routes leads commuters to make more efficient routing decisions, and Baseler (2023) shows that information about the income prospects in cities increases rural to urban migration. This paper contributes new evidence on the complementary nature of liquidity and information frictions in relieving mobility constraints for entrepreneurs in particular.

This paper proceeds as follows: Section 1.2 describes the setting and motivating facts, Section 1.3 describes the model setup, Section 1.4 explains the experimental design and data collection, Section 1.5 describes the empirical analysis and main results, Section 1.6 explores the underlying mechanisms that can

explain the pattern of results, and Section 3.4 details the estimation of model parameters and simulation results.

1.2 Baseline Setting and Motivating Facts

I study mobile microentrepreneurs, primarily street vendors and motorcycle taxi drivers (“boda bodas”), a typically low-income group who may have both relatively low costs of moving and relatively high marginal utility of additional earnings. I begin by documenting substantial variation in microfirm profits across Kampala for a spatially representative sample of mobile microentrepreneurs. I find that profits associated with the 90th percentile parish are more than three times those associated with the 10th percentile parish, controlling for entrepreneur characteristics (education level, years of experience, gender, sector, and asset stock), as shown in Figure 1.1.

1.2.1 Sampling Procedure

In a baseline survey in October and November 2022, I sampled 3,745 business owners who (1) have one or fewer paid employees, (2) have a cell phone (to allow for tracking and mobile money transfers in the treatment phase), and (3) do not operate from a permanent structure or stall, including their home. Surveys take place at randomly selected points (latitude-longitude pairs), stratified by city parish, such that the sample is spatially representative of microenterprises in Kampala.⁸ This baseline data collection served several purposes: it provided a spatially representative cross-section of microenterprise profits across the city from which I measure the extent of unexplained variation in pre-intervention profits in Figure 1.1, identified the sample from which I primarily drew participants for the experiment, and provided the data needed to construct the information arm of the experiment. Figure 1.2 maps average microenterprise profits in each parish across the city and depicts the geographic coverage of the baseline data collection.

⁸To draw the baseline sample, I imposed a square lattice of points, 100 meters apart, over the area of Kampala and randomly selected an average of 15 “nodes” per city parish, or 1300 nodes in total across 86 city parishes. I then moved each randomly selected node to the nearest city road to ensure that it was accessible (i.e., not in the middle of a field or body of water). Enumerators were instructed to survey the entrepreneur located closest to the node who satisfied the inclusion criteria and consented to participation in the study (and enumerators worked in pairs, so one could check to the left and the other to the right). After this first survey, enumerators would then survey the first entrepreneur who again satisfied the inclusion criteria located a minimum of 100 meters to the left of the initial survey point. They repeated the same process to the right, resulting in three surveys per node. The 100 meter condition is to avoid surveying three entrepreneurs operating in a cluster (for example, at the same boda stage) and to reduce the likelihood of spillovers in the treatment phase. This survey protocol yielded an average of 45 surveys per city parish. Occasionally, it was not possible to find 3 entrepreneurs at a node, so a replacement node was drawn.

1.2.2 Sample Characteristics

The selection criteria yield a baseline sample of retail vendors (56%), such as those selling fruits, vegetables, clothing, or packaged foods; motorcycle taxi drivers (“boda bodas”) (29%); and entrepreneurs working in non-transportation services (15%), such as those providing car washes, shoe shining, or freshly cooked food. I aggregate retail vendors and those working in services together in a group termed “vendors” and describe the characteristics of the sample overall, as well as separately for vendors and bodas, in Table 1.1.

As intended, the sample is quite mobile: the median commute time from home to business location is 30 minutes (the average is 34 minutes), 93% are unregistered and have no official permit to use their business location, and only 22% primarily rely on repeat customers. Two-thirds of the sample believes that they could earn higher profits by locating elsewhere within the city, and a lack of liquidity is the modal self-reported constraint preventing moving to another business location, among those who believe that they could earn higher profits elsewhere. Anticipated moving costs to a new location, however, are only 3,400 UGX per day, or less than \$1 in November 2022.⁹ Despite claiming that they expect to make higher profits elsewhere, respondents earn low scores on a quiz assessing knowledge of profit levels across the city: the average score is equivalent to the expected accuracy rate under random guessing.¹⁰ These facts inform the decision to ease liquidity and information constraints in the experiment.

Only 25% of the sample is female, perhaps the result of selecting for a mobile sample (i.e., not operating from a permanent structure, including the home). All of the female entrepreneurs in the sample are vendors; there are no female boda drivers. On average, entrepreneurs in the sample are 30 years old and have five years of experience. The entrepreneurs report low levels of education: only 15% have completed secondary school, while about 70% have completed primary school.

Average daily revenue is 60,000 Ugandan shillings (UGX), or \$16. Revenues are typically higher for vendors than bodas, but so are costs. Average daily profit across the sample is 23,000 UGX (\$6) but 21,000 UGX among vendors and 27,000 UGX among bodas. Asset levels also differ between vendors and bodas: while vendors typically have only 188,000 UGX (\$51), or about nine times their daily profit, in assets, boda drivers have nearly 3,000,000 UGX (\$810) in assets, where the bulk of their asset stock is the motorcycle that they use to transport customers.

No boda driver and nearly no vendor hires any paid labor, indicating that the entrepreneur supplies all of the labor to the business (an average of 67 hours per week). Inventory is the primary input cost incurred

⁹At the time of the baseline survey in November 2022, the exchange rate is approximately 3,700 UGX to 1 USD.

¹⁰The quiz asks a series of questions with binary responses, where the entrepreneur is asked whether they believe that typical profits are lower or higher in X other city parish for businesses operating in their same sector. I score the quizzes using the parish-sector profit rankings, calculated from the information collected in the baseline survey. A score of 50% is equivalent to the expected score under random guessing.

by vendors, while fuel is the primary input cost incurred by bodas. Only 7% of the sample pays any sort of rent, formal or informal, to use their space for their business.¹¹ Approximately 60% of the sample has any savings at all, with vendors having average savings of 212,000 UGX (\$57) and bodas having more than twice that, 443,000 UGX (\$120). This level of savings far exceeds entrepreneurs' anticipated moving costs, suggesting that the reported insufficient liquidity to move to a more profitable place may not result from incomplete savings markets. About a quarter of the sample reports having any credit.

The baseline data yield several facts that guide the model setup and experimental design:

1. The profits associated with the 90th percentile most profitable city parish are more than three times those associated with the 10th percentile parish, for observationally similar microentrepreneurs.
2. Microentrepreneurs in the sample are relatively mobile (i.e., the mean commute from home is over 30 minutes) and rarely pay rent for use of space for their business place.
3. Despite a lack of physical and contractual barriers to moving, the majority of entrepreneurs in the sample believe they could be more profitable elsewhere.
4. They identify liquidity as the primary constraint that prevents moving business locations. Although savings far exceed anticipated moving costs and almost a quarter of the sample has access to credit, liquidity may be desired for other reasons, e.g., as insurance if moving is profitable but risky.
5. While entrepreneurs believe that they could be more profitable elsewhere, they have low levels of knowledge about spatial profit variation in the city.

1.3 Model

I use a spatial equilibrium model (e.g., [Ahlfeldt et al., 2015](#); [Monte, Redding, and Rossi-Hansberg, 2018](#)), estimated with data collected during the experiment, to simulate aggregate outcomes under counterfactual policies. I adapt the model to reflect the business location decisions of microentrepreneurs who do not pay rent for business space. These entrepreneurs consume directly out of their business profits and choose locations to maximize utility, not profit, reflecting the well-documented lack of separation between consumption and production decisions in low-income settings ([Jayachandran, 2006](#); [LaFave and Thomas, 2016](#)).

¹¹I document the frequency with which microentrepreneurs pay rent at baseline to characterize the sample. Ultimately, those paying fixed monthly rent will be excluded from the experimental sample.

1.3.1 Entrepreneur's Preferences

The utility, U_{jknd} , of entrepreneur j operating in sector $k \in \{\text{Retail, Transport}\}$ with incumbent location (city parish) n who moves to destination parish d depends on income (i.e., business profit), amenities, and costs as follows:

$$U_{jknd} = \frac{A_{kd}z_{jkd}}{c_{nd}\nu_{kd}} \ln(Y_{kd}) \quad (1.1)$$

Here Y_{kd} captures average income for businesses in sector k at parish d , A_{kd} captures nonmonetary amenities, and c_{nd} captures the utility cost of moving. The parameter ν_{kd} is the coefficient of variation (standard deviation divided by the mean) of income. Its presence in the denominator indicates that variability in income lowers utility in a manner similar to moving costs. I assume that the idiosyncratic distribution of amenities is distributed Fréchet, $z_{jkd} \sim \text{Fréchet}(\theta_k)$, as is typical in the spatial literature, and I allow the shape parameter, θ_k , to vary by sector.¹²

In order to move, entrepreneurs must have exogenous liquid capital (i.e., savings), a_j , greater than or equal to $\bar{a}$. This could reflect a start-up cost incurred to explore alternative locations or a buffer stock that serves to insure risk in moving. Contingent on having $a_j \geq \bar{a}$, the entrepreneur faces an unconstrained, discrete-choice utility maximization. Otherwise, the entrepreneur cannot move, and their indirect utility simply equals that associated with their incumbent business location. The entrepreneur's *ex ante* indirect utility is as follows:

$$V(j, k, n) = \begin{cases} \max_d \frac{A_{kd}z_{jkd}}{c_{nd}\nu_{kd}} \ln(\delta_{jkd}Y_{kd}), & \text{if } a_j \geq \bar{a} \\ \frac{A_{kn}z_{jkn}}{\nu_{kn}} \ln(\delta_{jkn}Y_{kn}), & \text{otherwise} \end{cases} \quad (1.2)$$

Here $\delta_{jkd} \sim \text{Log-normal}(\lambda_{kd}, \sigma_{kd}^2)$ is idiosyncratic knowledge of income, where inaccurate knowledge has a mean distortion of λ_{kd} with variance σ_{kd}^2 . This knowledge shock may distort the entrepreneur's *ex ante* location decision but does not feature in realized, *ex post* welfare. Solving equation (1.2) yields N_{kd} , the number of firms of a given sector choosing to locate in parish d .

1.3.2 Determination of Profits in General Equilibrium

In equilibrium, as parish-level demand curves are downward-sloping, profits will be determined by jointly solving the consumer's problem and firm's problem, which I model as follows:

¹²Idiosyncratic amenities, or idiosyncratic preferences, could be, for example, a preference by a vendor who cooks fresh food to locate in a place with sufficient sidewalk space to set up a portable stove. In this setup, the Fréchet scale parameter is normalized to 1, as the A_d term will reflect average amenity value in each location. The Fréchet shape parameter, θ_k , is an inverse measure of dispersion where a smaller θ_k indicates greater dispersion between locations and a larger role for idiosyncratic preferences. I let it vary by sector to reflect that vendors ($k = \text{Retail}$) and bodas ($k = \text{Transport}$) may have different willingness to move, especially since bodas have direct access to capital that facilitates mobility via their motorbike.

Consumer's Problem. Households are assigned to exogenous locations d . Their total expenditure is divided between sectors $k \in \{\text{Retail, Transport, Other}\}$. Given prices $p_{kd}(j)$ and exogenous expenditure I_{kd} , households choose the quantity $x_{kd}(j)$ to purchase from each firm j :

$$\begin{aligned} U_{kd} &= \max \left(\int x_{kd}(j)^{\frac{\sigma-1}{\sigma}} dj \right)^{\frac{\sigma}{\sigma-1}} \\ \text{subject to } \int p_{kd}(j) x_{kd}(j) dj &\leq I_{kd} \end{aligned} \quad (1.3)$$

where σ is the elasticity of substitution between goods sold by different firms.

Firm's Problem. Given prices $p_{kd}(j)$ and cost function $C(\cdot)$, firms choose quantity $x_{kd}(j)$:

$$Y_{kd} = \max p_{kd}(j) x_{kd}(j) - C(x_{kd}(j)) \quad (1.4)$$

Profits result from jointly solving the consumer's problem and firm's problem:

$$\ln(Y_{kd}) = \ln(\Phi_{kd}) - \ln(N_{kd}) \quad (1.5)$$

Profits depend on a location-specific parameter, Φ_{kd} , which captures the baseline profitability of a parish, and the number of firms, N_{kd} , in parish d . The full derivation of equation (1.5) is in Appendix 1.A.

Solving equations (1.2) and (1.5) together yields entrepreneurs' optimal location choices in general equilibrium.

1.3.3 Model Predictions

The model implies the following about the entrepreneur's decision:

1. Entrepreneurs are more likely to locate in places with higher profits, Y_{kd} , and with greater amenity value, A_{kd} , all else equal.
2. They are less likely to locate in places with greater variation in profits, ν_{kd} , and to which they face higher costs of relocation, c_{nd} , all else equal.
3. For high values of θ_k , entrepreneurs will be more responsive to changes in the location-specific value of moving to destination d , V_{kd} . For low values of θ_k , idiosyncratic preferences will dominate.
4. An entrepreneur's knowledge of profits throughout the city, δ_{jkd} , can distort their location decision. On average, entrepreneurs are less likely to locate in places where profits are underestimated ($\lambda_{kd} < 0$) and more likely to locate in places where profits are overestimated ($\lambda_{kd} > 0$).

In Section 3.4, I use data collected through the experiment to estimate model parameters and simulate aggregate impacts on income and welfare.

1.4 Experimental Design and Data Collection

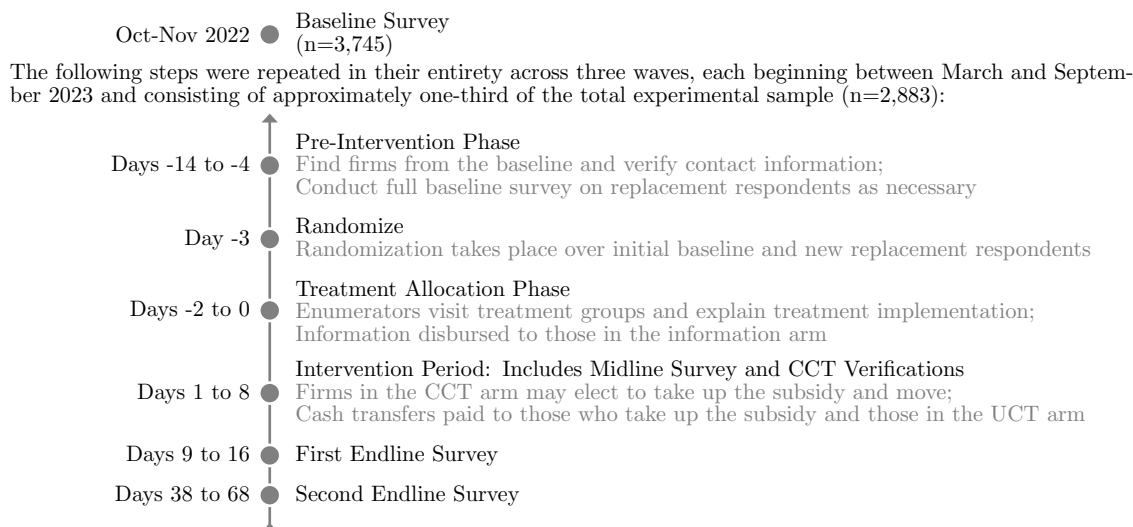
The experiment cross-randomizes an information intervention with each of two different liquidity interventions, a moving subsidy and an unconditional cash transfer of equivalent value, to test whether entrepreneurs can increase profits by moving business locations and whether liquidity and/or information constrain their location decisions. The moving subsidy, or a conditional cash transfer (CCT) that pays entrepreneurs who move their business location three or more kilometers from their pre-intervention location, induces random variation in business location. The moving subsidy explicitly increases the value of moving to a destination (V_d , in the model) that is three kilometers or more from the entrepreneur’s pre-intervention business location. The unconditional cash transfer alleviates liquidity constraints in general. Comparing the effect of taking up the unconditional transfer to taking up the moving subsidy will indicate whether entrepreneurs on average put the cash to a higher or lower return use than the return realized in a new business location. The information intervention tests whether information frictions constrain entrepreneurs’ location and investment decisions. This yields a 3x2 design, as depicted in Table 2.3.1.

1.4.1 Experimental Sample and Randomized Treatment Assignment

The experimental sample – those randomly assigned to one of the six treatment or control arms – consists of 2,883 entrepreneurs. When possible, I draw the experimental sample from the baseline sample ($n=3,745$). A baseline participant is eligible for retention in the experimental sample if they (1) successfully received a small mobile money transfer after participating in the baseline survey, as verification that they are able to receive mobile money transfers, (2) do not operate from the most profitable parish for their sector, i.e., there is at least one parish with higher average profits about which they can be informed if assigned to the information arm, (3) do not pay fixed monthly rent (a binding condition for less than 7% of the baseline sample), and (4) are successfully found and re-surveyed during a short interview after the baseline and immediately preceding randomization and treatment allocation (the “pre-intervention” phase). Ultimately, 1,691 (59%) of the 2,883 experiment participants are drawn from the baseline sample, while the other 1,192 (41%) are “replacement” respondents, who meet both the initial inclusion criteria for the baseline sample (as laid out in Section 1.2), as well as “retention” conditions (1) - (3).¹³ The sequence of the baseline survey,

¹³Retention condition (4) is irrelevant for replacement respondents, as they are solely interviewed in the pre-intervention phase. The most common reason for replacing baseline respondents was a failure to meet retention criterion (4): the baseline respondent could not be found again during the pre-intervention phase immediately preceding treatment allocation, reflecting

pre-intervention phase, randomization, treatment allocation, intervention, and data collection is summarized below:



Once the experimental sample has been selected through the baseline survey and pre-intervention phase, I randomly allocate each participant to one of the six arms, stratifying over business sector (vendor; boda; services), geographic division in Kampala (an administrative unit, of which there are five in the city), gender (female; male), and participation cohort (of which there are three).¹⁴ Randomization is at the level of the individual entrepreneur. Table 1.3 depicts balance between treatment arms on baseline sample characteristics. Characteristics are generally balanced across treatment arms, and the frequency with which I reject the null of equality (at the 95% confidence level) is in line with the number of rejections expected due to random chance. In a test for joint orthogonality across treatment arms, I reject the null of joint orthogonality for two of 19 baseline characteristics. In 95 pairwise tests (19 covariates $\times$ 5 tests) for equality on baseline characteristics between the control and each treatment arm, I reject the null in 7 instances. Table 1.B.1 summarizes the sample composition along sector and gender, two of the stratification variables, which thus are not featured in the balance table. The sample is heavily male (70%), perhaps due to selecting a mobile sample, and is composed of 62% vendors, 26% bodas, and 12% other services.¹⁵ The modal vendor

a high level of transience in a sample of low-income, urban microentrepreneurs operating largely informal businesses that lack a permanent structure. The baseline survey was conducted in October and November 2022, while the treatment phase was initiated in waves between March and October 2023, so 4 - 12 months elapsed between the initial baseline and the pre-intervention follow-up. Randomization is orthogonal to whether an entrepreneur is an original or replacement respondent, and so this does not confound the estimation of treatment effects.

¹⁴For logistical purposes of implementing the treatments, especially the conditional cash transfer, I divide the experimental sample into three cohorts of roughly equivalent size (approx. n=960 in each cohort). The pre-intervention phase, randomization, treatment allocation, intervention, and endline surveys take place in succession within a given cohort and are repeated in their entirety for each cohort. As I stratify treatment assignment over cohort, all six experimental arms are represented within a given cohort. This guarantees that no more than one-third of the sample participates in the experiment at any given time, enabling the enumerator team to carefully verify compliance with the three kilometer moving condition, among those assigned to and taking up the CCT.

¹⁵Note that the composition of the experimental sample (n=2,883) is slightly different from the full baseline sample (n=3,745), as it consists of only a subset of those surveyed at baseline plus some replacement respondents. The baseline characteristics of the experimental sample are summarized in Table 1.3, while the baseline characteristics of the full baseline sample are

sells fruits or vegetables.

1.4.2 Implementation of the Interventions

A “treatment allocation” visit takes place 1 - 3 days before the intervention period begins to explain the implementation logistics to all entrepreneurs randomly assigned to a treatment (non-control) arm. If an entrepreneur is randomly assigned to receive both a liquidity intervention (either the subsidy or cash transfer) and the information intervention together, then implementation is additive: the entrepreneur receives the liquidity intervention exactly as an entrepreneur receiving the liquidity intervention alone plus the information intervention in the exact form that it is disbursed to an entrepreneur receiving the information intervention alone.

Moving Subsidy, or Conditional Cash Transfer (CCT). The moving subsidy, or conditional cash transfer (CCT), allocates 7,000 Ugandan shillings (UGX), or approximately \$2 per day, on any day in an eight-day intervention period that an entrepreneur locates three or more kilometers from their initial (pre-intervention) business location. This distance is measured in road travel distance. The 7,000 UGX incentive is equivalent to 35% of median daily income and covers entrepreneurs’ anticipated daily moving costs to a more profitable location for more than 90% of the sample. The subsidy is paid out on a daily basis to mimic the time horizon on which entrepreneurs in the sample make investment decisions (the modal entrepreneur incurs 1-2 input costs, typically inventory or transport costs, and pays both on a daily basis). I select three kilometers as the distance at which an entrepreneur becomes eligible for the CCT based on the typical distance that entrepreneurs must travel to access more profitable parts of the city. As depicted in Figure 1.B.1, more than 60% of the sample is able to access a parish with average profits that are at least 20% higher than average profits in their own parish within three kilometers. Furthermore, I find that less than 20% of the sample moves more than three kilometers between the baseline survey and a pre-intervention follow-up survey (among those in the experimental sample who are retained from the baseline), suggesting that the conditionality will require most entrepreneurs to deviate from their typical habits. The CCT is disbursed via mobile money, to the mobile phone number confirmed during the treatment allocation visit, after an enumerator verifies the respondent’s new business location and records the geopoint of the new location. Detailed information on the process of verifying compliance with the 3 km conditionality is provided in Appendix 1.B. The requirements and logistics of verification are communicated to the entrepreneur during the treatment allocation visit, immediately before the eight-day intervention period begins.

Unconditional Cash Transfer (UCT). The unconditional cash transfer equivalently allocates 7,000 UGX per day during each day of the same eight-day intervention period and is also allocated using mobile

summarized in Table 1.1.

money. If the study team could not reach an entrepreneur in the UCT arm to conduct the treatment allocation visit – where the study team explained the entrepreneur’s treatment assignment and logistics of implementation, i.e., that a 7,000 UGX transfer would be disbursed each day via mobile money for eight days – then no transfer was made. This is both a cost-savings measure and to maintain comparability with the implementation of the CCT: if an entrepreneur in the CCT arm could not be reached during the treatment allocation phase, then they were not aware of their eligibility for the moving subsidy and thus would be unable to take it up. Likewise, if an entrepreneur in the UCT arm was unreachable during treatment allocation, then they were not aware of their assignment to the UCT arm and so no transfer was made. About 10% of entrepreneurs in both the CCT and UCT arms were unreachable during treatment allocation and so did not take up their respective liquidity intervention.

Information Intervention. The information treatment uses the baseline distribution of profits across the city to inform respondents about parishes where profits for businesses in their sector are typically higher than in their own parish. Respondents receive information about the average difference in profits for businesses of their same sector in up to four other parishes, randomly selected from among those where the average sector profit exceeds that in the entrepreneur’s own parish.¹⁶ The profit differential about which the respondent is informed is the difference in the parish-sector fixed effect from the entrepreneur’s own parish-sector to the randomly selected parish about which they receive information. To increase the likelihood of providing useful information, I restrict the set of parishes about which an entrepreneur could be informed to the parishes with a parish-sector fixed effect among the 90% most precise within a given sector, and I double the probability that an entrepreneur is informed about parishes in the top profit quartile within their sector.

In addition to the name of the parish and the profit differential, I also share a landmark in the parish for those who may not be familiar with administrative boundaries. Figure 1.3 shows an example “information card” that a respondent assigned to the information treatment would receive during the experiment. This information is provided verbally by the enumerator during the treatment allocation visit, on a physical card with text in both English and Luganda, and in a follow-up text message.

1.4.3 Data Collection

I collect survey data before, during, and shortly after the intervention period, to document how firms respond to the intervention while they are actively receiving subsidies and cash transfers, as well as whether impacts

¹⁶All respondents in the information arm receive information about four randomly selected parishes, unless there are fewer than four with higher average profits than the entrepreneur’s own parish, for businesses of their same sector. There is always at least one parish about which to inform the respondent, as entrepreneurs located in the highest profit parish for their sector were dropped from the experimental sample.

persist following the conclusion of the intervention. Surveys take place at the following points, and the sequence of data collection relative to the intervention is given in the timeline in Section 1.4.1:

1. Baseline survey (n=3,745): Comprehensive initial survey in October - November 2022 or immediately preceding the intervention period for replacement respondents
2. Pre-intervention survey (n=2,883): Brief follow-up for those retained from the baseline to experimental sample (“original” respondents) that verifies contact information and pre-intervention location

Randomization then takes place over the respondents surveyed in the pre-intervention phase (the 2,883 respondents who compose the experimental sample). After randomization, I attempt to re-survey all treatment and control respondents at the following points:

3. Midline survey (retained 2,410 of 2,883 respondents): Brief survey that takes place during the intervention
4. First endline survey (retained 2,500 of 2,883 respondents): Comprehensive follow-up survey that takes place in the week after the intervention concludes
5. Second endline survey (retained 2,596 of 2,883 respondents): Comprehensive follow-up survey that takes place 1-2 months after the intervention concludes

The baseline and pre-intervention survey are one and the same for replacement respondents (n=1,192), while they are two separate surveys for original respondents (n=1,691). After the pre-intervention survey and following randomization, replacement and original respondents are surveyed in an identical manner. I also conduct a very brief survey at the time of CCT compliance verification, for anyone taking up the CCT. These data lack any experimental variation, as they are only collected from those who comply with the CCT. Thus, these data are only used for descriptive purposes, especially related to describing where entrepreneurs moved. I use these data to characterize where and how far entrepreneurs taking up the CCT moved in Section 1.5.1.

Baseline Survey. The baseline survey captures GPS firm location, the business owner’s demographic characteristics (e.g., gender, age, education level, years of experience), characteristics of the business (e.g., business sector and type of goods sold, if in retail), typical daily revenue, value of and frequency with which costs are incurred (separately for each type of cost), number of hours that the entrepreneur works on their own business, mode of transportation and commute time from home to business location, business assets (including inventory), personal savings and credit, knowledge of business profits in other locations in the city, beliefs about barriers to accessing other locations, future business expansion aspirations, and finally patience and time consistency through an incentivized cash lottery.

Pre-Intervention Survey. The brief pre-intervention survey primarily serves to verify contact information and location of original baseline respondents, as 4-10 months have elapsed from the baseline, depending on the cohort. It takes place in the two-week period immediately preceding randomization and the intervention period. The survey once again records GPS firm location (the location from which compliance with the CCT will be determined, for those randomly assigned to the CCT arm), gives the respondent the opportunity to update their phone number if it has changed since the baseline, and collects information on a few dimensions that were not included in the baseline survey. These new questions include the number of competitors (other businesses selling the exact same good or service) within eyesight of the respondent’s business (counted and recorded by the enumerator), where vendors buy inventory (i.e., which market), frequency of trips to the market, and the type of fruit or vegetable sold by any vendors for whom this is their primary good. If the respondent is a replacement respondent, then the pre-intervention and baseline surveys are coalesced into one.

Midline Survey. The midline survey is also a brief follow-up survey that takes place during the eight-day intervention period. The midline survey collects GPS location, the number of competitors surrounding the entrepreneur at their business location, their mode of transportation to work, number of hours worked so far that day, revenue earned so far, costs so far, expected profit for the day, mobile money balance, and whether and where (i.e., which market) the entrepreneur has purchased inventory. The midline survey also collects profit earned in the prior day and approximate location in the prior day, but the aim of collecting revenue and costs so far at the moment of the survey is to tie firm performance to an exact GPS location. While the midline survey takes place during the intervention, it is entirely independent of the CCT verification and associated CCT verification survey. For example, an entrepreneur taking up the CCT might participate in the midline survey early in the day, before they have reached their primary business location for the day, or on a day during the intervention period in which they elect not to take up the CCT. For these reasons, the entrepreneur’s location in the midline survey may not be the same as at the time of CCT verification.

Endline Surveys. The first endline survey takes place in the week immediately following the conclusion of the intervention, and the second endline survey takes place one to two months following the conclusion of the intervention. These surveys closely resemble the baseline survey, collecting GPS location and detailed information on firm revenues and costs, assets, market access, savings, credit, knowledge of profit variation across the city, and business aspirations.

Attrition from the Sample. I retain 84%, 87%, and 90% of respondents at the midline, first endline, and second endline wave, respectively. Attrition is higher at midline than in subsequent survey rounds because there is a shorter period of time over which to track respondents: if a respondent cannot be found and surveyed during the eight-day intervention period, then they will have attrited from the midline sample.

In contrast, the first and second endline surveys do not have the same fixed end date after which a respondent can no longer be tracked. Midline respondents do, however, appear to be missing at random. The rate of attrition in each treatment arm is not statistically significantly different from that of the control group, and I cannot reject the null of joint orthogonality across all treatment arms (Table 1.B.2).

1.5 Empirical Analysis and Results

Of those assigned to the moving subsidy (CCT) arm, 55% take it up and move business locations on at least one day of the intervention period. This indicates that the majority of the sample perceives positive net value of moving when offered the subsidy, or more specifically that the net value of moving for one day exceeds -\$2 for 55% of the sample. I find that only firms offered both the moving subsidy (the CCT) and information are more likely than the control group to move into a city parish with higher average profits for businesses of their same sector. They also increase revenues by 14% relative to the control group, the only group to report statistically significant revenue gains. Those offered both the subsidy and information have differentially higher profits relative to those receiving information alone, and they also differentially realize a higher fraction of expected profits in the parishes they move to, suggesting that the combination of the subsidy and information enables a more profitable entrepreneur-location match.

1.5.1 Take-Up and Description of Moves

As shown in Table 1.B.3, 54% of those in the CCT only group take up the subsidy, while 56% of those in the CCT+Info group take it up, but the difference is not statistically significant. The baseline characteristics of who take up the subsidy do, however, vary depending on whether it is offered with or without information. Because the rate of compliance is similar between groups, this indicates a compositional change based on the presence of information with the subsidy offer. As shown in Table 1.B.4, those taking up the subsidy when it is offered with information have statistically significantly ($p < 0.10$) higher baseline knowledge of profit in nearby parishes and lower savings than those taking up the subsidy without information. They also have longer commutes ($p = 0.11$).¹⁷ These differential characteristics of compliers have implications for the impact of the subsidy, which I discuss further in Section 1.6.2.

Using the data collected from entrepreneurs during CCT verification, I find that firms taking up the CCT

¹⁷I analyze complier characteristics using the approach laid out in [Abadie \(2002\)](#) and [Angrist, Hull, and Walters \(2023\)](#), where I instrument for compliance with treatment assignment and use 2SLS to analyze when baseline characteristics differ because of complier status. Note that the never taker mean in Table 1.B.4 is solely descriptive: it is the average value observed for those who were offered the CCT but did not take it up. The complier mean, however, is identified using variation in complier status that results from treatment assignment.

do not solely move the minimum distance required: the median distance moved is 4.8 km.¹⁸ Entrepreneurs also experiment with several new locations. The average entrepreneur taking up the CCT moves to five different parishes in the eight-day intervention period. Only 9% of firms receiving both the CCT and information ever move to a parish about which they were informed, suggesting that to the extent that information may influence outcomes, it is not by directing entrepreneurs to locate in the particular parishes about which they were informed. Together with the fact that those taking up the subsidy when it is offered with information have higher baseline knowledge of profits across the city suggests that any impact of information in combination with the subsidy does not come from informing entrepreneurs about average differences in profit, but perhaps from a higher dimensional moment, such as variability in perceptions about profit across the city.

Of those assigned to the UCT only group, 90% take it up, while 87% of those in the UCT+Info group take it up (also not statistically significantly different between groups). Entrepreneurs who were unreachable during the treatment allocation phase account for the imperfect take-up of the UCT: as discussed in Section 1.4.2, if an entrepreneur could not be reached for mobile phone number verification, then the study team did not send the transfer.¹⁹

1.5.2 Estimation Strategy

I estimate treatment effects using the following OLS specification, in which I regress each outcome of interest on treatment assignment. For firm j in stratum s , the intent-to-treat (ITT) specification is as follows:

$$Y_{js} = \beta_0 + \beta_1 info_j + \beta_2 CCT_j + \beta_3(info_j \times CCT_j) + \beta_4 UCT_j + \beta_5(info_j \times UCT_j) + \delta X_j + \phi_s + \epsilon_{js} \quad (1.6)$$

where Y_{js} is an outcome for firm j in stratum s , $info_j=1$ if the firm is randomly assigned to the information arm, $CCT_j=1$ if the firm is randomly assigned to the CCT arm, $UCT_j=1$ if the firm is randomly assigned to the UCT arm, X_j is an optional vector of baseline characteristics, and ϕ_s are strata fixed effects.²⁰ I use double-selection lasso to select which of 28 pre-specified baseline controls are included in X_j , with strata fixed effects automatically included as non-penalized controls. Standard errors are heteroskedasticity-robust. I winsorize all continuous outcomes to the 2.5th and 97.5th percentiles, to mitigate the possibility of outliers driving average effects.

¹⁸Note that Section 1.5.1 reports descriptions of the moves made by those taking up the moving subsidy, not the effect of treatment, which I discuss beginning in Section 1.5.2.

¹⁹Recall that this was a cost savings measure, as well as for comparability with the CCT arm. If an entrepreneur assigned to receive the CCT could not be reached during the treatment allocation phase, then they would not have known about their eligibility and thus could not take up the subsidy.

²⁰Recall that randomization was stratified over participation cohort, geographic division, business sector, and gender.

The β terms are the coefficients of interest, where β_1 is the effect of receiving information alone, β_2 is the effect of the CCT offer alone, β_3 is the differential effect of receiving information and a CCT offer together, β_4 is the effect of the UCT offer alone, and β_5 is the differential effect of receiving information and a UCT offer together. While β_1 , β_2 , and β_4 indicate the total effect of being assigned to the information only, CCT only, or UCT only treatment arms, respectively, $\beta_1 + \beta_2 + \beta_3$ indicates the total effect of being assigned to the CCT with information arm, and $\beta_1 + \beta_4 + \beta_5$ indicates the total effect of being assigned to the UCT with information arm. β_3 alone indicates whether the CCT and information are substitutes (-), complements (+), or neither. Likewise, β_5 alone indicates whether the UCT and information are substitutes (-), complements (+), or neither.

Local Average Treatment Effect. I also estimate the local average treatment effect (LATE) by instrumenting take-up with treatment assignment in a two-stage least squares (2SLS) setup. I consider an entrepreneur as having taken up the CCT if they took up the subsidy by moving three or more kilometers on at least one day of the intervention period. I consider an entrepreneur as having taken up the UCT if our study team was able to reach the entrepreneur immediately before the intervention to verify their mobile money phone number, through which the transfer will be sent.²¹ I assume 100% compliance, or take-up, with the information treatment in that there was no relevant take-up decision: anyone randomized to the information arm was exposed to the information, as information was provided verbally, on a written card, and through text message, regardless of whether or not the study team could reach the entrepreneur during the treatment allocation phase. In this experiment, there is no possibility of “always taking” – that is, there is no possibility that the CCT or UCT was delivered to someone not assigned to the CCT or UCT arm, respectively, because the study team only made mobile money transfers to the set of phone numbers associated with respondents assigned to the relevant arm. Thus, the local average treatment effect (LATE) is the same as the treatment-on-treated (ToT) effect.

The exclusion restriction must hold for the LATE to yield an unbiased treatment effect for compliers: the offer of the CCT must only affect outcomes through take-up of the CCT, and the offer of the UCT must only affect outcomes through take-up of the UCT. I consider an entrepreneur to have taken up the CCT if they moved three or more kilometers and received the subsidy at *any point* during the intervention period. Even if an entrepreneur only took up the incentive for one day of the eight day intervention period, they are considered as having fully taken up the incentive (in the same way that an entrepreneur who takes up the CCT on every day of the intervention is considered as having taken up the incentive). Conceivably, an experience moving one day could affect outcomes on another day when the entrepreneur does not take up

²¹If the entrepreneur could not be reached, then the transfer was not sent, such that the entrepreneur did not take up the transfer, for the reasons described in Section 1.4.2.

the subsidy, and so this threat to the exclusion restriction is mitigated by considering anyone who took up the subsidy on any single day of the experiment as having fully taken it up. Concerns that exclusion might be violated for those offered the CCT with information – if information together with the offer of the CCT somehow influenced behavior even without taking up the CCT – are alleviated by the lack of an impact of the information intervention alone. It is unlikely that information alone has no impact but information with the offer of the CCT has an impact absent taking up the CCT. Entrepreneurs assigned to the UCT arm only failed to take it up when they could not be reached during the treatment allocation phase, and so they were unaware of their eligibility. Their random assignment to the UCT arm, thus, is unlikely to affect outcomes absent taking it up, given that those who did not take up the UCT did not know they were eligible for it. The second stage estimating equation that yields the ToT effect is given in Appendix 1.B.

By instrumenting take-up with random assignment to the CCT arm, the treatment-on-treated effect enables me to measure the size of the revenue and profit effects for those entrepreneurs who did in fact move in response to the CCT, the set of entrepreneurs who, by revealed preference, have a positive net value of moving when offered the subsidy. This is useful for determining the cost-effectiveness of the subsidy, as the cost of paying out the subsidy is only incurred for compliers. For those in the UCT arm, the ToT indicates the effect of actually receiving cash. The ToT, thus, identifies the effect of taking up the CCT for those complying with the CCT and, separately, the effect of taking up the UCT for those complying with the UCT, but these complier groups could be different. As a result, a difference in the effect of taking up the CCT relative to the effect of taking up the UCT could reflect either an effect of moving only realized through the CCT or differences in returns for the type of entrepreneur who takes up the CCT versus the UCT.

Machine Learning to Predict Take-Up with Baseline Characteristics. To address this, and to better understand how the effect of moving compares to the returns to cash for the set of people likely to take up the moving subsidy, I use lasso to predict which entrepreneurs are likely to take up the subsidy, when it is offered with information and, separately, when it is offered without information. I then construct a subgroup of “likely compliers”, as predicted with baseline characteristics, on which I test for heterogeneity in intent-to-treat effects. Consistent with the characterization of compliers in Table 1.B.4, lasso selects commute time and knowledge of spatial variation in profits as important predictors of taking up the subsidy when it is offered with information, but not when it is offered without information. Likelihood of incurring transit costs is the sole lasso-selected predictor of take-up that is common to both groups. The results of the lasso analysis are shown in Table 1.4.

Differences in complier characteristics for those taking up the subsidy with information relative to without suggests that the presence of information changes the characteristics of who chooses to take up the subsidy. Conducting heterogeneity analysis on the machine learning-predicted “likely complier” subgroup

allows me to test this directly, by comparing the differential effects for the group that is likely to comply with the CCT+Info relative the differential effects for the group that is likely to comply with the CCT alone. It also enables me to have a more direct comparison between the returns to cash in a new business location (i.e., the profit gains of taking up the CCT) and the returns to cash in general (i.e., the profit gains of taking up the UCT), all within the subgroup that is likely to take up the moving subsidy if offered it.

1.5.3 Effect on Business Location During the Intervention

Table 1.5 shows the effect of each treatment group on business location during the intervention. Note that columns 1 - 4 compose mutually exclusive groups: an entrepreneur either left Kampala, stayed in their same parish (within Kampala) as before the intervention, moved to a parish with lower average profits within Kampala, or moved to a parish with higher average profits within Kampala. Column 5 is the average profit for businesses of the entrepreneur's same sector in the parish where an entrepreneur locates during the intervention, computed using baseline survey data, and so provides a continuous measure of the profitability of the places where entrepreneurs move.²²

Focusing first on the intent-to-treat effects in the upper panel A: column 1 shows that about 8% of the total sample leaves Kampala during the intervention, and this likelihood does not change as the result of treatment assignment, indicating that moving outside of Kampala is unrelated to the entrepreneur's treatment group during the intervention. Column 2 shows that 54% of the control group remains in their same parish from baseline. Receiving information or a CCT offer, or both together, does not have a statistically significant effect on the likelihood of staying in the same parish, though those who receive the UCT only are 6 percentage points ($p < 0.10$), or 11%, less likely to stay in their baseline parish.²³ Those receiving the UCT with information are 7 percentage points ($p < 0.10$) less likely to stay in their baseline parish, and this rate is not statistically significantly different from those receiving the UCT alone. This suggests that those receiving unconditional cash may use it to leave their home parish – perhaps indicative of a general desire for mobility, which the UCT can fund – but entrepreneurs receiving the UCT alone are also 6 percentage points, or 32%, more likely than the control group to move into a lower profit parish (column 3). This effect is undone for entrepreneurs offered the UCT and information together, as they are no more or less likely than those receiving information alone to move into a lower profit parish, and neither those receiving information alone nor the UCT with information are statistically significantly more or less likely than the control group

²²I have no data on typical profits outside of Kampala, as the baseline sample consisted entirely of entrepreneurs located within Kampala's boundary. Thus, expected profits, in column 5, for any location outside of Kampala is computed as the average profits realized for entrepreneurs in the control group who left Kampala during the intervention. If this assumption is inaccurate, it will not bias treatment effects, as the likelihood of leaving Kampala does not vary with treatment assignment.

²³Note that the average city parish in Kampala is 2 square kilometers, so it's plausible that an entrepreneur taking up the moving subsidy could move 3 kilometers (as measured by road travel distance, which is typically longer than Euclidean distance, while staying within their baseline parish).

to move into a lower profit parish. Like those offered the UCT alone, those offered the CCT alone are also statistically significantly more likely ($p < 0.10$) to move into a lower profit parish, but again this is undone by the differential effect of receiving information and the CCT together, which reduces likelihood of moving into a lower profit parish by an additional 9 percentage points ($p < 0.05$).

The results in column 2 and 3 show that those offered the CCT or UCT alone are more likely than the control group to move out of their baseline parish, but they are also more likely than the control group to move into a parish with lower average profits, an impact which is undone when the subsidy is combined with information, suggesting that information is impactful in driving entrepreneurs to move strategically. This role of information is further underscored by the results in column 4. While the differential effect is not statistically significant, those receiving information in addition to the CCT are 5 percentage points more likely than those receiving information or the CCT alone to move into a higher profit parish, and in total they are 4 percentage points ($p < 0.10$), or 21%, more likely than the control group to move into a higher profit parish. The differential impact of combining the UCT with information is also positive, but neither the differential impact nor the total effect relative to the control group is statistically significantly different from zero. Finally, column 5 coalesces the extensive margin impacts documented in columns 1-4 into a single continuous measure of parish profitability and shows that those offered the CCT with information move to parishes with differentially higher profits ($p < 0.05$) than those offered information or the CCT alone. On average, those offered the CCT with information locate in parishes with profits that are only about 430 UGX ($p = 0.12$), or 2%, higher than average profits in the parishes where the control group locates.

The bottom panel B depicts treatment-on-treated effects, which only change meaningfully for those offered the CCT, where take-up is not universal (as is the case for information) or nearly universal (for the UCT). Column 3 of panel B shows that those taking up the CCT while receiving information are 10 percentage points ($p < 0.05$) less likely to move into a parish with lower average profits than those taking up the CCT without information or those receiving information alone. Only those taking up the CCT and receiving information are statistically significantly more likely to move into a parish with higher average profits: they are 7 percentage points ($p < 0.05$), or 37%, more likely to move into a parish with higher average profits than the control group. Average profits of the parishes where firms in the CCT with information group move are also 1,386 UGX, or 6%, higher than those in parishes where the CCT only group moves.

I also analyze the effect of each intervention on distance, in kilometers of road travel distance, from the pre-intervention business location. The distribution of distance traveled is right-skewed, with a small number of entrepreneurs moving very far distances. To avoid a small fraction of the sample driving the average treatment effect, I use distribution regressions to analyze the effect of treatment on distance traveled, in Figure 1.4. These distribution regressions show the effect of taking up each treatment on the probability

that distance moved exceeds X kilometers, where X spans the support of distance moved from the pre-intervention location, truncated at $X=6$ km, the 90th percentile. Figure 1.4 shows that entrepreneurs taking up the CCT revert towards their pre-intervention location, even during the intervention period. Those taking up the CCT without information are more likely than the control group to locate more than 1.5 km from their pre-intervention location, but they are less likely than the control group to locate more than 3.5 km from their pre-intervention location.²⁴ These probabilities are imprecise and never statistically significantly different from zero, again underscoring that entrepreneurs who take up the CCT and move vary in how long and where they stay during the intervention period. For those taking up the CCT with information, the effect of the intervention on the probability of moving more than X km from their pre-intervention location is positive for all $X > 1.5$ km, suggesting that their willingness to locate further away may be higher than those taking up the CCT alone. The distribution regressions of those taking up the UCT, with or without information, are similar to those taking up the CCT, which may indicate access to liquidity in general increases mobility.

Reversion back towards the pre-intervention location, even for those taking up the CCT, whose GPS location verifies that they moved more than 3 kilometers at some point during the intervention, could happen for several reasons. First, *bodas* in particular may use the subsidy to initially catch customers in a new location, for example during rush hour, but then return to their pre-intervention location between customers. Second, an entrepreneur could take up the CCT on some days but not others, and the midline survey may or may not take place on the particular day on which they took up the CCT, leading to high levels of variance in the impact of taking up the CCT on distance moved. Third, the 3 kilometer condition for taking up the moving subsidy may be inefficient: perhaps entrepreneurs would like to move, but not the full 3 kilometers (supported by the positive, though insignificant, effect of taking up the UCT on the probability of locating more than X km from the pre-intervention location, for small values of X). If the net value of the subsidy is positive despite this inefficiency, then entrepreneurs may take it up, temporarily operate elsewhere, and then move to the place where they would actually prefer to locate, which is less than 3 km from their initial location. The extensive margin impacts of the intervention on likelihood of locating in a more profitable parish indicate that the subsidy did influence place of business operation. Together with the muted impacts on distance traveled, these results provide evidence that entrepreneurs are willing to move but prefer to stay relatively close to their individually-chosen pre-intervention business location. This is consistent with a large role for idiosyncratic preferences in entrepreneurs' location decisions.

In sum, the combination of cash (either the CCT or the UCT) with information appears to mitigate

²⁴That is, the effect of taking up the CCT without information on the probability of locating more than X km from their pre-intervention location is positive for $X = 1.5$, but negative for $X = 3.5$.

the likelihood of making “bad moves,” as those offered information with cash are differentially less likely to move into lower profit parishes than those offered cash alone. But only those offered the CCT and receiving information are statistically significantly more likely than the control group to move into an on average higher profit parish than the control group. The magnitude of the impact is large: those taking up the subsidy and moving are 37% more likely than the control group to move into a parish with higher average profits than their pre-intervention parish.

1.5.4 Effect on Financial Outcomes During the Intervention

After examining the impacts of the intervention on location choice, I turn to financial outcomes, to test the effect on more “downstream” outcomes, after the entrepreneur does (or does not) move business locations. In Table 1.6, I show impacts of each intervention on entrepreneurs’ mobile money balance, revenues, costs, profit, fraction of expected parish profit that is realized, hours spent working per day, and profit per hour. Beginning with the ITT effects in panel A, column 1: here, I show how the intervention impacted mobile money balances. The CCT and UCT each delivered 7,000 UGX per day to entrepreneurs who took them up. The amount held in the mobile money account, thus, reflects to what extent entrepreneurs retained elevated balances by saving some part of the transfer or withdrew it. Entrepreneurs who were offered the CCT alone, the UCT alone, and the UCT with information hold statistically significantly higher mobile money balances during the intervention. Those offered the CCT alone have a 3,220 UGX ($p < 0.05$), or 48%, higher balance than the control group, those offered the UCT alone have a 4,727 UGX ($p < 0.01$), or 71%, higher balance than the control group, and those offered the UCT with information have a 5,922 UGX ($p < 0.01$), or 89%, higher balance than the control group. This indicates that for those in these three groups, funds received from the subsidy and transfers were not spent right away and did result in elevated balances, at least in the short run. In contrast, those offered the CCT with information do not retain a statistically significantly higher balance than the control group. Given that this is the only group that was more likely than the control to move into higher average profits, this may reflect that a greater portion of the subsidy was spent on investment in location than was spent by other groups.

Turning next to the effects on revenue, cost, and profit, I show that only those offered the CCT with information have statistically significantly greater revenue than the control group (column 2). Their daily revenue is 5,309 UGX ($p < 0.05$), or 14%, higher than in the control. No other treatment group realizes statistically significant revenue gains. I do not find a statistically significant effect on costs for any group relative to the control group (column 3), suggesting that to the extent that the intervention changed financial outcomes, it was on the demand side rather than supply side. In column 4, I show that the combination of

information with both the CCT and UCT differentially increases profits relative to those receiving information, or either the CCT or UCT, alone. Those offered the CCT and information together realize 7,396 UGX ($p < 0.05$) greater profit per day than those receiving either the CCT or information alone, and those offered the UCT and information together realize 7,999 UGX ($p < 0.10$) greater profit per day than those receiving either UCT or information alone. The total ITT effect on profit for those in the CCT and information group, as well as those in the UCT and information group, is positive but not statistically significant.

In addition to realizing differentially higher profits, those offered CCT or UCT together with information also realize differentially higher profit per hour (column 6) and a differentially higher fraction of expected parish profit (column 7), without working differentially more hours (column 5). The outcome in column 7 is constructed as an entrepreneur’s individually realized daily profit (in column 4 of Table 1.6) as a fraction of average profit for businesses of their same sector in the parish where they choose to locate during the intervention (in column 5 of Table 1.5). Together, these results suggest that entrepreneurs offered liquidity (either the CCT or UCT) with information are able to realize differentially higher profits than those receiving liquidity or information alone, and without simply working differentially longer hours. The differential increase in profit per hour may be reflective of differential improvements in productivity – particularly since revenue and profit increases occur without an increase in costs – and the differential increase in the fraction of expected parish profit that is realized could reflect a better entrepreneur-location match, where liquidity and information together enables entrepreneurs to move to parishes where they are able to capture a larger fraction of the location-specific profit premium.

Turning to panel B, which contains ToT estimates, I show that only entrepreneurs taking up the moving subsidy while also receiving information realize statistically significantly greater revenue, profit, and profit per hour than the control group. They earn 9,732 UGX ($p < 0.01$), or 26%, more in revenue and 7,705 UGX ($p < 0.05$), or 45%, more in profits than the control group. There is no statistically significant effect on costs, and because revenue increases without a proportionate increase in costs, the percent gain in profit is higher than the percent gain in revenue. They also earn 613 UGX ($p < 0.10$), or 41%, more in profit per hour than the control group, and the small positive effect on hours worked per day is not statistically significantly different from zero. Even accounting for the additional 0.36 hours worked on average by those taking up the CCT with information, would only explain 759 UGX $(= (1,495 \text{ UGX/hour} + 613 \text{ UGX/hour}) * 0.36 \text{ hours})$, or less than 10% of the total profit effect, underscoring that the profit gains do not come from working longer hours. The total revenue and profit impacts for those receiving the UCT and information together are also positive, though about half the size as those for the CCT and information group and not statistically significant at conventional levels. As discussed in Section 1.5.2, a difference in ToT effect of the CCT relative to the UCT could reflect an effect of moving not realized by those taking up the UCT, or a difference in

complier characteristics – that the set of people taking up the CCT are different on average than the set of people taking up the UCT. To gain insight into this issue, I will test for heterogeneous treatment effects for the machine learning-predicted “likely complier” group in Section 1.6.

In sum, only those taking up the CCT while receiving information realize statistically significant revenue and income gains during the intervention. This is consistent with their tendency to move into higher profit parishes, though they are also differentially likely to realize a higher fraction of expected profit in the parishes where they move, suggesting that higher profit is not just the result of moving to more profitable places, but also moving to places which make a more profitable idiosyncratic entrepreneur-location match.

1.5.5 Persistence: Effect of Treatment After the Intervention Ends

The impacts of the intervention fade quickly after the subsidies and cash transfers cease. In Table 1.7, I show the effect of each intervention on the expected profit of the parish where an entrepreneur locates, the entrepreneur’s revenue, and the entrepreneur’s profit at the first endline (1 week after the intervention) and the second endline (1-2 months after the intervention). In the first week after the intervention, entrepreneurs in the CCT+Info group differentially locate in parishes with higher average profits: they locate in parishes with 851 UGX ($p < 0.05$) greater profits than entrepreneurs receiving offered the CCT or information alone. This differential effect, however, has faded by the second endline, 1-2 months after the intervention. In the first week after the intervention, both those offered the CCT+Info as well as those offered the UCT+Info earn higher revenues than the control group: those in the CCT+Info group earn 5,038 UGX ($p < 0.05$), or 12%, greater revenue than the control group, while those in the UCT+Info group earn 4,948 UGX ($p < 0.10$), also about 12%, more in revenue than the control group. While the effect of both the CCT+Info and the UCT+Info on profit during the week after the intervention is positive, neither point estimate is statistically significantly different from zero. By one month after the intervention — one month after moving subsidies have ended — the total effect on revenue is statistically indistinguishable from zero, for both the CCT+Info group and the UCT+Info group. The differential effect on the profitability of the parishes in which entrepreneurs locate, observed during and one week after the intervention for those offered the CCT with information, is also statistically indistinguishable from zero, and no effects emerge for other groups, on any dimension.

The financial outcomes during the intervention show that entrepreneurs can make highly profitable moves: those taking up the moving subsidy when it is offered with information realize 45% higher profits. This indicates that entrepreneurs can capture location-specific profit premia when moving is subsidized and information frictions are relieved. These impacts are not sustained after subsidies cease, but they do not

entirely dissipate immediately upon the conclusion of the intervention. One week after the intervention, entrepreneurs receiving the CCT with information still differentially move into parishes with higher average profits, where they realize higher revenues. The mechanism through which entrepreneurs realized higher profits, thus, must entail both a non-durable role of the subsidy, which fades eventually but not immediately after the intervention ends, and a complementarity between the subsidy and information that uniquely unlocks an entrepreneur’s ability to make profitable moves.

1.6 Mechanisms

I have shown that there is a large and significant positive effect on revenue and profit of taking up the moving subsidy when it is offered with information. The effect of receiving the UCT with information, meanwhile, is smaller and imprecise, but still positive. Liquidity and information constraints were alleviated for those receiving the UCT with information to at least the same extent as those taking up the CCT with information. Both groups received the same information intervention, while the UCT group received cash that could be used for any purpose. Even if all entrepreneurs desire liquidity primarily for moving, those receiving the UCT are free to accept the funds and move, in exactly the same fashion as those taking up the CCT. Understanding the reason for differences in the average treatment-on-treated effect of the CCT with information relative to the UCT with information illuminates the role of the subsidy and information.

1.6.1 Moving During the Intervention: Profitable or Preferable?

Differences in the treatment-on-treated (ToT) effect of the CCT+Info versus the UCT+Info may reflect a large fraction of the sample who can make profitable moves but have a small negative valuation of moving to a new destination, $V_d \in (-7000 \text{ UGX}, 0 \text{ UGX}]$, which is overcome when they are offered the 7,000 UGX subsidy. In this case, moving would be profitable but not welfare-maximizing absent the subsidy, and entrepreneurs should only make profitable moves when the subsidy makes the net valuation of moving positive. Entrepreneurs taking up the UCT would not move, as the unconditional cash that they receive does not change the valuation of moving. The small and imprecise, but positive, effect of the UCT and information would then reflect that the welfare-maximizing use of cash is less profitable than moving, such that the role of the subsidy was to induce moves that were profitable but not preferable. Alternatively, the differences in the effect of taking up the CCT with information relative to the UCT with information could reflect differences in complier characteristics between the two groups. The ToT is the treatment effect for those taking up the treatment (which in this experiment is perfectly analogous to the treatment effect for those complying with the treatment), and compliance varies between those taking up the CCT and the UCT.

While anyone who could be reached by the enumerator team takes up the UCT, only those who, by revealed preference, have a positive net valuation of moving when offered the subsidy take it up.

Did the subsidy induce profitable moves by entrepreneurs who, absent the subsidy would rather not move, or do differences in the characteristics of the complier groups drive differences in the average effect taking up each liquidity intervention? To answer this question, I use lasso to predict compliance with the CCT, separately when it is offered with information and without, using baseline characteristics. Table 1.4 shows the characteristics which lasso selects as the most important in predicting compliance. The 55% of the sample who appear most likely on baseline characteristics to comply with the treatment compose a group that I term “likely compliers”. I separately construct “likely CCT+Info compliers” and “likely CCT only compliers”. Those likely to comply with the CCT with information treatment have longer commutes, higher baseline knowledge of profits across the city, and are more likely located in lower profit parishes, among other characteristics in Table 1.4. Those likely to comply with the CCT alone are more likely to incur transit costs and have higher savings. Note that this indicator is orthogonal to treatment, as it is predicted solely from baseline characteristics.

If differences in average ToT effects between those taking up the CCT+Info relative to those taking up the UCT+Info reflect an effect of moving only realized in response to the subsidy, then I should see no differential effect for the group of likely CCT+Info compliers in response to taking up the UCT+Info. When presented with the UCT+Info, they would have used it for a different purpose than those taking up the CCT+Info. If, in contrast, the CCT+Info and UCT+Info function similarly and differences in ToT effects reflect differences in complier characteristics, then the differential effect of the UCT+Info as a likely CCT+Info complier should be similar to that of the CCT+Info. I test this directly in Table 1.8. I find evidence in support of the latter: in response to both the CCT+Info and the UCT+Info, likely CCT+Info compliers experience significant and positive differential effects of similar magnitude on revenue, profit, fraction of expected profit that is realized, and profit per hour. The total effect on daily profit for likely CCT+Info compliers offered the CCT+Info is 8,825 UGX ($p < 0.05$), and the total effect on daily profit for likely CCT+Info compliers offered the UCT+Info is a remarkably similar 8,743 UGX ($p < 0.05$). Like the average effects in Table 1.6, likely CCT+Info compliers offered the CCT+Info or UCT+Info experience no effect on number of working hours. In total, likely CCT+Info compliers offered the CCT+Info locate in parishes with 640 UGX higher profits ($p < 0.10$), though the differential effect is not significant. Likely CCT+Info compliers offered either the UCT+Info or information alone also locate in parishes with higher average profits, though neither the total nor differential effect is significantly different from zero. In Figure 1.C.1, I show that likely CCT+Info compliers offered the CCT+Info, UCT alone, or UCT+Info are more likely than the control group to locate more than 1 km from their pre-intervention location, but this effect

is not significantly different from zero. The effect of the CCT+Info and the UCT+Info on the likelihood of locating more than X km away remains positive for all values of X greater than 1 km, but again the impacts are imprecise across most of the distribution (as in Figure 1.4). As discussed in Section 1.5.3, there are many reasons why impacts on distance moved during the intervention could be noisy.

For the subgroup that is likely to take up the CCT when it is offered with information, the CCT+Info treatment and the UCT+Info treatment appear to function similarly. Whether offered the subsidy with information or the unconditional cash with information, this subgroup realizes large, statistically significant increases in revenue and profit. This subgroup also moves to parishes with higher average profits and may move further distances (though both of these impacts are imprecisely estimated). For this subgroup, moving appears both profitable and welfare-maximizing, as those offered the UCT+Info did not have to move but chose to do so anyway. The 3 km conditionality of the moving subsidy, however, may be inefficient, given that entrepreneurs realize similar profit gains whether or not they are subject to it, and entrepreneurs in the CCT+Info group revert to a business location within 3 km of their pre-intervention location.

1.6.2 Role of Information: Inducing High-Return Entrepreneurs to Move

I have documented that entrepreneurs taking up the moving subsidy, or CCT, when it is offered with information differ in their baseline characteristics from those who take up the moving subsidy without information. I have also shown that the entrepreneurs with characteristics resembling those who take up the subsidy when it is offered with information have differentially higher profit as an effect of receiving the UCT+Info. Did the presence of information serve to induce these high-return entrepreneurs to select into moving when they would not have otherwise?

To address this, I test whether “likely CCT only compliers” experience differentially positive effects of the CCT or UCT. If they do, then the role of information cannot be to induce uniquely high-return entrepreneurs to take up the subsidy. If they do not, then information may serve to induce high-return entrepreneurs to self-select into moving when they would not have otherwise. The results of this analysis are in Table 1.9. In contrast to likely CCT+Info compliers, likely CCT only compliers are not differentially likely to realize higher revenue or higher profit in response to the CCT or UCT, with or without information. In fact, likely CCT only compliers do not have a statistically significant differential response to the CCT or UCT (with or without information) on any dimension. This is consistent with the presence of information serving in a selection role, where it induces high-return entrepreneurs to select into moving. Without information, the subgroup that selects into taking up the subsidy and moving is no more likely than others to move into higher profit parishes, realize higher revenue, or realize higher profit.

1.6.3 Role of Liquidity: Insurance for Risky Moves

What was the role of the subsidy and the unconditional transfer in relieving liquidity constraints that enabled entrepreneurs to realize more than 50% higher profit, but only during the intervention? Figure 1.5 reveals that while those in the CCT+Info arm are more likely than the control group to realize profit in the upper tail, they are no less likely than the control to realize profit in the lower tail of the distribution; that is, it is not clear that the CCT+Info treatment first order stochastically dominates the control. In Figure 1.6, I focus on profit realizations by those who moved to parishes with higher average profits (across all treatment groups) relative to those who did not. While this is solely correlational, it suggests that moving to higher profit places may not dominate the alternative. This may indicate that moving is risky, and the subsidy served to insure risky moves during the intervention. In Table 1.10, I test the extent to which changes in profit pass through to consumption, following [Townsend \(1994\)](#). I find that the correlation between fluctuations in profit and fluctuations in consumption is differentially lower for those offered the CCT or the UCT, relative to those who do not receive either liquidity intervention. This is consistent with the CCT and UCT serving to insure riskiness in moving.

Once the intervention ends, entrepreneurs may be able to save profit gains from the intervention period to self-insure in the short run. Given a plausible degree of risk aversion and probability of future adverse events, their ability to self-insure may dissipate in the long run. This is consistent with the pattern of results: entrepreneurs taking up the subsidy when it was offered with information, as well as the high-return “likely CCT+Info complier” subgroup offered either the UCT with information, realize higher profits during the intervention and sustain revenue gains but imprecise profit gains in the week after the intervention. A month later, all effects on movement, revenue, and profit have dissipated.

1.7 Model Estimation and Counterfactual Simulations

I use the model, set up in Section 1.3, to conduct simulations, which quantify the aggregate impacts of relieving constraints on microentrepreneur location, and to establish the features that can replicate the empirical result of a complementarity between liquidity and information. I find that a model with limited information, fixed costs of moving, and two types of entrepreneurs, one with high returns of moving and another with low returns of moving, can rationalize the empirical results.

1.7.1 Assigned Parameters

I estimate the model for vendors, i.e., sector $k = \text{retail}$, who compose the majority of the sample. For simplicity, I will drop the k subscript from here forward, as all estimation and simulation is within sector. I assume that entrepreneurs face a liquidity threshold of $\bar{a} = 7,000$ UGX, above which they can optimally choose their location. This threshold equals the 95th percentile amount of money that entrepreneurs report needing to move to a more profitable location in the baseline survey and is equivalent in value to the cash transfer in the experiment.

Recalling the initial setup in equation (1.2) in Section 1.3, entrepreneur j with incumbent business location n and sufficient liquid capital, $a_j \geq 7000$, chooses a destination location (city parish) d to maximize utility,

$$V(j, n) = \begin{cases} \max_d \frac{A_d z_{jd}}{c_{nd} \nu_d} \ln(\delta_{jd} Y_d), & \text{if } a_j \geq 7000 \\ \frac{A_n z_{jn}}{\nu_n} \ln(\delta_{jn} Y_n), & \text{otherwise} \end{cases} \quad (1.7)$$

where A_d , Y_d , and ν_d are the amenities, income, and coefficient of variation common to parish d , respectively, c_{nd} is the iceberg cost of moving from location n to location d , $z_{jd} \sim \text{Frechet}(\theta)$ is an i.i.d. preference shock, and $\delta_{jd} \sim \text{Log-normal}(\lambda_d, \sigma_d^2)$ is an i.i.d. knowledge shock such that:

$$\mathbb{E}_j[\delta_{jd} Y_d] = \int \delta_{jd} Y_d dF_j(\delta_{jd}) = \exp(\lambda_d + \frac{\sigma_d^2}{2}) Y_d = \tilde{Y}_d \quad (1.8)$$

The entrepreneur's problem yields the number of entrepreneurs in location d , N_d , and the resulting π_d , the fraction of entrepreneurs locating in d . π_d equals the probability that an entrepreneur has greater indirect utility in location d than in any other location. Following standard results ([Eaton and Kortum, 2002](#)):

$$\pi_d = \frac{(A_d \bar{c}_d^{-1} \nu_d^{-1} \ln(\tilde{Y}_d))^\theta}{\sum_{x \in D} (A_x \bar{c}_x^{-1} \nu_x^{-1} \ln(\tilde{Y}_x))^\theta} \quad (1.9)$$

where $\bar{c}_d = c_{nd} \pi_n$, the average cost of locating in d . I assign ν_d and c_{nd} from the values observed in baseline data. To solve for entrepreneurs' location decisions in general equilibrium, I endogenize parish profits, Y_d , as in equation (1.5). Taking parish profits as exogenous, i.e., fixed at their baseline level, yields entrepreneurs' location decisions in partial equilibrium.

1.7.2 Estimated Parameters

I use exogenous variation in business location to estimate θ , the elasticity of willingness to locate in a particular place with respect to the net benefits of locating there, V_d . I use responses to questions about knowledge of profit variation throughout the city in the baseline survey to estimate the parameters of the knowledge shock, λ_d and σ_d . I use the location realizations of entrepreneurs in the UCT+Information group, for whom liquidity and information constraints were relieved, to estimate amenities, A_d .

Estimating the Fréchet Shape Parameter, θ . Using experimental variation from the offer of conditional cash transfers, I estimate θ as the elasticity of the likelihood of locating in a destination, π_d , with respect to the net benefits of locating there, V_d , as follows,

$$\ln \pi_d = \theta \ln \Omega_d \quad (1.10)$$

where Ω_d is the sum of the distance-normalized value of all CCTs offered three kilometers or more from location d (i.e., offered to those for whom the CCT increased the value of locating in d). The distance-normalized value of each CCT is 7,000 UGX divided by the kilometers between destination d and the pre-intervention location of the business owner to whom the CCT was offered, thus weighting CCT offers in nearby locations more heavily. This reflects the exogenous increase in the value of moving created by the offer of a CCT (normalized by distance, which is orthogonal to the CCT offer). I estimate $\theta=1.4$, a relatively low elasticity that implies a large role for idiosyncratic preferences. This estimate of θ is lower than the values of approximately 2.5 to 4 estimated in other low-income cities by [Franklin et al. \(2024\)](#) and [Tsivanidis \(2023\)](#). However, this is the first paper, to my knowledge, to use experimental variation to estimate willingness to move business locations by entrepreneurs, rather than workers. The additional production considerations faced by entrepreneurs—for example, sensitivity to profit variability not internalized by workers—could plausibly reduce willingness to move in response to a change in the value of doing so.

Estimating the Parameters of the Knowledge Shock, λ_d and σ_d . In the baseline survey, all entrepreneurs were asked a series of questions to assess their knowledge about profits in five randomly selected parishes near the parish in which the entrepreneur locates.²⁵ The question asked: “Do you think a business like yours would make more or less in parish X?” I use survey data to assess responses to this question by comparing mean sector-specific profits in parish X to those in the respondent’s own parish at baseline.²⁶ Using this scoring mechanism, I code correct responses as $\text{misperception}_d = 0$ to reflect

²⁵Five additional questions of the same format asked the entrepreneur about their perceptions of profits in far away parishes (outside their geographic division, an administrative unit that is less fine than the parish), but average responses to questions about far away parishes are on par with random guessing, overall and for any subgroups of interest. Thus, only responses to questions about nearby parishes are used to estimate the parameters of the knowledge shock.

²⁶Results are similar if profit rankings are assessed after conditioning on other individual entrepreneur characteristics in

non-distorted perceptions of profits in parish d . If the respondent overestimated profits in parish d , then $\text{misperception}_d = 1$, and if the respondent underestimated profits, then $\text{misperception}_d = -1$. I estimate λ_d as the mean value of misperception_d among the respondents asked about profits in parish d . If entrepreneurs on average underestimate profits in parish d , then $\lambda_d < 0$, while $\lambda_d > 0$ for parishes where entrepreneurs on average overestimate profits. I estimate σ_d as the bootstrapped standard error when the mean is estimated over 1000 samples of size n , drawn with replacement from the original sample of n respondents asked about profits in parish d . This reflects variability in perceptions at the parish level.

Estimating Amenities, A_d . I solve for amenities using the business locations realized during the intervention by those offered the UCT and information (i.e., those for whom liquidity and information constraints were relieved). Specifically, I solve for amenities in equation 1.9, where π_d is the fraction of entrepreneurs of sector k offered the UCT and information who then choose to locate in parish d and $\tilde{Y}_d = Y_d$, such that expected profits equal true profits, to reflect the easing of information constraints. Those offered the UCT and information may locate anywhere they like—unlike those offered the CCT, where receipt of the subsidy is conditional on location choice—and their location choices are realized after liquidity and information constraints have been eased. Thus, the location realizations of those in the UCT and information group reflect amenity value under minimal distortion by other frictions.

Estimated amenity values, in Figure 1.7, reflect utility cost. As an example, locating in a parish with amenity value of 0.86 offers 14% lower welfare than locating in a parish with amenity value of 1.00, all else equal. The estimation results reveal a distribution with a fat left tail, indicating that a large fraction of city parishes have relatively low amenities. I find that higher amenity value parishes are associated with more women entrepreneurs, more educated entrepreneurs, and more experienced entrepreneurs, groups who may have a greater idiosyncratic preference for high-amenity places. Qualitative discussions reveal that factors like tree cover and congestion may be important drivers of amenity value.

1.7.3 Model Simulations

I first simulate the spatial distribution of entrepreneurs under baseline conditions, where entrepreneurs face information frictions and fixed costs. In the baseline scenario, expected profit is distorted by an i.i.d. knowledge shock (with mean λ_d and variance σ_d^2), and any entrepreneur with savings less than 7,000 UGX (approximately \$2 and equivalent in value to the subsidy and unconditional transfer disbursed in the experiment) cannot move. This latter condition binds for 37% sample. The model fits the data reasonably well: the observed spatial distribution of entrepreneurs in the control group explains 50% of the variation in the model-simulated spatial distribution of entrepreneurs under baseline conditions, as shown in Figure 1.8.

addition to sector.

I then simulate entrepreneur locations under scenarios which provide correct information and liquidity such that fixed costs do not bind:

1. Corrected information: Expected profits are equivalent to true profits, i.e., $\tilde{Y}_d = Y_d$
2. Fixed costs do not bind: All entrepreneurs can move, as though they all have liquidity $> 7,000$ UGX
3. The combination of (1) and (2)
4. The combination of (1) and (2), but where a group of high-return entrepreneurs realizes differentially higher returns in moving
5. Simulation (4) under improved amenities across the city, where below-median amenity values are winsorized to the median

In scenario (4), entrepreneurs have exogenous type $T \in \{H, L\}$, where they have either high returns of moving (denoted by H) or low returns of moving (denoted by L). The optimization problem of “high” types now features idiosyncratic but on average higher returns, $r_j \sim \text{Log-normal}(0.52, 0.16^2)$, to be realized in any destination where a high-return entrepreneur might move.²⁷ This simulation attempts to mimic the empirical finding that high-return entrepreneurs can be induced to move and realize large gains from doing so, but only in response to both information and liquidity. Given correct knowledge of profits, sufficient liquidity such that fixed costs do not bind, and differentially higher returns realized by a subset of entrepreneurs, the entrepreneur’s problem in scenario (4) is as follows:

$$V(j, k, n, T) = \begin{cases} \max_d \frac{A_{kd} z_{jkd}}{c_{nd} \nu_{kd}} \ln(r_j Y_{kd}), & \text{if } T = H \\ \max_d \frac{A_{kd} z_{jkd}}{c_{nd} \nu_{kd}} \ln(Y_{kd}), & \text{if } T = L \end{cases} \quad (1.11)$$

These counterfactual scenarios intend to simulate the role of alleviating liquidity and information frictions and assess which features of the model are necessary to match empirical results. The simulation results appear in Table 1.11. The first two columns of Table 1.11 depict aggregate income and welfare under various counterfactual scenarios when parish profits are exogenous (partial equilibrium): they are fixed at their baseline level. The final two columns depict aggregate income and welfare when parish profits are endogenous (general equilibrium), responding to firm entry and exit across the city, per equation (1.5).

Providing information alone (simulation 1) offers no gain to either aggregate income or welfare. In general, entrepreneurs tend to overestimate profits in higher welfare places (higher expected profits, for

²⁷The mean, 0.52, is estimated as the percent increase in the fraction of expected parish-sector profits realized by “likely CCT+Info compliers,” those who empirically earn higher returns from moving, per the discussion in Section 1.6.1. The standard deviation, 0.16, is the bootstrapped standard error around the point estimate.

example, are associated with higher true profits, as in Figure 1.9a), so providing corrected information may not change the outcome of their location decision. When it does, entrepreneurs trade off higher income against amenity value, as amenities are negatively associated with true profit (Figure 1.9b). The coefficient of variability is less negatively associated with expected profits (Figure 1.9c) than with true profits (Figure 1.9d), again suggesting that entrepreneurs tend to overestimate profits in places which are relatively more desirable on the variability dimension. Thus, correcting incorrect notions of expected profit would drive entrepreneurs into more profitable parishes, all else equal, but trades off with lower amenity value and higher variability that was suppressed by the entrepreneur’s prior (incorrect) notions of expected profits.

Providing liquidity alone (simulation 2) increases aggregate welfare by 8% with no gains to aggregate income when parish profits are exogenous. This indicates that even when liquidity constraints on the ability to move are relieved, entrepreneurs face greater utility gains from moving into places with better amenities, lower variability, lower moving costs, and for which they have a higher idiosyncratic preference than into higher profit places. When parish profits are endogenous, there is no change in either aggregate income or welfare, reflecting saturation in welfare-maximizing location as profits decline in response to new entrepreneurs moving into a place. The simple additive combination of corrected information and liquidity (simulation 3) maintains this result, consistent with information alone offering no aggregate gains.

However, relieving information and liquidity constraints together can induce substantial income and welfare gains when their combination unlocks a complementarity that induces high-return entrepreneurs to move, as in simulation 4. In this simulation, 55% of entrepreneurs are a “high-return” type, where they have differentially higher returns of moving. They realize these higher returns only in response to the combination of information and liquidity. The differentially higher return means that they now face greater utility gains from moving into higher profit places than before, resulting in a willingness to accept lower amenity values, higher variability, and greater moving costs, as well as a willingness to move into places for which they have a lower idiosyncratic preference. In this simulation, aggregate income increases by 39% and aggregate welfare increases by 10%, if parish profits are fixed at their baseline levels. When parish profits decline as additional entrepreneurs move into a place, aggregate income increases by 37% but the aggregate welfare gain declines to just over 1%. This reflects congestion in high-welfare locations and indicates that profit gains are traded off against other components of the entrepreneur’s objective function. Nonetheless, this result underscores the spatial misallocation of entrepreneurs: even if profits decline as more entrepreneurs move into a place, it is possible to increase aggregate income without reducing welfare.

In simulation 5, I show that improving urban amenities in low-amenity places reduces the welfare drop-off from partial equilibrium to general equilibrium that is seen in simulation 4. Simulation 5 maintains the large aggregate income gains of simulation 4 while increasing aggregate welfare by 9% in general equilibrium,

only a 2.5 percentage point reduction from partial equilibrium. This emphasizes that low amenity values across the city dampen welfare gains in simulation 4, as entrepreneurs must trade off higher profits against amenities.

In sum, a model with fixed monetary costs of moving, information frictions, and heterogeneous returns to moving can rationalize the empirical results. I find large aggregate income gains of relieving liquidity and information frictions at scale. Welfare only increases marginally, due to low amenities in high-profit places.

1.8 Conclusion

I document large spatial variation in profits earned by observationally similar entrepreneurs within Kampala and then use an experiment and structural model to show that spatial profit gaps reflect, at least in part, spatial misallocation of entrepreneurs due to simultaneous financial and information frictions. When liquidity and information constraints are relieved together, entrepreneurs move into city parishes with higher average profits and realize higher profits of their own. Neither liquidity nor information alone is sufficient, indicating a complementarity between the two in overcoming the frictions that otherwise restrict entrepreneurs' location decisions. Entrepreneurs do not, however, sustain impacts on their location, revenues, or profits after the intervention ends. Impacts fade out gradually within a month after the liquidity interventions cease. These results are consistent with liquidity providing insurance for profitable but risky moves during the intervention and risk-averse entrepreneurs gradually losing the ability to self-insure as they face adverse events over time. The role of information is consistent with a selection tool: the presence of information changes the characteristics of the entrepreneurs who move locations, and these entrepreneurs realize differentially higher returns to moving and to cash in general.

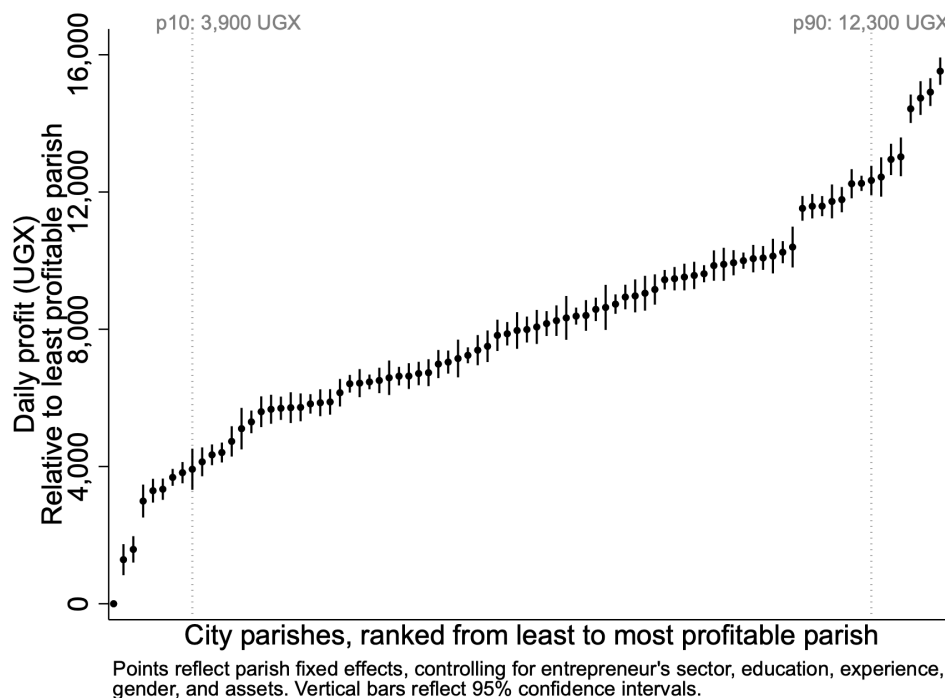
The complementary nature of liquidity and information frictions in constraining firm location decisions, and the possibility of using information as a screening tool, are useful insights for policy. Existing work finds low average effects of programs that relieve liquidity constraints, such as microcredit, for microentrepreneurs. Perhaps other contexts and investment decisions also require liquidity together with information to unlock benefits. Information provision as a possible tool to screen for high-return entrepreneurs is an attractive possibility in its own right, since identifying high growth potential microentrepreneurs has proven difficult. In addition, information provision tends to be inexpensive relative to other development programs, and when information can be used to induce take up by those with the highest returns, additional savings arise from the more efficient allocation of resources.

I show that a spatial equilibrium model that features fixed costs of moving, limited knowledge about profit variability, and a complementarity between relieving liquidity and information constraints that leads

high-return entrepreneurs to move can replicate the empirical results. I estimate large aggregate income gains of relieving liquidity and information constraints simultaneously, demonstrating spatial misallocation of microentrepreneurs.

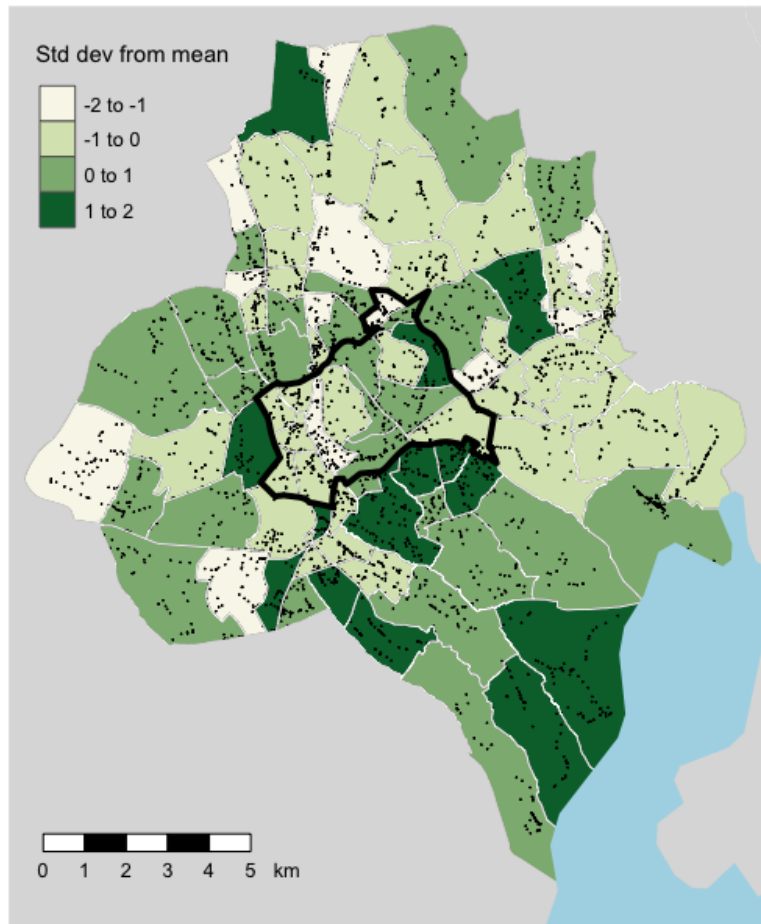
1.9 Figures

Figure 1.1: Between-Parish Profit Variation for Observationally Similar Microenterprises



Notes. This figure graphs the profit in Ugandan shillings (UGX) associated with locating in each city parish, relative to the least lucrative parish. The circular markers correspond to parish fixed effects (where the bars are 95% confidence intervals) from a regression that predicts profit with parish fixed effects, conditional on the business owner's sector, education, experience (in years and years squared), and asset stock. The city parish is the parish in which the business was located prior to the experiment.

Figure 1.2: Parish Profits at Baseline



Notes. Map depicts average parish profits, in terms of standard deviations above or below the mean. Dots reflect baseline survey locations, and the bold black outline is the central business district.

Figure 1.3: Example of Information Treatment Allocation



Information card personalized for: FID:

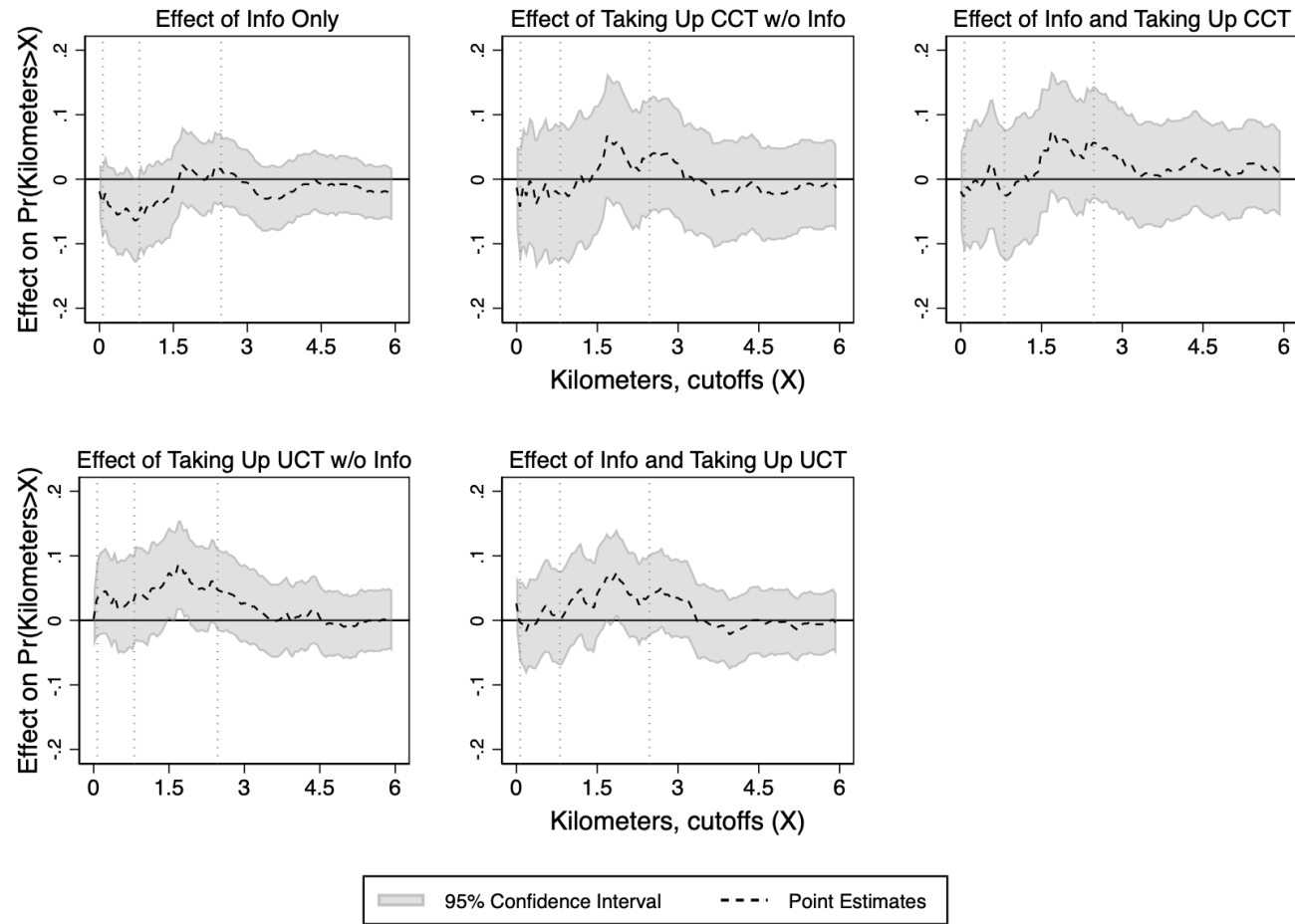
Businesses like yours report higher profits in other places. For example:

- Businesses locating in LUNGUJJA parish (the area around Uganda Youth Football Association) typically earn 2,200 UGX more than businesses in your parish.
- Businesses locating in KAMWOKYA I parish (the area around City oil) typically earn 9,300 UGX more than businesses in your parish.
- Businesses locating in BWAISE III parish (the area around Babu plaza) typically earn 6,900 UGX more than businesses in your parish.
- Businesses locating in BUGOLOBI parish (the area around Bugolobi flats) typically earn 5,900 UGX more than businesses in your parish.

*These numbers are estimates and do not represent a guaranteed higher profit in another place.

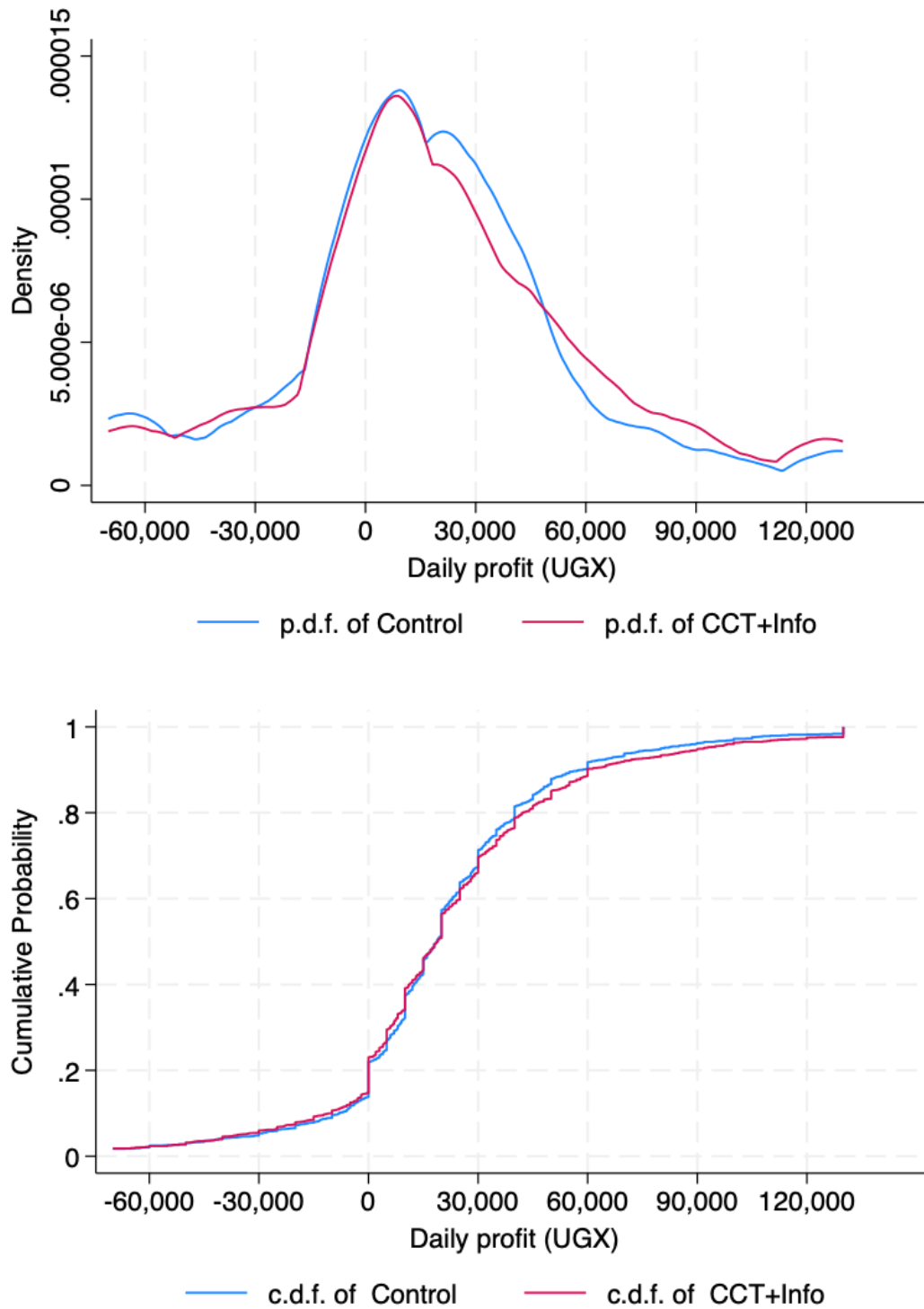
If you have any questions, please contact our study team at: 0702 935 679 or 0773 450 160

Figure 1.4: Distribution Regressions: Effect of Treatment on Probability of Locating $>X$ Kilometers from Pre-Intervention Business Location



Notes: Vertical dashed lines reflect the 25th, 50th, and 75th percentile of the outcome. Plots depict treatment on the treated (TOT) effects in distribution regressions, where take-up is instrumented with indicators for random assignment to the CCT, CCT with information, UCT, and UCT with information arms. Standard errors are heteroskedasticity-robust and all estimating equations include controls selected via double-selection (DS) lasso, as well as non-penalized strata FEs.

Figure 1.5: Profit Distributions During the Intervention



Note: Control mean = 17,000 UGX

Figure 1.6: Profit Distributions: Moving to a Higher Profit Parish Versus Not

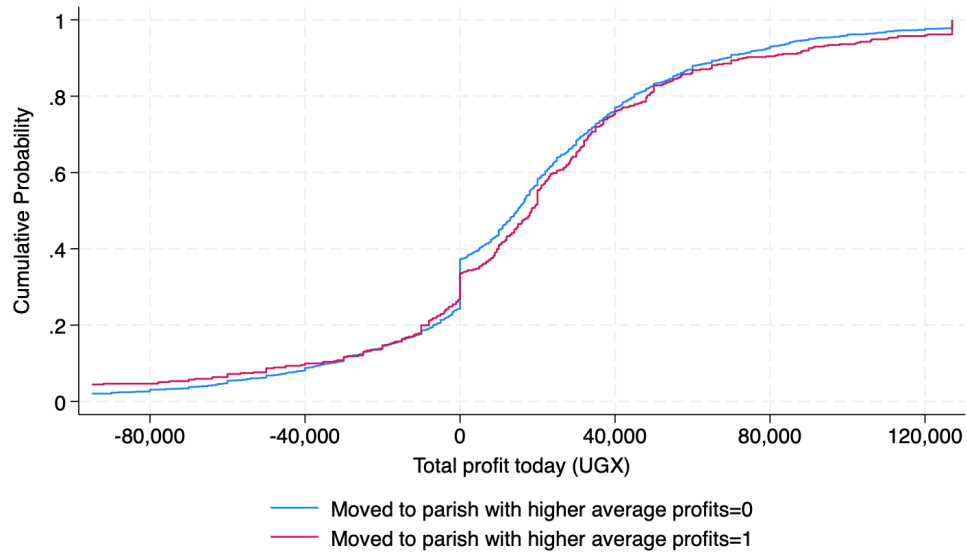
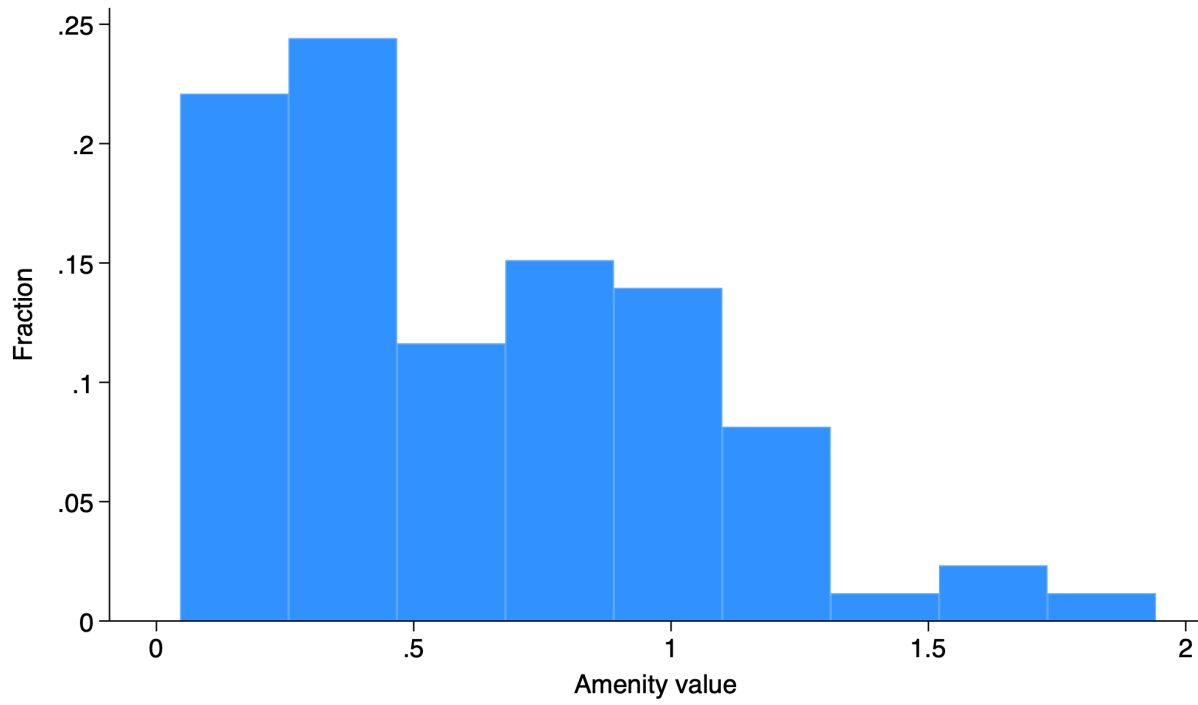


Figure 1.7: Distribution of Parish Amenities



Notes. This figure shows the distribution of parish-level amenities estimated in the model. Values are in terms of utility cost, e.g., a parish with amenity value of 0.50 reduces welfare by 50% relative to a parish with amenity value of 1.00, all else equal.

Figure 1.8: Model Fit

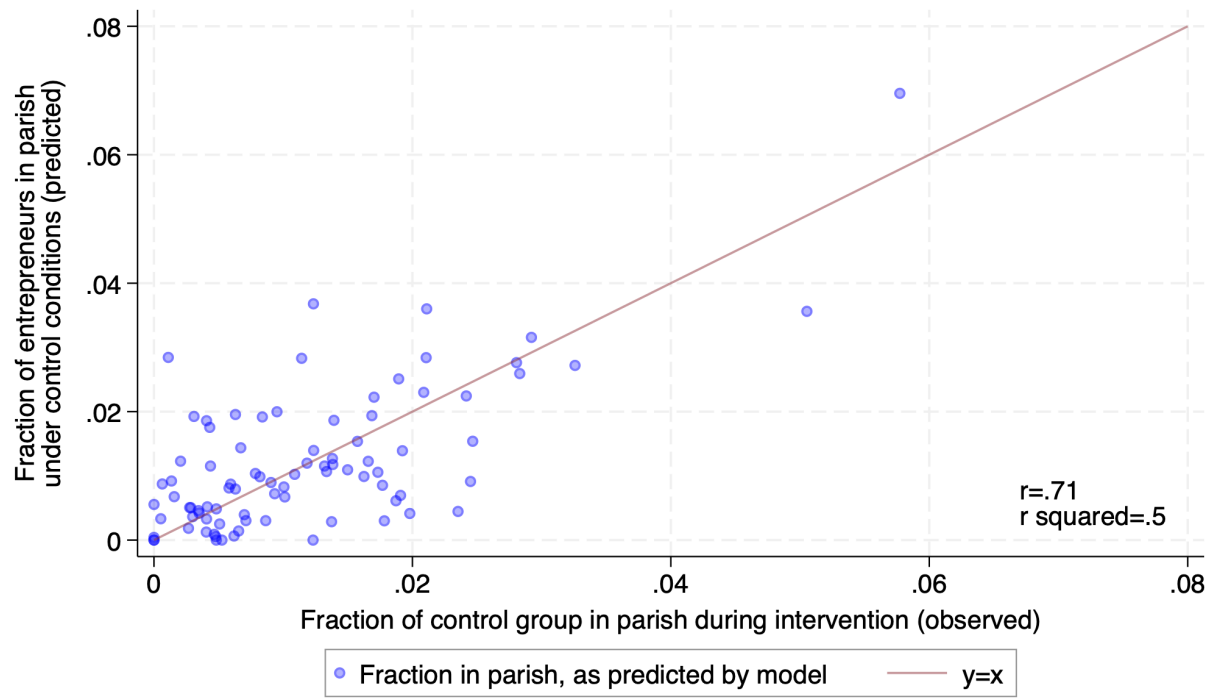
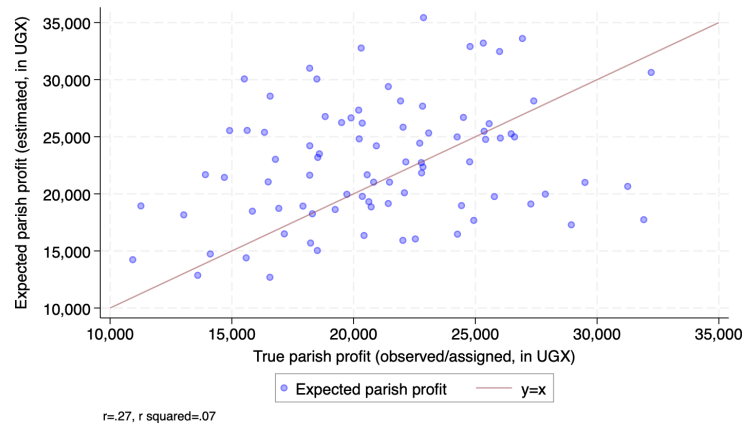
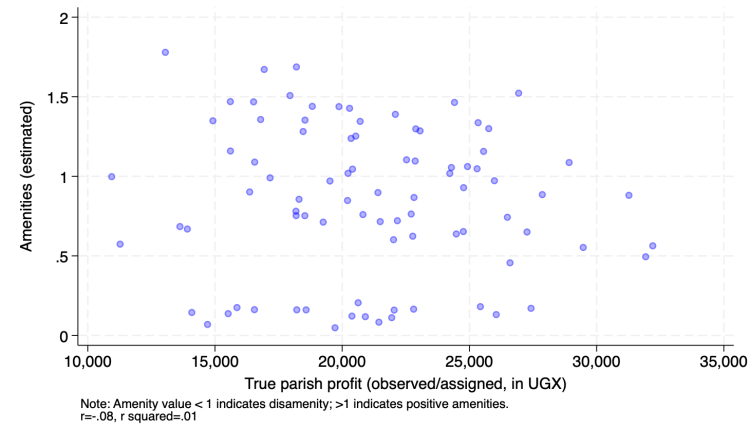


Figure 1.9: Correlations Between Estimated and Assigned Parameters

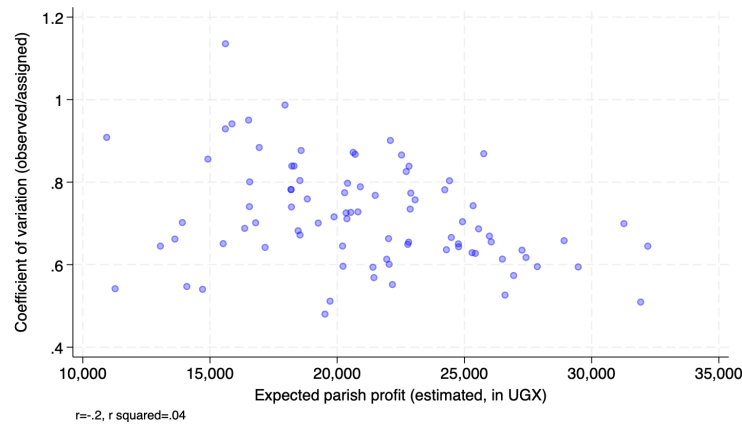
(a) Expected Profit v. True Profit



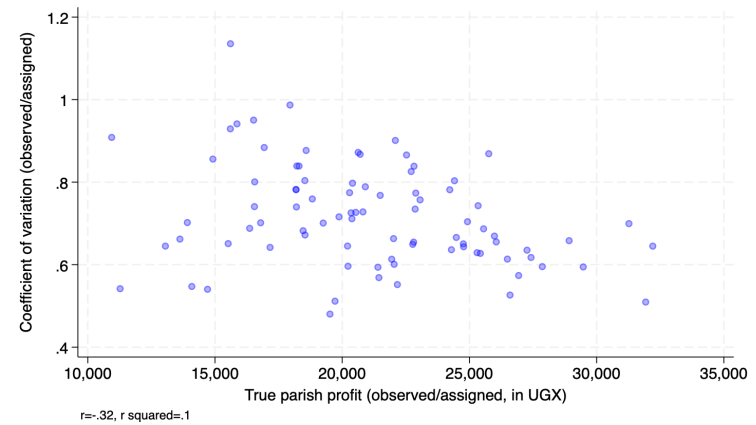
(b) Amenities v. True Profit



(c) Coefficient of Variation v. Expected Profit



(d) Coefficient of Variation v. True Profit



1.10 Tables

Table 1.1: Baseline Sample Characteristics

	All	Vendors	Bodas
<i>Business Location</i>			
Commute (min) from home to business location	34	37	25
Formally registered business with authority=1	.068	.025	.17
Primarily relies on repeat customers=1	.22	.25	.15
Thinks could be more profitable elsewhere=1	.66	.67	.62
Primary constraint to moving is cash=1 [†]	.62	.67	.49
Anticipated daily moving costs (UGX) to more profitable location [†]	3,433	3,105	4,285
Fraction correct on quiz about spatial profit dispersion [§]	.49	.48	.49
<i>Business Owner Demographics</i>			
Female=1	.25	.35	0
Business owner age	30	30	31
Years running any business	4.8	4.5	5.5
Completed primary school=1	.68	.64	.78
Completed secondary school=1	.15	.13	.2
<i>Business Finances</i>			
Daily revenue (UGX)	60,223	65,437	47,716
Daily costs (UGX)	34,239	40,655	18,852
Daily profit (UGX)	22,820	21,177	26,761
Asset stock (UGX)	1,011,321	187,800	2,986,427
<i>Type of Costs Incurred</i>			
Pays for labor=1	.0043	.0061	0
Pays for inventory=1	.66	.94	0
Pays for fuel or transport=1	.45	.23	.98
Pays any sort of rent to use space=1	.066	.09	.01
Own hours worked per week	67	62	78
<i>Personal Finances</i>			
Has any savings=1	.61	.6	.64
Total savings (UGX)	280,107	212,242	442,871
Has any credit=1	.22	.19	.31
Total credit (UGX)	201,123	63,246	531,802
Observations	3,745	2,643	1,102

Notes: 1 USD $\approx$ 3,700 UGX at time of baseline survey (November 2022). [†]These are the only dimensions where the mean is computed over a subsample of the total reported in the bottom row: only those who responded positively to the prior question on whether the entrepreneur believed she could be more profitable elsewhere were asked the questions on the primary constraint to moving and anticipated moving costs. All other variables reflect responses from the entire sample, of the size noted in the last row. [§]A score of 50% on the quiz is the expected score under random guessing.

Table 1.2: Experimental Design

	No Information	Information	
No Subsidy/Transfer	<i>Pure Control</i> n=499	<i>Information Only</i> n=498	997
Moving Subsidy a.k.a. Conditional Cash Transfer (CCT)	<i>CCT Only</i> n=585	<i>CCT + Info</i> n=586	1,171
Unconditional Cash Transfer (UCT)	<i>UCT Only</i> n=354	<i>UCT + Info</i> n=361	715
	1,438	1,445	2,883

Table 1.3: Balance Table

Variable	(1) Control Mean/SE	(2) Info Only Mean/SE	(3) CCT Only Mean/SE	(4) CCT+Info Mean/SE	(5) UCT Only Mean/SE	(6) UCT+Info Mean/SE	(7) Total Mean/SE	F-test for joint orthogonality
Daily profit (UGX)	21,377 (626)	21,499 (617)	21,931 (582)	20,817 (605)	22,421 (744)	21,582 (758)	21,551 (264)	.64
Daily revenue (UGX)	60,632 (1,918)	62,108 (2,056)	60,992 (1,792)	58,363 (1,736)	60,282 (2,225)	60,687 (2,397)	60,463 (811)	.79
Daily costs (UGX)	35,597 (1,528)	37,570 (1,647)	35,354 (1,425)	33,915 (1,316)	33,126 (1,611)	36,018 (1,769)	35,296 (628)	.26
Incurs transport costs=1	.44 (.022)	.41 (.022)	.4 (.02)	.39 (.02)	.39 (.026)	.44 (.026)	.41 (.0092)	.052*
Inventory value (UGX)	83,150 (5,075)	89,627 (5,520)	79,871 (4,615)	82,750 (4,937)	78,234 (5,854)	88,573 (6,483)	83,598 (2,177)	.39
Total bus. asset value (UGX)	894,124 (61,682)	928,042 (61,029)	907,286 (56,997)	889,718 (56,690)	901,381 (73,265)	879,512 (70,661)	900,820 (25,491)	.94
Total savings value (UGX)	199,783 (14,657)	239,753 (17,922)	186,252 (13,421)	253,299 (17,414)	197,408 (18,216)	209,064 (19,805)	215,690 (6,875)	.027**
Has formal savings=1	.45 (.022)	.43 (.022)	.42 (.02)	.46 (.021)	.41 (.026)	.45 (.026)	.44 (.0092)	.57
Has any credit=1	.23 (.019)	.23 (.019)	.22 (.017)	.2 (.017)	.23 (.022)	.21 (.021)	.22 (.0077)	.78
Hours worked per week	67 (.93)	66 (.96)	66 (.85)	68 (.9)	66 (1.2)	67 (1.1)	67 (.4)	.42
Num competitors in eyesight	4 (.35)	3.8 (.33)	3.7 (.3)	3.3 (.23)	3.3 (.32)	3.8 (.42)	3.6 (.13)	.15
Age of entrepreneur	32 (.41)	31 (.4)	31 (.39)	32 (.36)	31 (.5)	32 (.51)	31 (.17)	.8
Years as a business owner	5.5 (.23)	5.3 (.25)	5.2 (.22)	5.2 (.23)	5.3 (.31)	5.6 (.29)	5.3 (.1)	.92
Completed primary school=1	.64 (.022)	.69 (.021)	.68 (.019)	.68 (.019)	.68 (.025)	.67 (.025)	.67 (.0088)	.57
Commute to work (min)	33 (1.1)	32 (1.1)	33 (1.1)	32 (1.1)	33 (1.4)	32 (1.4)	32 (.48)	.96
Walks to work=1	.61 (.022)	.6 (.022)	.61 (.02)	.59 (.02)	.62 (.026)	.62 (.026)	.61 (.0091)	.92
Desire to expand=1	.95 (.01)	.94 (.011)	.93 (.01)	.92 (.011)	.92 (.014)	.91 (.015)	.93 (.0048)	.19
Patient=1	.2 (.018)	.24 (.019)	.26 (.018)	.26 (.018)	.21 (.022)	.27 (.023)	.24 (.008)	.027**
Time-inconsistent=1	.11 (.014)	.13 (.015)	.13 (.014)	.13 (.014)	.099 (.016)	.12 (.017)	.12 (.0061)	.66
N	499	498	585	586	354	361	2883	

Notes: Standard errors are heteroskedasticity-robust. Strata fixed effects are included in all estimation regressions. Asterisks denote statistical significance as follows: ***1%, **5%, *10%.

Table 1.4: Lasso-selected characteristics of compliers

	CCT+Info	CCT Only
Minutes from home to business location	.041	
Percent score on info quiz about nearby parishes (50%=guessing)	.030	
Total daily fuel costs (UGX)	.028	
In below median profit parish=1	.021	
Number of competitors	.021	
Has credit=1	.020	
Profit per hour	-.017	
Female=1	-.014	
Completed secondary school=1	.011	
Incurs transit costs=1	.009	.021
Completed primary school=1	.008	
Wants to expand business=1	-.005	
Years working on this business	.003	
More transient=1	-.002	
Total asset value		.028
Total savings		.013
Years of experience as a business owner		.001

Notes: Table displays standardized penalized coefficients selected by lasso, in descending order. Penalty parameter selected using cross-validation with 10 folds.

Table 1.5: Effect on Location During the Intervention

	(1) Left Kampala=1	(2) Same parish as baseline=1	(3) Moved to lower profit parish=1	(4) Moved to higher profit parish=1	(5) Expected (avg) profit in current parish
<i>Panel A. Intent-to-Treat (ITT):</i>					
Info=1	-0.02 (0.02)	-0.01 (0.03)	0.03 (0.03)	-0.002 (0.03)	-110.8 (287.8)
CCT=1	-0.02 (0.02)	-0.02 (0.03)	0.04* (0.03)	-0.007 (0.02)	-310.0 (268.7)
Info=1 × CCT=1	0.02 (0.02)	0.02 (0.05)	-0.09** (0.04)	0.05 (0.03)	853.3** (377.8)
UCT=1	-0.01 (0.02)	-0.06* (0.04)	0.06* (0.03)	0.02 (0.03)	110.9 (358.9)
Info=1 × UCT=1	0.03 (0.03)	0.006 (0.05)	-0.06 (0.04)	0.02 (0.04)	265.1 (476.8)
Info+CCT+Info*CCT	-0.02	-0.01	-0.01	0.04	432.47
P-value: I+C+I*C=0	0.29	0.66	0.62	0.08	0.12
Info+UCT+Info*UCT	0.00	-0.07	0.03	0.03	265.22
P-value: I+U+I*U=0	0.93	0.06	0.26	0.26	0.40
Num. of DS Lasso Controls	0	0	2	1	2
<i>Panel B. Treatment-on-Treated (ToT):</i>					
Info=1	-0.02 (0.02)	-0.01 (0.03)	0.03 (0.03)	-0.002 (0.02)	-110.9 (285.0)
Took up CCT=1	-0.03 (0.03)	-0.03 (0.05)	0.07* (0.04)	-0.01 (0.04)	-507.4 (437.9)
Info=1 x Took up CCT=1	0.03 (0.04)	0.03 (0.07)	-0.1** (0.06)	0.08 (0.06)	1386.4** (610.8)
Took up UCT=1	-0.02 (0.02)	-0.06* (0.04)	0.06* (0.03)	0.02 (0.03)	114.7 (364.8)
Info=1 x Took up UCT=1	0.04 (0.03)	0.006 (0.05)	-0.06 (0.04)	0.02 (0.04)	278.5 (488.0)
Info+CCT+Info*CCT	-0.02	-0.01	-0.04	0.07	768.06
P-value: I+C+I*C=0	0.43	0.76	0.23	0.04	0.04
Info+UCT+Info*UCT	0.00	-0.07	0.03	0.03	282.23
P-value: I+U+I*U=0	0.90	0.05	0.27	0.24	0.38
Control Mean	0.08	0.54	0.19	0.19	21,946.10
Observations	2,410	2,410	2,410	2,410	2,410

Notes: Columns 1 - 4 compose mutually exclusive groups. Heteroskedasticity-robust standard errors in parentheses. Any non-binary (or otherwise bounded) outcomes are winsorized to the 2.5 and 97.5 percentiles, and any financial outcomes are in Ugandan shillings (UGX). Controls are selected via double-selection (DS) lasso, where lasso selects from among 28 pre-specified baseline controls, given the inclusion of strata FE's as non-penalized controls. ***1%, **5%, *10%.

Table 1.6: Effect on Financial Outcomes During the Intervention

	(1) Mobile money balance	(2) Daily revenue	(3) Daily costs	(4) Daily profit	(5) Work hours per day	(6) Profit per hour	(7) Fraction of exp. profit realized
<i>Panel A. Intent-to-Treat (ITT):</i>							
Info=1	-425 (1,400)	-1,880 (2,535)	703 (2,008)	-3,773 (2,745)	-.37* (.21)	-411 (282)	-.21 (.14)
CCT=1	3,220** (1,414)	3,232 (2,464)	2,458 (1,944)	-300 (2,537)	.12 (.2)	-47 (258)	-.012 (.14)
Info=1 × CCT=1	-1,663 (1,924)	3,958 (3,512)	-2,318 (2,698)	7,396** (3,698)	.34 (.29)	681* (375)	.32* (.19)
UCT=1	4,727*** (1,342)	-88 (2,839)	1,039 (2,153)	-841 (3,020)	.23 (.22)	-106 (303)	-.046 (.16)
Info=1 × UCT=1	1,620 (2,014)	6,145 (4,127)	-1,785 (3,032)	7,999* (4,348)	.47 (.31)	724* (438)	.5** (.23)
Info+CCT+Info*CCT P-value: I+C+I*C=0	1,132 0.39	5,309 0.03	842 0.65	3,323 0.20	.083 0.68	222 0.39	.1 0.44
Info+UCT+Info*UCT P-value: I+U+I*U=0	5,922 0.00	4,176 0.16	-43 0.98	3,385 0.27	.32 0.12	207 0.50	.24 0.14
Num. of DS Lasso Controls	0	2	1	0	1	0	0
<i>Panel B. Treatment-on-Treated (ToT):</i>							
Info=1	-423 (1,387)	-1,868 (2,509)	704 (1,988)	-3,769 (2,718)	-.37* (.21)	-411 (279)	-.21 (.14)
Took up CCT=1	5,295** (2,299)	5,324 (3,986)	4,041 (3,160)	-471 (4,134)	.2 (.32)	-76 (420)	-.019 (.22)
Info=1 x Took up CCT=1	-2,786 (3,103)	6,275 (5,653)	-3,823 (4,353)	11,945** (5,980)	.54 (.46)	1,100* (606)	.52* (.31)
Took up UCT=1	4,862*** (1,363)	-91 (2,884)	1,067 (2,188)	-843 (3,073)	.24 (.23)	-107 (308)	-.046 (.16)
Info=1 x Took up UCT=1	1,752 (2,061)	6,393 (4,227)	-1,851 (3,107)	8,311* (4,454)	.49 (.32)	752* (449)	.52** (.23)
Info+CCT+Info*CCT P-value: I+C+I*C=0	2,086 0.25	9,732 0.00	922 0.72	7,705 0.04	.36 0.19	613 0.09	.29 0.10
Info+UCT+Info*UCT P-value: I+U+I*U=0	6,191 0.00	4,434 0.14	-80 0.97	3,699 0.23	.35 0.09	234 0.45	.26 0.11
Control Mean	6,650	37,604	21,253	17,130	10	1,495	.77
Observations	2,410	2,410	2,410	2,410	2,410	2,410	2,410

Notes: Heteroskedasticity-robust standard errors in parentheses. Any non-binary (or otherwise bounded) outcomes are winsorized to the 2.5 and 97.5 percentiles, and any financial outcomes are in Ugandan shillings (UGX). Controls are selected via double-selection (DS) lasso, where lasso selects from among 28 pre-specified baseline controls, given the inclusion of strata FE's as non-penalized controls. ***1%, **5%, *10%.

Table 1.7: Effect on Location and Financial Outcomes After the Intervention

	Endline 1: A Week Post-Intervention			Endline 2: $\approx$ A Month Post-Intervention		
	(1) Expected (avg) profit in current parish	(2) Daily revenue	(3) Daily profit	(4) Expected (avg) profit in current parish	(5) Daily revenue	(6) Daily profit
<i>Panel A. Intent-to-Treat (ITT):</i>						
Info=1	-329 (294)	2,868 (2,596)	1,655 (2,509)	77 (123)	33 (2,349)	1,946 (2,256)
CCT=1	-622** (289)	2,931 (2,397)	-66 (2,329)	39 (120)	1,113 (2,234)	1,087 (2,273)
Info=1 $\times$ CCT=1	851** (392)	-760 (3,552)	1,671 (3,444)	35 (170)	286 (3,274)	-4,879 (3,196)
UCT=1	-399 (329)	-876 (2,534)	-3,089 (2,591)	46 (141)	-2,499 (2,456)	-2,929 (2,565)
Info=1 $\times$ UCT=1	747 (456)	2,956 (3,873)	5,600 (3,904)	-3.7 (191)	2,116 (3,565)	751 (3,584)
Info+CCT+Info*CCT	-100	5,038	3,260	151	1,432	-1,846
P-value: I+C+I*C=0	0.74	0.04	0.18	0.22	0.55	0.43
Info+UCT+Info*UCT	18	4,948	4,166	119	-350	-232
P-value: I+U+I*U=0	0.96	0.08	0.14	0.35	0.89	0.93
Num. of DS Lasso Controls	1	2	0	2	2	0
<i>Panel B. Treatment-on-Treated (ToT):</i>						
Info=1	-329 (291)	2,880 (2,567)	1,659 (2,483)	77 (122)	41 (2,327)	1,939 (2,236)
Took up CCT=1	-998** (462)	4,715 (3,802)	-98 (3,703)	66 (197)	1,841 (3,644)	1,765 (3,711)
Info=1 $\times$ Took up CCT=1	1,372** (629)	-1,218 (5,665)	2,701 (5,502)	58 (280)	474 (5,367)	-8,080 (5,255)
Took up UCT=1	-420 (343)	-920 (2,642)	-3,248 (2,706)	49 (147)	-2,624 (2,557)	-3,092 (2,673)
Info=1 $\times$ Took up UCT=1	798* (482)	3,159 (4,103)	5,966 (4,137)	-2.4 (203)	2,197 (3,802)	695 (3,819)
Info+CCT+Info*CCT	45	6,377	4,262	201	2,356	-4,376
P-value: I+C+I*C=0	0.91	0.07	0.22	0.25	0.49	0.19
Info+UCT+Info*UCT	49	5,119	4,377	123	-386	-457
P-value: I+U+I*U=0	0.89	0.08	0.14	0.36	0.89	0.87
Control Mean	22,222	40,692	21,466	22,025	45,804	22,926
Observations	2,500	2,500	2,500	2,596	2,596	2,596

Notes: Heteroskedasticity-robust standard errors in parentheses. Any non-binary (or otherwise bounded) outcomes are winsorized to the 2.5 and 97.5 percentiles, and any financial outcomes are in Ugandan shillings (UGX). Controls are selected via double-selection (DS) lasso, where lasso selects from among 28 pre-specified baseline controls, given the inclusion of strata FE's as non-penalized controls. ***1%, **5%, *10%.

Table 1.8: Heterogeneity in Treatment Effects on Profit During the Intervention: By Likelihood of Complying with the CCT+Info

	(1) Expected (avg) profit in current parish	(2) Daily revenue	(3) Daily costs	(4) Daily profit	(5) Frac. of exp. profit realized	(6) Work hours per day	(7) Profit per hour
Info Only $\times$ Likely CCT+Info complier=1	957 (589)	4,690 (5,236)	-3,462 (4,188)	8,453 (5,659)	.42 (.29)	.24 (.44)	1,028* (585)
CCT Only $\times$ Likely CCT+Info complier=1	457 (568)	1,792 (5,069)	-5,289 (4,035)	6,350 (5,272)	.3 (.28)	.66 (.41)	681 (535)
CCT+Info $\times$ Likely CCT+Info complier=1	454 (570)	12,312** (5,005)	-787 (3,874)	12,839** (5,347)	.69** (.27)	.39 (.4)	1,129** (530)
UCT Only $\times$ Likely CCT+Info complier=1	-253 (753)	6,614 (5,813)	99 (4,441)	4,353 (6,177)	.38 (.32)	-.13 (.45)	583 (622)
UCT+Info $\times$ Likely CCT+Info complier=1	201 (671)	10,119* (6,091)	-2,747 (4,440)	12,588** (6,272)	.65* (.34)	.11 (.43)	1,512** (629)
Info Only	-618 (485)	-5,100 (4,222)	2,592 (3,430)	-9,056* (4,621)	-.47* (.24)	-.52* (.32)	-1,046** (487)
CCT Only	-567 (460)	1,790 (4,034)	5,133 (3,314)	-4,086 (4,305)	-.19 (.23)	-.25 (.31)	-459 (442)
CCT+Info	185 (473)	-2,045 (3,786)	1,136 (3,098)	-4,015 (4,167)	-.29 (.22)	-.14 (.3)	-430 (415)
UCT Only	237 (613)	-3,981 (4,459)	971 (3,541)	-3,541 (4,832)	-.26 (.25)	.31 (.32)	-459 (498)
UCT+Info	153 (561)	-1,637 (4,780)	1,503 (3,526)	-3,845 (4,924)	-.13 (.26)	.27 (.33)	-659 (500)
Likely CCT+Info complier=1	-492 (520)	-660 (4,229)	2,864 (3,368)	-2,213 (4,470)	-.071 (.24)	-.08 (.32)	-319 (446)
Info+Info*Het	338	-410	-870	-603	-.048	-.28	-18
P-val: Info+Info*Het=0	0.32	0.89	0.71	0.85	0.77	0.34	0.96
CCT Only+CCT Only*Het	-111	3,582	-156	2,264	.1	.4	222
P-val: CCT Only+CCT Only*Het=0	0.73	0.24	0.95	0.45	0.53	0.12	0.46
(CCT+Info)+(CCT+Info)*Het	640	10,268	349	8,825	.4	.26	700
P-val: (CCT+Info)+(CCT+Info)*Het=0	0.05	0.00	0.88	0.01	0.02	0.33	0.03
UCT Only+UCT Only*Het	-16	2,632	1,070	812	.12	.18	124
P-val: UCT Only+UCT Only*Het=0	0.97	0.48	0.69	0.83	0.58	0.57	0.74
(UCT+Info)+(UCT+Info)*Het	354	8,483	-1,243	8,743	.52	.37	853
P-val: (UCT+Info)+(UCT+Info)*Het=0	0.33	0.02	0.64	0.02	0.01	0.16	0.02
Control Mean (Level)	21,946	37,604	21,253	17,130	.77	10	1,495
Observations	2,410	2,410	2,410	2,410	2,410	2,410	2,410

Notes: The table depicts mutually exclusive indicators of treatment assignment (i.e., these are intent-to-treat effects) interacted with an indicator for likelihood of complying with the CCT and information treatment (as predicted by baseline characteristics). “Het” is an abbreviation for the heterogeneity variable of interest. Heteroskedasticity-robust standard errors in parentheses. Any non-binary (or otherwise bounded) outcomes are winsorized to the 2.5 and 97.5 percentiles, and any financial outcomes are in Ugandan shillings (UGX). Likelihood of complying is predicted with lasso from baseline covariates. Controls are selected via double-selection (DS) lasso, where lasso selects from among 28 pre-specified baseline controls, given the inclusion of strata FE’s as non-penalized controls. ***1%, **5%, *10%.

Table 1.9: Heterogeneity in Treatment Effects on Profit During the Intervention: By Likelihood of Complying with the CCT Only

	(1) Expected (avg) profit in current parish	(2) Daily revenue	(3) Daily costs	(4) Daily profit	(5) Frac. of exp. profit realized	(6) Work hours per day	(7) Profit per hour
Info Only × Likely CCT Only complier=1	-29 (585)	4,466 (5,238)	-3,709 (4,260)	10,158* (5,694)	.51* (.3)	.5 (.44)	1,109* (584)
CCT Only × Likely CCT Only complier=1	120 (550)	655 (5,060)	-1,799 (4,003)	5,417 (5,244)	.18 (.28)	.65 (.4)	586 (533)
CCT+Info × Likely CCT Only complier=1	-668 (561)	5,760 (4,968)	-442 (3,863)	7,327 (5,337)	.26 (.27)	.56 (.4)	591 (528)
UCT Only × Likely CCT Only complier=1	-764 (738)	1,007 (5,792)	-505 (4,403)	2,749 (6,221)	.12 (.33)	.44 (.46)	387 (622)
UCT+Info × Likely CCT Only complier=1	-218 (652)	-2,085 (6,062)	-351 (4,391)	-994 (6,237)	-.3 (.34)	.32 (.42)	137 (623)
Info Only	-103 (445)	-4,649 (4,138)	2,481 (3,570)	-9,508** (4,561)	-.49** (.25)	-.64* (.33)	-1,040** (471)
CCT Only	-364 (429)	2,770 (4,024)	3,575 (3,223)	-3,473 (4,141)	-.12 (.23)	-.22 (.3)	-395 (423)
CCT+Info	797* (443)	1,597 (3,759)	928 (3,071)	-1,165 (4,150)	-.06 (.22)	-.22 (.3)	-155 (414)
UCT Only	536 (560)	-761 (4,551)	1,209 (3,520)	-2,424 (4,948)	-.11 (.27)	.0013 (.36)	-332 (494)
UCT+Info	380 (517)	5,386 (4,690)	360 (3,411)	3,625 (4,847)	.4 (.27)	.14 (.32)	1.0e+02 (486)
Likely CCT Only complier=1	234 (537)	1,104 (4,657)	8,880** (3,782)	-10,210** (5,005)	-.46* (.27)	-.62* (.34)	-1,118** (507)
Info+Info*Het	-132	-183	-1,228	650	.015	-.14	69
P-val: Info+Info*Het=0	0.73	0.95	0.60	0.85	0.93	0.62	0.84
CCT Only+CCT Only*Het	-245	3,425	1,776	1,944	.062	.43	191
P-val: CCT Only+CCT Only*Het=0	0.47	0.26	0.45	0.54	0.70	0.10	0.55
(CCT+Info)+(CCT+Info)*Het	129	7,356	486	6,162	.2	.34	436
P-val: (CCT+Info)+(CCT+Info)*Het=0	0.71	0.02	0.83	0.07	0.22	0.20	0.18
UCT Only+UCT Only*Het	-228	246	705	324	.0071	.44	55
P-val: UCT Only+UCT Only*Het=0	0.63	0.95	0.79	0.93	0.97	0.12	0.88
(UCT+Info)+(UCT+Info)*Het	161	3,300	9.1	2,630	.093	.45	237
P-val: (UCT+Info)+(UCT+Info)*Het=0	0.68	0.39	1.00	0.50	0.64	0.09	0.54
Control Mean (Level)	21,946	37,604	21,253	17,130	.77	10	1,495
Observations	2,410	2,410	2,410	2,410	2,410	2,410	2,410

Notes: The table depicts mutually exclusive indicators of treatment assignment (i.e., these are intent-to-treat effects) interacted with an indicator for likelihood of complying with the CCT only treatment (as predicted by baseline characteristics). “Het” is an abbreviation for the heterogeneity variable of interest. Heteroskedasticity-robust standard errors in parentheses. Any non-binary (or otherwise bounded) outcomes are winsorized to the 2.5 and 97.5 percentiles, and any financial outcomes are in Ugandan shillings (UGX). Likelihood of complying is predicted with lasso from baseline covariates. Controls are selected via double-selection (DS) lasso, where lasso selects from among 28 pre-specified baseline controls, given the inclusion of strata FE’s as non-penalized controls.

***1%, **5%, *10%.

Table 1.10: How Pass Through of Profit to Consumption Varies by Treatment Group (Panel)

	(1) Spending on irregular expenses in past week	(2) Spending on irregular expenses in past week	(3) Spending on irregular expenses in past week
Typical daily profit, past week	0.31 (0.35)	0.49 (0.53)	0.49 (0.53)
Daily profit x offered any cash (CCT or UCT)	-0.97** (0.43)	-1.08* (0.64)	
Daily profit x received info		-0.37 (0.69)	
Daily profit x offered cash and received info		0.23 (0.87)	
<i>Profit interacted with mutually exclusive treatment dummies:</i>			
Daily profit x Info only			-0.37 (0.69)
Daily profit x CCT only			-1.16* (0.69)
Daily profit x CCT+Info			-1.19* (0.72)
Daily profit x UCT only			-0.95 (0.81)
Daily profit x UCT+Info			-1.25 (0.76)
Total effect of (any) cash	-0.66		
P-value in F-test	0.01		
Total effect of cash only		-0.59	
P-value in F-test		0.11	
Total effect of info only		0.12	0.12
P-value in F-test		0.79	0.79
Total effect of cash + info		-0.72	
P-value in F-test		0.05	
Total effect of CCT only			-0.67
P-value in F-test			0.13
Total effect of CCT+Info			-0.70
P-value in F-test			0.15
Total effect of UCT only			-0.46
P-value in F-test			0.46
Total effect of UCT+Info			-0.76
P-value in F-test			0.17
Firm FE's	Yes	Yes	Yes
Outcome mean in control	68,632	68,632	68,632
R ²	0.16	0.16	0.16
No. of Unique Firms	2,666	2,666	2,666
Observations	5,096	5,096	5,096

Notes: Table depicts panel regressions of firms surveyed at endline 1 and 2, as consumptive expenditures were only collected in the endlines. Outcomes are in UGX and winsorized to the 2.5 and 97.5 percentiles. ***1%, **5%, *10%.

Table 1.11: Simulation Results

	<i>Partial equilibrium</i>		<i>General equilibrium</i>	
	Aggregate income (1)	Aggregate welfare (2)	Aggregate income (3)	Aggregate welfare (4)
0. Baseline: Firms face info frictions and fixed costs	1.000	1.000	1.000	1.000
1. Provide correct info [†]	1.002	0.997	0.999	0.934
2. Provide liquidity to relieve fixed costs [†]	0.998	1.083	0.989	0.995
3. (1) + (2) [†]	1.001	1.078	0.989	0.992
4. (1) + (2) + high-return movers induced to move	1.388	1.100	1.371	1.014
5. (4) + improved amenities	1.403	1.118	1.390	1.093

Notes: “Improved amenities” sets the amenity value in parishes with below-median amenities to the median.

[†]Changes of less than 0.005 (or less than half of a percentage point) reflect a loss of precision due to computing aggregate outcomes as the fraction of entrepreneurs located in each of 86 parishes times the income or welfare associated with each parish and should not be thought of as meaningful differences (in other words, they are rounding errors). For example, entrepreneurs redistribute across city locations to hold income and welfare constant in partial equilibrium in simulation 1.

Appendix

Appendix 1.A Derivation of Profit Response to Firm Entry in Equilibrium

In equilibrium, as parish-level demand curves are downward-sloping, profits will be determined by jointly solving the consumer's problem and firm's problem, which I model as follows:

Consumer's Problem. Households are assigned to exogenous locations d . Their total expenditure is divided between sectors $k \in \{\text{Retail, Transport, Other}\}$.²⁸ Given prices $p_{kd}(j)$ and exogenous expenditure I_{kd} , households choose the quantity $x_{kd}(j)$ to purchase from each firm j :

$$\begin{aligned} U_{kd} &= \max \left(\int x_{kd}(j)^{\frac{\sigma-1}{\sigma}} dj \right)^{\frac{\sigma}{\sigma-1}} \\ &\text{subject to } \int p_{kd}(j) x_{kd}(j) dj \leq I_{kd} \end{aligned} \quad (1.1)$$

where σ is the elasticity of substitution between goods sold by different firms. This yields the following first order condition:

$$\begin{aligned} \text{FOC: } \quad U_{kd}^{1/\sigma} x_{kd}(j)^{-1/\sigma} &= \lambda_{kd} p_{kd}(j) \\ U_{kd}^{1/\sigma} x_{kd}(j)^{1-1/\sigma} &= \lambda_{kd} p_{kd}(j) x_{kd}(j) \\ U_{kd}^{1/\sigma} \int x_{kd}(j)^{1-1/\sigma} dj &= \lambda_{kd} \int p_{kd}(j) x_{kd}(j) dj \\ U_{kd} &= \lambda_{kd} I_{kd} \\ \implies p_{kd}(j) &= x_{kd}(j)^{-1/\sigma} U_{kd}^{1/\sigma-1} I_{kd} \end{aligned} \quad (1.2)$$

²⁸Their utility function over these sectors is given by:

$$\begin{aligned} W_d &= \alpha_R \ln(X_{Rd}) + \alpha_T \ln(X_{Td}) + (1 - \alpha_T - \alpha_R) \ln(X_{Od}) \\ &\text{subject to } E_d = P_{Rd} X_{Rd} + P_{Td} X_{Td} + P_{Od} X_{Od} \\ \implies \alpha_k E_d &= P_{kd} X_{kd} \equiv I_{kd} \end{aligned}$$

Firm's Problem. Given prices $p_{kd}(j)$ and cost function $C(\cdot)$, firms choose quantity $x_{kd}(j)$:

$$\begin{aligned} Y_{kd} &= \max p_{kd}(j) x_{kd}(j) - C(x_{kd}(j)) \\ &= \max \Delta_{kd} x_{kd}(j)^{1-1/\sigma} - C(x_{kd}(j)) \end{aligned} \quad (1.4)$$

$$\text{where } \Delta_{kd} = U_{kd}^{1/\sigma-1} I_{kd} \quad (1.5)$$

This yields the following first order condition:

$$\text{FOC:} \quad \frac{\sigma-1}{\sigma} \Delta_{kd} x_{kd}(j)^{-1/\sigma} = MC(x_{kd}(j)) \quad (1.6)$$

The FOC can be rewritten as:

$$\begin{aligned} \frac{\sigma-1}{\sigma} p_{kd}(j) &= MC(x_{kd}(j)) \\ \frac{p_{kd}(j)}{MC(x_{kd}(j))} &= \frac{\sigma}{\sigma-1} \equiv \text{Markup}_{kd}(j) \end{aligned} \quad (1.7)$$

Suppose that marginal cost, $MC(x_{kd}(j))$, is of the form:

$$MC(x_{kd}(j)) = \delta x_{kd}(j)^\theta \quad (1.8)$$

Then, substituting between (1.6) and (1.8) yields:

$$\frac{\sigma-1}{\sigma} \Delta_{kd} x_{kd}(j)^{-1/\sigma} = \delta x_{kd}(j)^\theta \quad (1.9)$$

$$\implies x_{kd}(j) = \left(\frac{\sigma-1}{\sigma} \frac{\Delta_{kd}}{\delta} \right)^{\frac{\sigma}{1+\sigma\theta}} \quad (1.10)$$

Note that $MC(x_{kd}(j)) = \delta x_{kd}(j)^\theta$ implies:

$$\begin{aligned} C(x_{kd}(j)) &= \frac{\delta}{1+\theta} x_{kd}(j)^{1+\theta} \\ &= \frac{x_{kd}(j)}{1+\theta} MC(x_{kd}(j)) \end{aligned} \quad (1.11)$$

Recall the objective function of the firm:

$$\begin{aligned}
Y_{kd} &= p_{kd}(j) x_{kd}(j) - C(x_{kd}(j)) \\
&= \left(p_{kd}(j) - \frac{MC(x_{kd}(j))}{1+\theta} \right) x_{kd}(j) \\
&= \left(\text{Markup}_{kd}(j) - \frac{1}{1+\theta} \right) MC(x_{kd}(j)) x_{kd}(j) \\
&= \delta \frac{1+\sigma\theta}{(\sigma-1)(1+\theta)} x_{kd}(j)^{1+\theta} \\
&= \delta \frac{1+\sigma\theta}{(\sigma-1)(1+\theta)} \left(\frac{\sigma-1}{\sigma} \frac{\Delta_{kd}}{\delta} \right)^{\frac{\sigma(1+\theta)}{1+\sigma\theta}} \\
\text{where } \Delta_{kd} &= U_{kd}^{1/\sigma-1} I_{kd} = \left(\int x_{kd}(j)^{\frac{\sigma-1}{\sigma}} dj \right)^{-1} I_{kd}
\end{aligned} \tag{1.12}$$

All firms of sector k in parish d make the same decision, so by symmetry:

$$\begin{aligned}
\Delta_{kd} &= \left(\frac{\sigma-1}{\sigma} \frac{\Delta_{kd}}{\delta} \right)^{\frac{1-\sigma}{1+\sigma\theta}} I_{kd}/N_{kd} \\
\Delta_{kd}^{\frac{\sigma(1+\theta)}{1+\sigma\theta}} &= \left(\frac{\sigma-1}{\sigma\delta} \right)^{\frac{1-\sigma}{1+\sigma\theta}} I_{kd}/N_{kd}
\end{aligned} \tag{1.13}$$

Therefore, the profit response to firm entry in equilibrium depends on a location-specific parameter, Φ_{kd} , which captures the baseline profitability of a parish, and the number of firms, N_{kd} , in parish d :

$$\begin{aligned}
Y_{kd} &= \delta \frac{1+\sigma\theta}{(\sigma-1)(1+\theta)} \left(\frac{\sigma-1}{\sigma} \frac{\Delta_{kd}}{\delta} \right)^{\frac{\sigma(1+\theta)}{1+\sigma\theta}} \\
&= \Phi_{kd}/N_{kd}
\end{aligned} \tag{1.14}$$

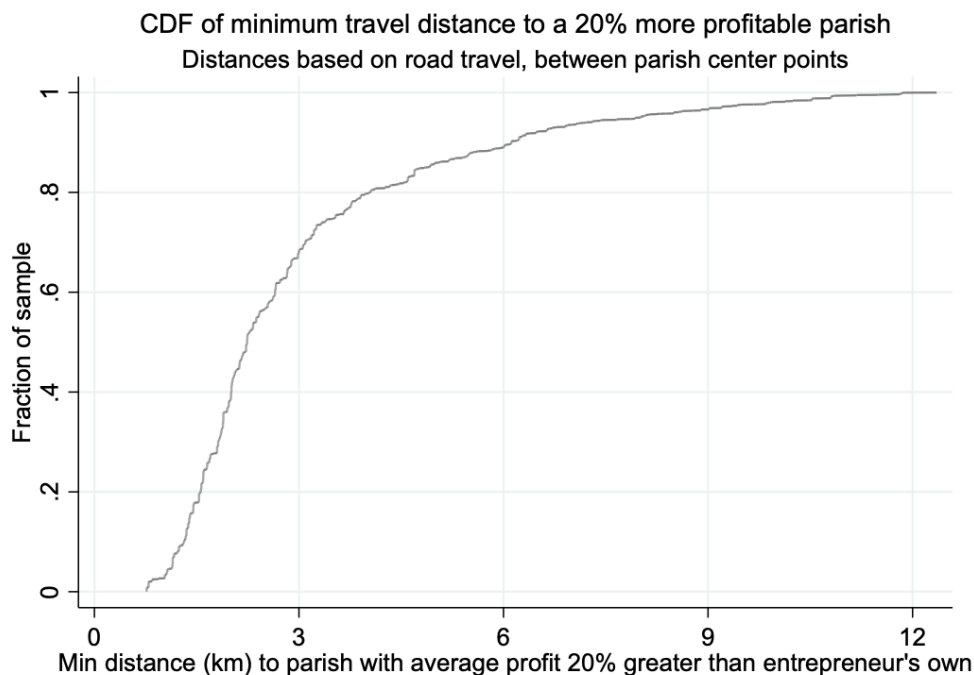
$$\text{where } \Phi_{kd} = \delta \frac{1+\sigma\theta}{(\sigma-1)(1+\theta)} \left(\frac{\sigma-1}{\sigma\delta} \right)^{\frac{\sigma(1+\theta)}{1+\sigma\theta}} \left(\frac{\sigma-1}{\sigma\delta} \right)^{\frac{1-\sigma}{1+\sigma\theta}} I_{kd} \tag{1.15}$$

$$\ln(Y_{kd}) = \ln(\Phi_{kd}) - \ln(N_{kd}) \tag{1.16}$$

Appendix 1.B Data Appendix

1.B.1 The Conditional Cash Transfer

Figure 1.B.1: Distance to More Profitable Parishes



Verifying Compliance with the 3-Kilometer Rule of the CCT. Entrepreneurs assigned to the CCT arm receive contact information for enumerators on the study team whose responsibility is to verify compliance with the CCT. Upon moving to a business location that is three or more kilometers from their pre-intervention location, the respondent calls the study team to report that they are ready for verification. The verification enumerators are located throughout the city such that one is likely to be able to reach the respondent fairly quickly, typically within 1-2 hours. The slight lag, i.e., that verification is not immediate, helps to prevent gaming, e.g., where the entrepreneur tries to take up the subsidy for a brief time during their standard commute. With a lag, the entrepreneur must actually stay in their new location long enough for the verification enumerator to arrive. Upon arrival, the enumerator uses Google Maps to measure the distance between the entrepreneur’s new business location, where they are currently located, and the entrepreneur’s pre-intervention GPS location, of which the enumerator has a record. Enumerators are instructed to allow for up to a 10% deviation in compliance with the three kilometer condition, allowing any distance of 2.7 kilometers or more to “pass” verification. The enumerator must also verify that the

business is truly in operation at the new location. For example, if the respondent sells tomatoes, then the enumerator verifies that the tomatoes are on display to be sold. After verifying that the respondent has complied with the moving requirement, the enumerator submits an electronic record of compliance, including the respondent's new GPS location and a screenshot of the Google Maps distance from the entrepreneur's pre-intervention location. I then initiate an 7,000 UGX mobile money transfer and double check compliance with the three kilometer condition remotely, using road travel distance between the entrepreneur's current and pre-intervention GPS locations, as computed by the Open Source Routing Machine (OSRM).

1.B.2 Estimating Treatment-on-Treated Effects

The second stage estimating equation that yields the local average treatment effect for compliers (which in this case is also the treatment-on-treated effect, as there is no possibility of always-taking) is as follows:

$$Y_{js} = \gamma_0 + \gamma_1 info_j + \gamma_2 \widehat{TookUpCCT}_j + \gamma_3 \widehat{TookUpCCTwithInfo}_j + \gamma_4 \widehat{TookUpUCT}_j + \gamma_5 \widehat{TookUpUCTwithInfo}_j + \delta X_j + \phi_s + \epsilon_{js} \quad (1.1)$$

where $\widehat{TookUpCCT}_j=1$ if an entrepreneur took up the CCT (when offered either alone or with information), $\widehat{TookUpCCTwithInfo}_j=1$ if the entrepreneur took up the CCT when it was offered with information, $\widehat{TookUpUCT}_j=1$ if an entrepreneur took up the UCT (when offered either alone or with information), and $\widehat{TookUpUCTwithInfo}_j=1$ if the entrepreneur took up the UCT when it was offered with information. Thus, the γ terms retain the same additive interpretation as the β terms in equation (1.6). For example, γ_2 is the effect of taking up the CCT alone, while $\gamma_1 + \gamma_2 + \gamma_3$ is the effect of receiving information and taking up the CCT. Four first stage regressions for each outcome X_{js} ,

$$X_{js} \in \{\widehat{TookUpCCT}_{js}, \widehat{TookUpCCTwithInfo}_{js}, \widehat{TookUpUCT}_{js}, \widehat{TookUpUCTwithInfo}_{js}\}$$

take the same form as the primary ITT specification given in equation (1.6):

$$X_{js} = \beta_0 + \beta_1 info_j + \beta_2 CCT_j + \beta_3 (info_j \times CCT_j) + \beta_4 UCT_j + \beta_5 (info_j \times UCT_j) + \delta X_j + \phi_s + \epsilon_{js} \quad (1.2)$$

These first stage regressions yield the predicted values $\widehat{TookUpCCT}_j$, $\widehat{TookUpCCTwithInfo}_j$, $\widehat{TookUpUCT}_j$, and $\widehat{TookUpUCTwithInfo}_j$ in equation (1.1).

Data Appendix Tables

Table 1.B.1: Experimental Sample: Composition along Stratification Variables

	Male	Female	Total
Vendors	1,058	742	1,800
	59%	41%	100%
Boda bodas	750	0	750
	100%	0%	100%
Other services	211	122	333
	63%	37%	100%
Total	2,019	864	2,883
	70%	30%	100%

Table 1.B.2: Attrition by Treatment Arm

	(1) Attrit at midline=1
Info Only	.028 (.024)
CCT Only	.013 (.023)
CCT+Info	-.00091 (.022)
UCT Only	-.026 (.024)
UCT+Info	-.022 (.025)
P-value in F test for joint orthogonality	0.23
Overall Attrition Rate	.16
Observations	2,883

Notes: Heteroskedasticity-robust standard errors in parentheses.
*** 1%, ** 5%, * 10%.

Table 1.B.3: Take-Up of the Intervention

	(1) Took up CCT=1	(2) Took up UCT=1
Info=1	-0.00 (0.00)	0.00 (0.00)
CCT=1	0.54*** (0.02)	0.00 (0.00)
Info=1 × CCT=1	0.02 (0.03)	-0.00 (0.00)
UCT=1	-0.00 (0.00)	0.90*** (0.02)
Info=1 × UCT=1	0.00 (0.01)	-0.03 (0.02)
Info+CCT+Info*CCT	0.56***	-0.00
P-val: I+C+I*C=0	0.00	0.89
Info+UCT+Info*UCT	0.00	0.87***
P-val: I+U+I*U=0	0.87	0.00
Strata FE's	Yes	Yes
Control Mean (Level)	0.00	0.00
R squared	0.44	0.86
Observations	2,883	2,883

Notes: Heteroskedasticity-robust standard errors in parentheses.

*** 1%, ** 5%, * 10%.

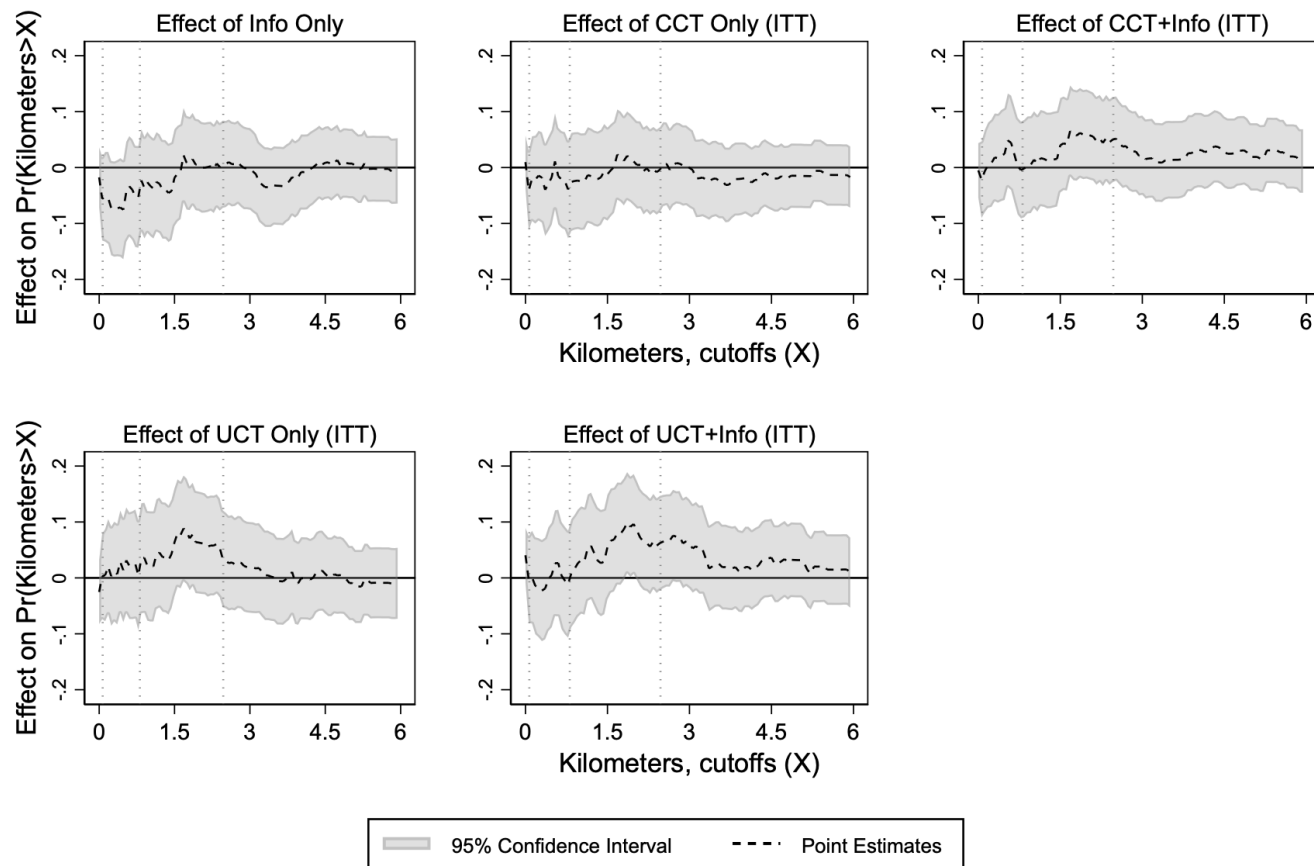
Table 1.B.4: Baseline Characteristics of Compliers

	CCT+Info			CCT Only			
	Complier Mean	Never Taker Mean	P-Value: Complier = Never Taker	Complier Mean	Never Taker Mean	P-Value: Complier = Never Taker	P-Value: CCT+Info Complier = CCT Only Complier
<i>Demographics</i>							
Boda=1	.33	.18	0	.33	.18	0	.97
Female=1	.25	.36	0	.27	.33	.09	.49
Business owner age	32	31	.9	32	31	.1	.41
Completed primary school=1	.7	.64	.08	.69	.66	.44	.54
Completed secondary school=1	.15	.1	.03	.15	.11	.12	.81
Replacement respondent=1	.39	.45	.1	.41	.41	.85	.38
<i>Business chracteristics</i>							
Years running current business	5	4.2	.04	5	4.2	.03	.97
Years running any business	5.5	5	.24	5.8	4.7	0	.4
Min to bus. location from home	36	29	0	33	33	.97	.11
Number of competitors in eyesight	4.3	2.4	0	4.3	2.9	0	.83
Hrs worked last week	67	67	.84	67	65	.34	.85
<i>Revenue, costs, and profit</i>							
Daily revenue (UGX)	63,655	63,326	.95	61,481	66,288	.25	.56
Daily costs (UGX)	38,469	35,879	.49	36,937	38,959	.56	.6
Incurs transport costs=1	.47	.33	0	.48	.32	0	.69
Total daily transport spending	6,095	3,617	0	6,135	3,530	0	.93
Daily profit (UGX)	22,350	22,033	.82	22,546	22,897	.8	.87
Profit per hour	2,186	2,448	.22	2,306	2,402	.53	.37
Below median profit parish=1	.52	.44	.05	.49	.54	.16	.28
<i>Assets and financial services</i>							
Inventory value	96,254	104,736	.66	96,977	101,442	.76	.96
Total asset value	1,080,060	726,433	0	1,133,984	680,349	0	.54
Total savings	221,623	378,303	.12	312,205	146,022	0	.06
Mobile money balance	17,583	25,398	.42	21,632	14,286	.13	.46
Has formal savings=1	.44	.45	.83	.47	.38	.02	.35
Has any credit=1	.26	.16	0	.24	.2	.19	.42
Total credit	223,222	94,291	.01	208,240	133,286	.15	.74
<i>Aspirations and knowledge</i>							
Would like to expand business=1	.91	.94	.16	.92	.94	.16	.7
Patient=1	.26	.24	.55	.24	.26	.47	.52
Time-inconsistent=1	.12	.13	.84	.13	.12	.81	.69
Pct score on info quiz: Close parishes	.54	.45	0	.49	.53	.18	.01
Pct score on info quiz: Far parishes	.51	.48	.15	.48	.52	.1	.14
Pct score on info quiz: Total	.53	.47	0	.49	.52	.11	.02

Notes: There is no possibility of always-taking.

Appendix 1.C Mechanisms Appendix

Figure 1.C.1: Distribution Regressions: Effect of Treatment on Probability of Locating $>X$ Kilometers from Pre-Intervention Business Location, Among Likely CCT+Info Compliers



Notes: Vertical dashed lines reflect the 25th, 50th, and 75th percentile of the outcome. Plots depict intent-to-treat effects for the “likely CCT+Info complier” subgroup. Standard errors are heteroskedasticity-robust and all estimating equations include controls selected via double-selection (DS) lasso, as well as non-penalized strata FEs.

Chapter 2

How Important are Investment Indivisibilities for Development? Experimental Evidence from Uganda

Joseph P. Kaboski, Molly Lipscomb, Virgiliu Midrigan, and
Carolyn Pelnik

2.1 Introduction

Theory suggests that high-yield, indivisible investment opportunities can play crucial roles in development. In particular, at the micro level, they can help explain entrepreneurship decisions and savings dynamics (Buera, 2009; Midrigan and Xu, 2014), can lead to poverty traps (Banerjee and Newman, 1993a; Buera, Kaboski, and Shin, 2014), and are important in predicting the effects of financial interventions and poverty programs (Kaboski and Townsend, 2011; Buera, Kaboski, and Shin, 2015; Banerjee et al., 2024; Buera, Kaboski, and Shin, 2020). At the aggregate level, they are also critical in understanding the quantitative importance of financial frictions on output, productivity, and investment (Buera, Kaboski, and Shin, 2014; Midrigan and Xu, 2014).

Indivisibilities in high-yield investments imply that some households could significantly increase profitability if they had a large amount of additional capital to invest, but may not be able to do so with only small amounts of capital if financial frictions are severe (Dercon, 1998; Barrett et al., 2006; Santos and Barrett, 2011). If poor households lack access to credit, they cannot borrow money to make the indivisi-

ble investment. Likewise, if returns to incremental savings are low, households may not be able or find it worthwhile to save for indivisible investments even when they are high-yield. In the presence of financial frictions, households with high-yield indivisible investment opportunities may fall into poverty traps, and the indivisible investment opportunities may induce the same households to make risk-loving choices for the chance to obtain a sufficiently high payout to make the investment (Kwang, 1965; Chetty and Szeidl, 2007; Lybbert and Barrett, 2011).

We use a field experiment to test the extent to which households will select into a riskier, lower expected value lottery for the chance to obtain a much larger payoff than a safer but lower payoff lottery in peri-urban Uganda where financial services are scant. The experiment enables us to empirically assess the rate at which participants take up the riskier lottery, the characteristics of those selecting the riskier lottery, the investment choices of lottery winners, and the effect of winning the lottery on income and consumption. We motivate the experiment with a model of indivisible investment opportunities in an environment without credit or interest-bearing savings. The model shows that households with a high-yield invisible investment opportunity may be unable to escape poverty traps if they lack access to finance and sufficient wealth to fund the investment.

Next, we show that although poor households are generally risk averse (Rosenzweig and Binswanger, 1993; Donovan, 2021), the presence of high-yield indivisible investment opportunities can induce some households to make risk-loving choices. Our model predicts that high productivity households with assets below what is needed for the indivisible investment are characterized by risk-loving behavior with varying levels of patience and impatience, due to high potential returns. They may also have higher initial savings, in an attempt to finance the indivisible investment good. The lottery experiment poses households with an actual choice between a lottery that pays out approximately \$100 with 50% probability or lottery with lower expected value that pays out approximately \$475 with 10% probability, allowing us to characterize who chooses the riskier lottery. The randomly selected winners receive the payout of their chosen lottery as a cash grant via mobile money. We follow up at 4 months, 18 months, and 6 years with a survey which allows us to empirically assess how investment goods vary by lottery choice.

We find that a substantial share of the sample, 27 percent, does indeed prefer the riskier option. These riskier choices are more common among males, well-known to have lower underlying risk aversion (e.g., Meissner et al. (2023)), but consistent with the model, the risky option is also preferred by high-saving, investment-oriented households. Those choosing the riskier lottery have higher crop income, business assets, and total wealth and are more likely to report a desire for credit to make a business investment. We also find that only winners of the riskier, larger payoff lottery sustain higher levels of investment in indivisible goods at the second endline. Large lottery winners invest in durable business assets (e.g., a generator) and land,

and land exhibits substantial capital gains. In contrast, small lottery winners make short-lived investments in small livestock (e.g., chickens) and business inventory (e.g., retail goods).

In evaluating impacts of winning the grant on the household, we use both a basic OLS specification and a seemingly unrelated regression (SUR). The necessity of a household budget summing up to equal income plus assets implies that the equations which estimate impacts on consumption, savings, investment, income, and borrowing are not independent, and so we use SUR to impose a cross-equation budget constraint. In doing so, we highlight the application of a method commonly used in the estimation of demand systems to the analysis of cash grants.

The model suggests that high-productivity households with new access to a large sum of money invest it and should be able to make large gains in income over time. In the short run (4 months after the lotteries pay out to winners), we find no significant impact on either consumption or income, by small or large lottery winners. In the medium run (18 months after payout), however, we find that large lottery winners have higher consumption, while winning the small grant still has no significant impact on either consumption or income. In a long run follow-up survey (6 years after payout), we find a substantial increase in income for large lottery winners with some (imprecise) evidence of gains to consumption. This is consistent with findings from [Balboni et al. \(2022\)](#) in which a household's ability to leverage a randomized transfer of wealth to work its way out of poverty depends on initial levels of wealth.

Finally, we calibrate the model to reflect participants' actual lottery choices and income and savings distributions, and we use the calibrated model to simulate the impact of financial services expansion. We show that the impacts of relaxing financial frictions on aggregate development in the area depend critically on the elasticity of the supply of the investment good. If the lumpy investment is elastically supplied to a community (e.g., tractors for a farm or refrigerators for a business), then access to credit or savings expands the set of people able to invest, and the increase in demand expands the amount of real investment in the community. In this case, financial services can increase income, investment, and economic mobility. If instead the lumpy investment is inelastically supplied in the community (e.g., land), then increased demand from expanding financial services manifests largely in an increase in prices rather than real investment. The model's results relate closely to our experimental results, since general equilibrium effects may be quite relevant for an investment like land, which is in fixed supply.

Our paper contributes to the literature on the aggregate importance of financial frictions, poverty traps, cash grants, and the impacts of financial services expansion. We add empirical and quantitative guidance to a macro literature on financial frictions and poverty traps. [Banerjee and Newman \(1993a\)](#), [Galor and Zeira \(1993\)](#), and [Piketty \(1997\)](#) are examples of models with indivisibilities, where financial frictions lead to poverty traps. Later work showed that indivisible decisions, when embedded in quantitative models with

intensive margins and mapped to the data, did not lead to macro multiplicities, but only micro poverty traps as in Buera (2009). This resulted in lower aggregate output (Buera, Kaboski, and Shin, 2014; Midrigan and Xu, 2014) and slower convergence (Buera and Shin, 2013). These theories all lead to demand for financing but also risk. We contribute to these findings by showing the empirical importance of both indivisibilities and risk-loving behavior linked together, and using these empirical results to discipline their quantitative importance.

Other closely related papers present empirical evidence on poverty traps and indivisibilities in developing economies. Balboni et al. (2022) examine the impact of a uniform livestock asset grant in Bangladesh and show that the impact is an s-shaped function of the initial assets of the recipient. Our work complements theirs: we allow for an endogenous relationship between productivity and initial wealth, which would confound empirics based on the initial level of assets, so we focus instead on risk preference. Through our use of a risky versus safer lottery, we also examine whether some groups may experience differentially larger returns to investment. Relatedly, Banerjee et al. (2024) present empirical evidence that despite low average returns, microcredit does indeed break a poverty trap for “gung-ho” entrepreneurs, those with prior business experience who exhibit higher than average returns to business investment. In non-experimental settings, Lybbert et al also examine livestock-based poverty traps empirically in a series of papers (e.g., Lybbert et al., 2004; Carter and Lybbert, 2012), and they highlight the role of risk taking (Lybbert and Barrett, 2011). McKenzie and Woodruff (2006) present evidence for poverty traps with observational data on Mexican microenterprises. These papers emphasize business investment and livestock as indivisible investments driving poverty trap dynamics. To this literature, we contribute novel findings on the role of land and capital gains in generating poverty traps.¹ In some cases, rental markets can be a way of overcoming investment indivisibilities (e.g., Bassi et al., 2022; Caunedo and Kala, 2021), but rental is less useful if capital gains are an important part of the return. Our findings on land also complement recent work by Acampora, Casaburi, and Willis (2022), who show that market imperfections lead to the under-utilization of land rentals and lower aggregate returns to land.

Cash grants have become a popular approach to identifying the marginal returns of capital for entrepreneurs (De Mel, McKenzie, and Woodruff, 2008, 2014) and the returns to poverty programs more generally (Blattman, Fiala, and Martinez, 2014; Blattman et al., 2016; Haushofer and Shapiro, 2016, 2018; Egger et al., 2019). Recent work by Banerjee et al. (2023) shows that cash grants of equivalent value have greater impact on revenues and enterprise ownership when delivered as a lump sum (rather than through multiple disbursements of smaller size). This is consistent with work by Casaburi and Macchiavello (2019),

¹Although land is in principle divisible, whether property rights are customary or formal, it is generally divided into discrete plots both in its use and any transactions and titling.

Herskowitz (2021), and Gertler et al. (2023) who show demand for lumpiness, as evidenced by demand for less frequent but larger payments, the use of gambling to finance lumpy expenses, and greater take-up of prize-linked savings when there is the possibility of winning a large cash lottery, respectively. Other related work by Beaman et al. (2023) shows evidence of selection into borrowing that reveals which farmers have high returns to capital. We combine experimental variation in grant payouts with a choice between a safer, lower payoff lottery and a riskier, higher payoff lottery to reveal which respondents face high-yield, indivisible investment opportunities.

Our simulations of the aggregate impacts of financial frictions are an additional contribution. Methodologically, we contribute to an emerging literature using experiments in conjunction with macro development models (e.g., Buera, Kaboski, and Shin, 2014; Bergquist et al., 2019; Donovan, 2021; Lagakos, Mobarak, and Waugh, 2018; Buera, Kaboski, and Shin, 2020).² While these papers use the results of RCTs to discipline simulated impacts of scaled policies in the macroeconomy, our paper is novel in using an experiment to actually test an existing theory of macrodevelopment and quantify its implications. We focus on financial services and illustrate that, even when indivisibilities can lead to micro-level poverty traps where financial services may seem particularly needed and powerful, their scaled impact can depend critically on the elasticity of the supply of the relevant capital.

The rest of the paper is organized as follows. Section 2.2 presents our simple motivating model. Section 2.3 summarizes our experimental methodology and data collection. Section 2.4 characterizes the types of households that select into the large lottery. Section 2.5 describes the effect of lottery winnings on investment and the household budget in general. Finally, Section 2.6 calibrates the model to our data and simulates the macroeconomic consequences of credit injections and redistribution.

2.2 Model and Motivation

In this section, we develop a simple model to illustrate mechanisms through which the combination of high-yield indivisible investments and financial frictions can lead to both poverty traps and behavior that is risk-loving and impatient. This provides motivation for our empirical experiment. Although we focus on qualitative patterns of individual decisions here, the model is parametric, since it will be later used for quantitative analysis in Section 2.6.

2.2.1 Set Up

We discuss the preferences, income streams, and investment decisions of households in turn.

²See Buera, Kaboski, and Shin (2020) for a review of this literature.

Preferences

Households have time-additive preferences over consumption:

$$V_0 = \mathbb{E}_t \sum_{t=0}^{\infty} \beta^t \frac{c_t^{1-\nu}}{1-\nu},$$

where ν is the (positive) coefficient of relative risk aversion, reflecting the underlying assumption that individuals are risk averse by nature. We assume that a fraction $1-p$ of them die each period and are replaced by an equal mass of newborn households with no initial wealth. Therefore, $\beta = \hat{\beta}p$ reflects time discounting, a product of the discount factor $\hat{\beta}$ and the survival probability p . The expectation is over realizations of death but also shocks to labor income, e_t , and capital income, y_t . (Here we loosely refer to capital income and capital to capture any investment income, but keep in mind that empirically, land will be the relevant investment good.)

Income

Labor income is a product of permanent and transitory components, both of which are stochastic. Specifically, we assume that

$$e_t = \bar{e}\varepsilon_t$$

where $\bar{e}$ is a permanent component and ε_t are idiosyncratic, independently- and identically-distributed (iid), lognormal innovations with log mean equal to zero and (log) variance of σ_e^2 .

Capital income is a function of the amount of capital k_t and entrepreneurial ability z_t :

$$y(z, k) = z_t k_t^\alpha - \delta k_t.$$

We assume $0 < \alpha < 1$ so that the production function exhibits diminishing returns to scale. The term, δk_t , is the maintenance expense associated with having k_t units of capital. Like labor income, capital income is also stochastic, since z_t , in parallel fashion, also follows a stochastic process:

$$z_t = \bar{z}v_t$$

where v_t are iid., lognormal innovations with mean zero and variance σ_v^2 .

Financial Environment and Investment Decisions

Households choose their liquid assets and capital. While capital produces capital income, liquid assets a_{t+1} earns an interest rate of r . In principle, liquid wealth can be negative but savings decisions are bounded below, since borrowing is limited by the value of available capital as collateral. We assume that households can borrow up to a multiple θ of this value

$$a_{t+1} \geq -\theta P_t k_{t+1}.$$

Capital can be bought and sold at the price P_t , but, critically, to capture the indivisibility of investment in a simple fashion, we assume it is divided into discrete units and for simplicity we assume that agents can only purchase one unit of indivisible capital. Hence, agents face the binary choice: $k_t \in \{0, k_1\}$. where k_1 represents the size of capital relative to typical income.

Denoting the price of capital P_t , the household's budget constraint is:

$$c_t + a_{t+1} + k_{t+1} - k_t = e_t + (1+r)a_t + z_t k_t^\alpha + (1-\delta)P_t k_t. \quad (2.1)$$

One can reformulate the household's problem recursively with cash on hand, $l = e + (1+r)a + zk^\alpha + (1-\delta)Pk$, as the only endogenous state variable. The value function then becomes:

$$V(l; \bar{e}, \bar{z}) = \max_{c, a', k' \in \{0, k_1\}} \frac{c^{1-\nu}}{1-\nu} + \beta \mathbb{E}V(l'; \bar{e}, \bar{z})$$

subject to:

$$c + a' + Pk' \leq l, \quad (2.2)$$

$$a' \geq -\theta Pk', \quad (2.3)$$

and

$$l' = e' + (1+r)a' + z'(k')^\alpha + (1-\delta)Pk'. \quad (2.4)$$

We normalize P to one in the initial steady state prior to interventions.

2.2.2 Model Results: Poverty Traps and Risk Loving Behavior

We first show how the model generates poverty traps and risk-loving lottery choices. We then summarize how the model motivates our experiment and empirics.

Optimal Choices and Poverty Traps

Poverty traps arise in the model due to the indivisible investment. One can show that the household will never invest if productivity is so low that the return from investing in capital is less than the return to saving in assets. That is, a necessary (though not sufficient) condition for the household to invest is that the expected return to investing k_1 exceed the value of saving that same amount:

$$\exp(\sigma_v^2/2)\bar{z}k_1^\alpha + (1 - \delta)\underline{k} \geq (1 + r)k_1$$

The return to saving (right-hand side) is constant, whereas the return to investing (left-hand side) is increasing in $\bar{z}$. When this constraint holds with equality, it defines a lower bound $\bar{z}^* = (r + \delta)k_1^{1-\alpha} \exp(-\sigma_e^2/2)$ on entrepreneurial productivity such that for lower levels of $\bar{z}$, no investments are made.³ The poverty trap and risk-loving behavior on which we focus are therefore characteristic of those with high entrepreneurial productivity and a desire to invest, a key finding that we will examine empirically.

We now discuss the model's more interesting qualitative predictions with respect to poverty traps, savings behavior, and risk. We present simulated examples of optimal behavior for a particular value of $\bar{z} > \bar{z}^*$, e , and $k = 0$ in Figure 2.1. We also set the parameters of the financial system such that households have no access to credit, i.e., $\theta = 0$ or interest bearing savings, i.e., $r = 0$.⁴ (All parameters follow our later empirical calibration in Section 2.6.) Moreover, for easier exposition, we present the value functions and optimal policies (i.e., behavioral decisions) as functions of liquid assets, $a = l - e$, a simple transform of cash-on-hand. The four panels shows the value function (upper-left), capital choice (upper-right), liquid assets — or savings choice (lower-left), and consumption choice. For each panel, the minimum amount of assets needed to invest in the indivisible capital, k_1 , is critical, so we indicate it with the vertical, dashed, red line. (Though labeled k_1 for simplicity, this line is actually $k_1 - e$ under the transform.)

The upper-left panel shows the value function increases with the level of liquid assets but in a very non-linear way around the threshold, as we discuss below. The upper-right panel shows that if households' liquid assets exceed the minimum threshold, then they invest immediately in the indivisible capital. All households would like to, but poorer households are unable to invest given the borrowing constraint.

Looking at lower panels, we see that savings (left) and consumption (right) decisions are non-monotonic, with both dropping around the investment threshold. At the threshold, savings drops to zero, as the house-

³Again, this is a necessary though not sufficient condition because households are risk averse. The precise threshold to invest therefore depends on the level of liquidity and labor income, i.e., a $\bar{z}^{**}(l, \bar{e}) \geq \bar{z}^*$ defined by $V_0(l; \bar{e}, \bar{z}^{**}) = V_k(l; \bar{e}, \bar{z}^{**})$, where the two value functions are for $k = 0$ and $k = k_1$, respectively.

⁴These stark assumptions are simply illustrative at this point, but they will not be too far off from our empirical setting, where credit levels are quite limited and formal savings accounts bear no interest. We will maintain these parameter values for a benchmark in our quantitative work of Section 2.B.

hold puts all available resources into the indivisible capital because of its higher return. Indeed, consumption also falls to nearly zero, indicating that the threshold asset level to invest is only slightly higher than the needed investment; households are willing to sacrifice consumption, when high-yield investment opportunities are available. As assets increase beyond this threshold, however, consumption increases one-for-one as the marginal utility of present consumption exceeds the expected marginal utility of future consumption (which will likely have higher income because of the investment). Thus, households, though wealthy, are hand-to-mouth.⁵

The dashed blue line in the lower-left panel is the 45-degree line. The intersection of the savings policy function with this line defines a critical value of assets, a^* . Since the savings function crosses the 45-degree line from below, this intersection defines an unstable steady state but also characterizes poverty trap dynamics: a household with assets above a^* will (in expectation) save up and increase its liquid assets over time to eventually invest in the high-yield indivisible investment. In contrast, a household with liquid assets below a^* will (in expectation) see its assets decline to the stable steady state at zero over time.

Looking at consumption decisions in the lower-right panel, one can see the importance of the threshold, which corresponds to a local peak in consumption. However, at asset levels below this peak, consumption is lower because assets are lower. At asset levels above this peak, consumption is lower despite assets being higher. Hence, households saving to invest display an apparent high level of patience, not because of an innate higher discount factor but because the household is saving to self-finance the high-yield indivisible investment in the future.

Finally, we return to the upper-left panel, to see how this behavior impacts the value function. Note that the slope of the value function reflects the marginal value of assets, which varies considerably. The marginal value of assets is highest just to the right of the threshold, given the low level of consumption. Indeed, to the right of this threshold, the value function exhibits the standard concavity in assets, reflective of the diminishing marginal utility of consumption implied by the assumed logarithmic utility function. However, to the left of the $\underline{k}$ threshold the function is convex, reflecting the declining consumption as assets approach the threshold level.

Lotteries and Risk

The convex region of the value function leads to a preference for risk which will motivate our empirical study. In the lottery experiments of the next section, we will offer a choice between four different lotteries: (i) a “safe now” lottery that pays Δ^S with probability π^S immediately; (ii) a “safe delayed” that pays off with

⁵This nonmonotonicity in consumption with indivisible investment was emphasized by [Kaboski and Townsend \(2011\)](#), and such people in advanced economies were later coined the “wealthy hand-to-mouth” by [Kaplan and Violante \(2014\)](#) and [Kaplan, Violante, and Weidner \(2014\)](#).

3 percent interest, i.e., $1.03\Delta^S$, next period with the same π^S probability; (iii) a “risky now” lottery that pays off $\Delta^R > \Delta^S$ but with only probability $\pi^R < \pi^S$; and (iv) a corresponding “risky delayed” lottery that pays $1.03\Delta^R$ next period with the same π^R probability. To assess these lotteries in the context of theory, we denote these four options as SN , SD , RN , and RD , respectively, and use these as superscripts on the value function to denote the value of choosing each lottery.⁶ As in our empirical experiment, the risky lotteries have a slightly lower expected payout, so typically risk aversion would push households to choose the safe lottery. However, the indivisible investment opportunity, together with the borrowing constraint, lead to this choice depending on available cash-in-hand (and therefore, in our simplified example, liquid asset) levels.

The value function of being offered the lottery, V^L , is then:

$$V^L(l; \bar{e}, \bar{z}) = \max\{V^{SN}(l; \bar{e}, \bar{z}), V^{SD}(l; \bar{e}, \bar{z}), V^{RN}(l; \bar{e}, \bar{z}), V^{RD}(l; \bar{e}, \bar{z})\}.$$

Figure 2.2 illustrates V^L (black, solid line) relative to the original V (dashed, black line) and lottery choices by the level of available liquidity for various parameter values of β , $\bar{z} = z$, and $\bar{e} = e$. To connect with the points in the previous graph, we have labeled the level of k_1 as well as the critical value of a^* (within the l space). Clearly, since each lottery has nonnegative payoffs, V^L exceeds V at all points and across all figures. However, the value of the lottery choice (i.e., $V^L - V$) is greatest for a household whose assets are just below the threshold, since the lottery gives them the possibility of being able to quickly make the high-yield, indivisible investment. Which lottery the household chooses depends on its asset levels, however, as well as parameter values.

The point of different values of parameters is to illustrate identification of β and heterogeneity with respect to z and e . The left-hand side figures are for impatient households, ($\beta = 0.8377$) while the right-hand side is for more patient households ($\beta = 0.96$). The top row demonstrates a high productivity household ($z = 1.609$) with average labor income ($e = 1$), the middle row a low productivity household ($z = 0.367$) with

⁶The values of the immediate lotteries are simply:

$$V^{SN}(l; \bar{e}, \bar{z}) = \pi^S V(l + \Delta^S; \bar{e}, \bar{z}) + (1 - \pi^S) V(l; \bar{e}, \bar{z})$$

and

$$V^{RN}(l; \bar{e}, \bar{z}) = \pi^R V(l + \Delta^R; \bar{e}, \bar{z}) + (1 - \pi^R) V(l; \bar{e}, \bar{z})$$

The delayed lotteries require defining continuation values conditional on winning the lottery, since current choices will anticipate future winnings. Define this for risk level $X \in \{S, R\}$ as

$$W^{XD}(l; \bar{e}, \bar{z}) = \max_{c, a', k'} \ln c + \beta \mathbb{E} V\left(l + \frac{1.03\Delta^X}{1+r}; \bar{e}, \bar{z}\right)$$

subject to the budget constraint and law of motion, constraints (2.2) and (2.4), respectively. The value of the delayed lotteries are then

$$V^{SD}(l; \bar{e}, \bar{z}) = \pi^S W^{SD}(l; \bar{e}, \bar{z}) + (1 - \pi^S) V(l; \bar{e}, \bar{z})$$

and

$$V^{RD}(l; \bar{e}, \bar{z}) = \pi^R W^{SD}(l; \bar{e}, \bar{z}) + (1 - \pi^R) V(l; \bar{e}, \bar{z}).$$

average labor income, and the bottom a high productivity household with low labor income. Note that in all cases, the underlying preferences are risk-averse, as we assume, and any risk-loving behavior comes from presence of the high-yield, indivisible investment and the possibility that the large lottery would enable them to make the investment k_1 .

Consider the upper-left panel. When the household already has enough liquid assets to invest, it chooses the safe now lottery, since investing is irrelevant and the value function is concave. Similarly, households just below the threshold choose the “safe now” lottery, since winning it yields enough to make the investment, it pays off with a higher probability, and it pays off now (so the investment can be made immediately). In the intermediate ranges, as assets get further below the threshold, the choice moves to the “safe delayed”, “risky now”, and “risky delayed” choices. When assets are low enough that winning the “safe now” lottery will not quite enable immediate investment, the “safe delayed” is chosen because delaying increases the available resources next period. For still lower levels of assets, the “risky now” is preferred, as the household is willing to take the lower odds of winning because only the higher payout will enable it to invest. Again, at even lower levels, the “risky delayed” is preferred, when winning the “risky now” would still not enable immediate investment. Note that the risky lottery choices are largely confined to those regions below a^* , i.e., households for whom saving out of poverty is suboptimal. Finally, at the lowest level of investment, the “safe now” is again preferred. Here the household is so poor that even winning the highest payout of the “risky delayed” would not be enough for it to be close to investing, the investment opportunity is effectively irrelevant, and the value function is locally concave. It chooses not to delay because it is inherently impatient.

Moving down the panels on the left, we see that both the ordering and range of the various lottery choices and the poverty trap specifics are sensitive to both capital productivity and labor income. A low productivity household dissaves if it cannot immediately afford capital. The benefit of the lottery is much smaller, since moving over the threshold and investing is of less value when the household is low productivity. The risk-loving region is also much smaller. In the lower panel, the agent is productive, but it has lower labor income, and so it is more difficult to accumulate assets. A larger poverty trap exists, i.e., a^* is larger.

Looking at the right panels, we can see the impact of patience. These households are much more willing to save, and the assets accumulate relatively quickly. The poverty trap region therefore disappears, except for the poorest households with low-labor income in the lower-right panel because saving would take too long. Nevertheless, in all cases the choices for the large lottery disappear. Again, it is better for these households to take the small lottery and simply save up since even the small lottery helps them get past the poverty trap threshold.

The contrast between the left- and right-columns illustrates why we will need to assume that households are highly impatient in order to yield risk-loving behavior. The observed lottery choices will facilitate

identification of the model. The heterogeneity in choices across different productivity and income underscores that the relationship between liquidity and lottery choice cannot be taken directly to the data in a simplistic fashion.

Experimental and Empirical Motivation

In summary, the model shows a connection between high-yield, indivisible investments, financial frictions, and poverty traps. The model leads to the following predictions that guide our empirics:

- high productivity households may have higher savings to eventually finance indivisible investment opportunities;
- risky lotteries can be attractive, even though they are lower-expected value, in order to help finance indivisible investments;
- those choosing risky lotteries are high productivity; and
- high productivity households may exhibit either patience or impatience in their lottery choice.

Our empirics in turn use a field experiment to measure the extent to which households select into a riskier lottery, which types of households choose the riskier lottery, what types of indivisible investments they make upon winning, and whether households choosing and winning the riskier lottery are able to escape poverty.

2.3 Field Experiment

Our experiment presents participants with actual lotteries that mirror the theoretical lotteries of the previous section. This section provides details of the design and implementation of our lottery experiment as well as our data collection.

2.3.1 Experimental Design

To draw our sample, we worked with a prominent microfinance bank that was hoping to expand services. They identified three geographically dispersed, underbanked districts where marginal populations lacked financial services and they were considering expanding their services which we used for our sample: Ntung-amo, Ibanda, and Kagadi.⁷ All three are district capitals with populations of roughly 20,000-30,000. The surveyed neighborhoods are best described as peri-urban and rural. Using a neighborhood census of each of the targeted neighborhoods, we randomly selected a sample of 1,048 participants, each from a distinct

⁷We piloted a related project in Mpigi in order to evaluate our survey instruments and other protocols.

household.⁸ We oversampled entrepreneurs and those who lacked formal financial services (i.e., had no formal loan or savings account at the time). All three districts are evenly represented in the sample: 350 participants come from Ntungano, 349 from Ibanda, and 349 from Kagadi.

The timing of the experiment was as follows. In October and November 2015, we conducted a baseline interview of households including questions on their household demographics, income, consumption, agricultural and business activities, assets, borrowing and savings, and hypothetical preferences over income, investment, risk and willingness to delay payment for higher returns. At the end of the survey, all households were offered a free, zero-interest formal savings account at a microfinance bank with a local branch. In February and March of 2017, approximately 16 months after participants received their savings accounts, we conducted the midline survey, where we resurveyed participants about their updated responses to select baseline questions and then implemented the cash lottery experiment described below. We then conducted a comprehensive first endline survey in June and July of 2017. We surveyed those who had chosen to receive delayed payments a month after the impatient sample who received non-delayed payments, so that all participants were surveyed 4 months after receiving their cash grants. The first endline survey consisted of similar questions to the baseline, but included additional questions about how recipients had used their grant money. We conducted a second endline survey, meant to track medium-run household outcomes and observe changes in grant effects over time, in September and November 2018, approximately 18 months after cash grant receipt. Finally, we conducted a third endline survey to track long-run outcomes in June and July 2023, 6 years after grant receipt.

The experiment at midline asked the participant to choose between the following two lotteries: (i) a 50 percent chance of winning a grant of 350,000 Ugandan shillings or UGX (\$98 based on an exchange rate of 3,570 UGX = 1 USD in February, 2015) or (ii) a 10 percent chance of a grant of 1,700,000 UGX (\$476).⁹ The sizes of grants were chosen based on baseline questions about desired investment amounts and demand for credit.¹⁰ As in the model of the previous section, the second lottery had both higher risk and a lower expected value (\$47.60 vs. \$49.00). We visually primed participants through “practice” lotteries in order to assist them in understanding the probabilities and the lottery. Finally, we gave participants the choice to receive their grants (conditional on winning) via mobile money either the following day or to delay payment

⁸Neighborhoods were randomly selected by placing a grid across each city and randomly selecting gridpoints overlaid on a map of each city. Gridpoints that fell in manufacturing zones, parks, or other unpopulated areas were omitted. We then conducted a census of the neighborhood around gridpoints. From the census, we selected the sample, stratifying by use of formal financial services, expressed desire for a savings account, and occupation (entrepreneur, salaried employee, or farmer). We also stratified over gender and whether or not the recipient was a head of household.

⁹The effect of winning the lottery, thus, is identified conditional on lottery choice — whether the respondent selected into the small or large lottery. We do not cross-randomize participants into winning the lottery not of their choice due to budget limitations.

¹⁰Specifically, we asked at baseline whether they would invest if they had access to credit, and how much they would need to make their desired investment. The sizes of the grants approximately match the 10th and 75th percentile responses.

for 30 days in return for 3 percent interest. Our use of mobile money and a minimum of one-day delay was designed to limit differential perceived risk in whether they would actually receive funds since no one received cash immediately in hand. After asking questions about how they would use their grants, we ran the lotteries using random number generators on tablets. Participants learned their outcomes immediately.

Table 2.1 shows the lottery choice preferences of the sample. 765 people (73% of the full sample) chose the small grant. The choice of the small, less risky lottery with no delay was therefore the most common. Despite its riskiness and lower expected payout, 283 people (27% of the full sample) chose the lottery for the larger grant. This is again quite a powerful result since the large lottery had a lower expected value and risk aversion tends to increase in lottery size (Holt and Laury, 2002). The small lottery payout was equivalent to roughly 1 month’s average household income, while the large lottery equated to 4 – 5 months of average household income. As shown in the theory of high-yield indivisible investments in Section 2.2, participants may be willing to accept higher risk lotteries in order to enable a large, high return investment.

In the context of the model, we expect some individuals to be patient, i.e., those for whom the higher payoff from delayed payment enables the indivisible, high-yield investment. This patience may be among those choosing either the small or large lottery. (Recall the RD and SD regions in Figure 2.2.) Of those who chose the small lottery, 144 people (14% of the full sample) chose to delay the payment by a month in return for a larger payment, while the remaining 621 (59%) chose to receive the payment the next day. Of those who chose the large lottery, 78 participants (28% of the high risk sample, 7% of the full sample) were patient and willing to wait a month for a larger payout, while 205 (20% of the full sample) were impatient. The high level of impatience is perhaps surprising given the high foregone rate of return (43%, annualized and compounded) and non-negligible absolute return (equivalent to 2-6 days of income for the median household, depending on the size of the grant) and will require high discount rates in order to reconcile, as shown in Figure 2.2. The sizable fraction of people willing to choose a lower expected value, large lottery is consistent with the model’s prediction driven by the presence of a high yield lumpy investment. We call these people and their choice “risk loving” throughout, though it is important to note that risk-loving behavior is expressed even though we assume risk aversion. While their choice of the lower expected value, large lottery is “risk loving” they are selecting it because of potential large investment payoffs not because of a relative preference for risk.

Knowing that the larger lottery (i.e., “risk loving” choice) was less likely to pay out and correctly anticipating that it would also be less popular, we increased the actual probability of winning the larger grant to 30% rather than the 10% probability communicated to participants, in order to increase the sample size of winners.¹¹ Table 2.2 presents the number of lottery winners in each category. We present the number

¹¹This change was not disclosed to either the participants or the data collection team. One could be concerned about the

of winners as a percentage of the people who made the choices. Of the 765 who chose the small lottery, 373 won the lottery for a win rate of 49%. Of the 283 who chose the large lottery, 85 won, for a win rate of 30%. Because of budget and sample size limitations, we did not randomize any participants into winning the small grant if they lost the large grant, or into winning the large grant if they lost the small grant.

The lotteries themselves provide our experimental variation. The lottery *choices* are of course endogenous, so we control for lottery choice in all specifications. Conditional on the lottery choice, the lottery *outcomes* are random. As discussed in Section 2.3.2, Tables 2.3 and 2.4 confirm balance between the treatment (grant recipients) and control (non-recipients) for both the small and large lotteries.

2.3.2 Data

We collected data over five survey waves: baseline, midline, first endline, second endline, and third endline. Each survey wave includes modules on income, consumption, agricultural activities, non-farm business activities, assets, borrowing and savings.¹² For several of the financial outcomes, we collect data in a multi-step process in order to improve measurement. First, we use questions about subcategories in order to prime the respondent about the many different specific sources of income, expenses, assets, and liabilities that may be relevant. Second, we summed the subcategories to get the total value and confirmed with the respondent whether the aggregate of the category reflects their perception. For example, income is collected as “typical monthly income” (revenues net of costs) using detailed questions about typical monthly revenues on subcategories — i.e., business, crop, livestock, and labor income — and typical costs, normalized to a monthly frequency, as well as direct questions on aggregate income and following up to see whether the participant viewed the aggregate or sum of the components as a better measure. We collect the respondent’s income separately from household income because the former may be measured with less noise. The Data Appendix (Appendix 2.A) describes these steps and measures in more detail.

Using these data, we look first at the use of the grant on investments: consumptive assets (home durables), productive divisible investments (small livestock and agricultural tools and business inventory), and productive indivisible investments (large livestock and ploughs, durable business assets, and land). Next we observe the impact of the grant on the other components of a household budget: consumption, savings, income, and net credit. The distinction between flows and stocks is important from a theoretical standpoint,

possibility that the higher payout rate was inferred by participants, but there is no statistically significant difference in the probability that participants choose the large lottery in the last three days of surveying (accounting for 24% of the sample) within each district, among those surveyed on the enumerator’s first visit attempt. This is consistent with participants maintaining the same evaluation of the large lottery’s expected value over time.

¹²The midline survey, however, is somewhat briefer than other surveys and does not collect updated information on all of the household assets included in the baseline survey. Specifically, we do not update the value of home durables, land, livestock and other agricultural assets from baseline to midline. Tables 2.3 and 2.4 depict balance on the most recent pre-intervention measurement of each outcome.

as well as for our construction of the household’s budget constraint. We measure income as a flow, and we measure savings, assets (business, agricultural, and land), and net credit as stocks. Income is the monthly flow of realized household income (crop profit, livestock profit, business profit, wages, and remittances). Savings and assets reflect current total levels at each respective endline. Net credit is current debt less any lending to others that the household expects to be paid back. At the first and second endline, we measure consumption as total spending on household expenditures (food, transportation, fuel, airtime, and any irregular expenses, like school, hospital or marriage fees) since the midline plus value of home durables. We combine household expenditures, which are collected as flows, and home durables, which are collected as stocks, together into this single measure of total spending between the midline and the respective endline.¹³ At the third endline, we analyze impacts on the flow of weekly consumption separately from the stock of home durables, given that we are more concerned with understanding how potential long-run impacts might manifest, rather than detecting the impact of the grant on household budget components.

We also measure land values as estimated by the participant in all three endline waves. In the third endline, we add questions on purchases, sales, and investment (collected separately from total value). As with many peri-urban areas in developing countries, land values in Uganda increased significantly over time (e.g., [Wineman and Jayne \(2018\)](#); [Gochberg \(2021\)](#)). Because there was significant overall appreciation in land, we estimate rates of appreciation by imputing capital gains using *control* households. That is, we calculate the average increase in land values over time using the land values of control households.¹⁴ On average, land values appreciate an estimated 41% over the 18 months between midline and second endline, or 2% per month (see Table 2.A.1 and Appendix 2.A for more detail on capital gains and their calculation).

Land value appreciation may either be a result of general local price increases, average general investment in land, or from spillovers of the experiment itself, i.e., increased demand for land as a result of increased availability of capital due to the lottery. To estimate the extent to which the experiment itself may have impacted land prices, we calculate the percent of the increase in land value that can be attributed to nearby households receiving a grant and purchasing land in Appendix 2.B. We use a two-stage least squares approach that instruments the change in an area’s land values with the number of grants disbursed in that area. Table 2.B.1 shows the first stage results, and Table 2.B.2 shows the second stage results. As detailed in Appendix 2.B, we find that land appreciated roughly 20% over the entire 18 month period between the midline and

¹³This differentiates household durable spending from spending on *productive* assets, which we include in the investment category because they may lead to positive income returns. Data on durables were not collected at midline, so we simply add the stock of durables at endline to the sum of consumption flows realized between midline and endline. This should lead to no issues with our application of a household budget constraint, as described in Section 2.5, as long as home durables were balanced at midline, as they were at baseline per Tables 2.3 and 2.4.

¹⁴Note that the overall increase in land values may not be entirely attributable to capital gains. It may be that, on average, all households in these peri-urban areas of Uganda are investing in their land (thereby increasing the land value). This would over-state the total level of capital gains over the period, but it would not affect our estimates of gains in land value for treatment households *relative* to control households.

second endline as a result of the grants themselves, or approximately $(20\%/41\%=)49\%$ of the total land appreciation during the period. This price increase will play a pivotal role in interpreting our aggregate simulations, as it implies that the supply of improved land is relatively inelastic.

Sample Balance

We return to midline data to examine sample characteristics and experimental balance between treatment and control. Again, participants self-selected into either the small or large lottery, but the realization of whether they won was random. We check balance across 26 baseline and midline outcomes and demographic characteristics within the sample. Table 2.3 presents the balance results for those who chose the small lottery, whereas Table 2.4 presents the balance results for those who chose the large lottery.

Before comparing treatment and control, note some sample characteristics. The average midline income of 362,000 UGX translates to \$101, and households average five members. Households are therefore quite poor, and even the small grants are sizable relative to income. Well over 70 percent of households are farmers, and about half are women. Financial intermediation is low with only about one-third of households having loans outstanding, and roughly 10 percent having savings in formal accounts. The overall levels of consumption, income, and savings have implications for the demand for financial services. Though few utilized their zero interest formal savings accounts provided by the project during the baseline, the data show a substantial underlying level of savings that could indicate a demand for financial services. Specifically, comparing the levels of savings to that of consumption, the average household savings is about 7 weeks of average household consumption. We also find large increases in savings from baseline to midline and substantial demand for credit for investment (see Table 2.C.1).¹⁵

Comparing treatment and control, balance is quite good. For the 26 outcomes in the two tables, we would expect roughly 1 variable per table to be significant at a 5 percent level based on type I errors, but indeed none are statistically significant at the 5 percent level, and only 1 - 2 per table are significant at the 10 percent level. Households winning small grants tend to consist of a slightly smaller number of adult females, 1.1, as opposed to 1.2 for the small lottery control. They also have a higher level of business durables, 405,000 UGX in non-stock business assets relative to 286,000 UGX in the small lottery control. Among those selecting the large lottery, winners have greater divisible investments, 642,000 UGX relative to 458,000 UGX (significant at the 10 percent level). Thus, balance between the treatment and control group is as expected.

¹⁵The lack of use of the formal accounts may have several causes: low trust in formal accounts as currently offered, the fact that the accounts are zero interest, or that traditional formal accounts are simply not convenient enough to use, even when offered free of charge. This may be because of the costs associated with transactions like transportation costs, or a mistrust of institutions. Novel approaches like savings groups or cell phone-related savings vehicles may therefore be more appropriate for meeting this underlying demand.

Still there are a few variables in which the differences may be economically significant. In Table 2.3, the average monthly income is roughly 22,000 UGX less among the small lottery “winners” relative to the control, and savings is 21,000 UGX higher, which are both small relative to the size of the 350,000 UGX grant received but not negligible. Similarly, in Table 2.4, we see that average monthly income is 45,000 UGX more in the large lottery treatment relative to the risk-loving control and likewise midline savings is 71,000 UGX more, which is small but not negligible compared to the grant size of 1,700,000 UGX. Land values are 2,843,000 UGX larger (or 22% higher) in the large lottery treatment group relative to the high-risk control, though this difference is not statistically significant. This is part of our motivation for controlling for baseline and midline levels in our endline analyses, as well as other economic and demographic characteristics such as midline income and the number of adult males, adult females, and children. We find similarly good balance among those receiving the grant and those not receiving the grant when partitioning by the choice of early or delayed reception.

Attrition from the Sample

From the midline to the first endline, we retain 983 of 1,048 (94%) original respondents. Of those, we retain 867 respondents (83% of the original respondents) through the second endline, which becomes our primary sample of analysis, allowing us to compare outcomes between the first and second endline for the same group. In the third endline, occurring 6 years after grant receipt, we were able to track 838 (80%) of original respondents.

The rate of retention between midline and second endline does vary between treatment (87%) and control (79%) households, though it does not vary based on choice of small or large lottery (i.e., we retain 87% of both small and large lottery winners). We do check balance on baseline characteristics between the retained and attrited sample, for both those who chose the small lottery and were retained relative to those choosing the small lottery who attrited, and again for those who chose the large lottery and were retained relative to those choosing the large lottery who attrited, in Tables 2.D.1 and 2.D.2. In general, balance is good, and we only find significant differences on age, which we include, along with age², as a demographic control in our outcome analyses. The differential rate of attrition between treatment and control, however, may be indicative of unobservable characteristics that are correlated with treatment status and driving the likelihood that a household remains in the sample over time. To address this possibility, we compute Lee Bounds ([Lee, 2009](#)) around all of our primary OLS point estimates in Appendix 2.D. Our key results — that large lottery winners differentially invest in indivisible assets and increase their incomes in the long run — are robust to the bounds.

2.4 Who selects the large lottery?

The model predicts that entrepreneurs who have large, indivisible investment opportunities are more likely to choose a lower expected return large lottery and that these investments may increase their income over time. In this section, we test the extent to which we see this behavior in our participants. We find that participants choosing the large lottery tend to be people with higher income, higher wealth, and whose wealth has increased quickly since the baseline.

2.4.1 Methods of Predicting Lottery Choice

We examine the significant predictors of the lottery choice in our data. We first run simple OLS regressions to test whether the characteristics that the model predicts matter in the risk decision are in fact correlated with the risk decision. We run the OLS regressions:

$$D_i^m = \beta P_i^m + \gamma X_i^m + \varepsilon_i, \quad (2.5)$$

where D_i^m is the decision of household i at midline m (e.g., a dummy for the risky choice, a dummy for the impatient choice) and P_i^m are the midline predictors on which we test for significant differences between those choosing the small lottery relative to those choosing the large lottery, as well as the impatient relative to the patient. X_i^m is a vector of household-level demographic controls: household income at midline, female, household head status, age, age², number of adult females, number of adult males, number of children, district fixed effects, and patience (when testing for differences in the risk decision). In the OLS specification, we first test whether each predictor P_i^m is statistically significantly correlated with the lottery decision unconditionally, and second whether the predictor is correlated with the lottery decision conditional on the set X_i^m of demographic characteristics. We also use lasso (Tibshirani, 1996) to select which among 160 covariates are most important for predicting the household's lottery choice given a parameter, selected using adaptive lasso, that penalizes additional model complexity.¹⁶ Table 2.5 presents statistically significant predictors of the large lottery choice in the OLS regression.¹⁷ Predictors also selected by the much higher dimension lasso regression are designated with an asterisk.

¹⁶A list of the full set of covariates included in the lasso specification is available on request.

¹⁷The full list of covariates on which we use OLS to test for significant differences between those choosing the small lottery versus the large lottery is in Table 2.C.1 in the Appendix. The point estimates in both Table 2.5 and Table 2.C.1 are from the unconditional specification of Equation 2.5, i.e. without X_i^m . The estimates conditional on X_i^m are available by request.

2.4.2 Correlates of Lottery Choice

The risk choice does not appear to be random but reflects true approaches to risk. In particular, we find that the lottery choice is correlated with proxies for underlying risk preference but also economic variables. We discuss each in turn.

Demographics

Demographics are a reasonable proxy for underlying risk preference. Those choosing the large lottery tend to be men, older, heads of household, and come from larger households with more children and adult males.

Income

In the model, those with high-productivity select the large lottery; we find that this is borne out in the data in that those who select the large lottery have higher incomes. They have significantly higher (40 percent) crop income, significantly higher (13 percent) consumption, and significantly larger (14 additional log points) growth in consumption from baseline to midline.¹⁸ Those choosing the large lottery tend to have higher crop income both in absolute terms and as a fraction of their total income, perhaps suggesting that those with higher levels of home production are more willing to take on risk. While crop income and consumption are both higher for those taking the large lottery, recent changes in total monthly income and business income in particular tend to be smaller for those selecting the large lottery over the small lottery, perhaps reflecting their need to invest in order to increase income. Crop income, in absolute level and per adult equivalent, and growth in consumption, business income, and total monthly income are also selected by lasso.¹⁹

Wealth

We also find suggestive evidence that those selecting the large lottery are wealthier and that they experienced faster increases in their savings and wealth from baseline to midline. Those who choose the large lottery are somewhat wealthier, a per adult equivalent difference of 94,000 UGX (\$26) in total wealth (assets plus savings), and 100,000 UGX (\$28) in net wealth (business assets plus savings minus debt) — approximately 30% of average monthly income. When accounting for an average of 3.4 adult equivalents in the household,

¹⁸On average, incomes increased significantly between baseline and midline. This may be due to seasonal fluctuations in measured monthly income or seasonal variations in the components of income (crop income was especially high, while business income was lower).

¹⁹A natural question is whether the financial predictors are ultimately driven by the demographics. For example, if men are both more likely to choose the large lottery and have a higher propensity to invest or accumulate assets, then perhaps the accumulation of assets is no longer predictive after controlling for gender. We analyze this by looking at the same predictions but controlling for household demographics. The findings are largely robust. One exception is the level of savings is no longer a predictor of selecting the large lottery once income is added as a control, though growth in savings from baseline to midline, as well as the level of business assets and wealth remain statistically significantly correlated with lottery choice. Table available on request.

these differences are smaller than the size of the small grant (350,000 UGX) and are only marginally significant at a 10 percent level.

Those who choose the large lottery do, however, have 43 percent more business assets and hold a higher fraction (5 percentage points) of their total wealth in business assets. The higher growth in wealth and savings before the midline is even more striking. Between baseline and midline, those selecting the large lottery experience significantly larger increases in total wealth (70 additional log points) and total savings (86 additional log points).²⁰ Thus, the accumulation of assets, also selected by lasso, seems to be a strong predictor of the choice of the large lottery.

Desire to Invest

Consistent with the model, the choice of the large lottery is associated with a pre-existing interest in investing. A significantly greater fraction (6 percentage points) of the participants who choose the large lottery report that they could increase their income if they had access to credit. Moreover, a significantly larger fraction (4 percentage points) want to invest an amount greater than \$100, the size of the small grant. Finally, a significantly larger fraction (7 percentage points) indicate that they would use credit for business investment. These predictors are significant at the 5% level in the OLS regression and all selected by lasso. Thus, it appears that, at least for some, the demand for the large lottery could be driven by a desire to invest.

Baseline Risk Choices

We have also have two *hypothetical* risk decisions involving investment collected at baseline. In contrast to the midline lottery choice, the riskier options involve the risk of absolute losses but with less variance and positive expected gains for the riskier option.²¹ That is, neither decision involves actual risk-loving behavior, as the midline choice does. We find higher willingness to take these risks, and the percentage is higher for the expected higher payoff. As with the actual lottery choice, regressing these risk choice measures on demographics and economic states yields similar relationships though with slightly different variables showing up in lasso. Namely, demographics predict risk choice in sensible ways, and those with higher income, wealth, and consumption are more willing to take risks. In all three risk choices, we therefore interpret that willingness to take risk depends on both underlying preferences (proxied by demographics) and by economic conditions. (See Table 2.C.3 in the Online Appendix.) The correlation between baseline and midline choices is low, however, consistent with the stochastically changing economic situations, as in

²⁰These changes are large because of the very low base levels of wealth in the sample: 25% of the sample has 0 wealth at baseline. Before taking logs, we set any negative or 0 values equal to the minimum value + 1 Ugandan shilling.

²¹Specifically, the first question asks whether the participant would make a risky investment option that yields either 10,000 UGX or 300,000 UGX with equal probability (expected value of 155,000 UGX) at a cost of 100,000 UGX. A follow up question asks the same payoffs but at a cost of 150,000 UGX.

the model.

Though the theory is more ambiguous, we also test for significant predictors of the impatient versus patient, but find few differential characteristics. We control for patience in the remainder of our analyses, but do not interact the lottery outcome with time preference. In sum, those whose wealth increased quickly between baseline and midline and those who expressed a pre-existing interest in investing were more likely to choose the large lottery. Those choosing the large lottery also tended to have higher crop income, both in absolute terms and as a proportion of their overall income. These predictors are robust to the inclusion of demographic controls. The level of crop income, growth in wealth from baseline to midline, desire to invest, and gender are also selected by lasso, underscoring the importance of these attributes for predicting risk-loving behavior. These results support the hypothesis from the model that a pre-existing business investment is important for the choice between lotteries.

2.5 How are winnings spent and what are the household impacts?

The theory predicts that those who win large lotteries should make large investments with high rates of return which increase their income. Based on the model, we expect to see impacts of the large grant on large assets, where households use the proceeds of the lottery to invest. For small lottery winners, we may expect transitory impacts on savings (or other divisible investments that perhaps mirror savings, such as inventory and small livestock), while households accumulate additional funds to invest in a large asset or for those who simply smooth their consumption of small lottery winnings over time before returning to poverty traps.

To evaluate the effect of winning the lottery, we start with OLS regressions to estimate the impact of the large and small grants, conditional on lottery choice (between the small and large lottery). For each household i in district d , we regress:

$$Y_{id}^e = \beta_0 + \beta_1 Win_i^m + \beta_2 Win_i^m * D_i^m + \beta_3 D_i^m + \rho Y_i^b + \gamma X_i^m + \lambda_d + \varepsilon_{id} \quad (2.6)$$

where b denotes baseline, m denotes midline, and e denotes endline.²² D_i^m reflects the household's lottery selection, where $D_i^m = 1$ if the household selected the large lottery, $Win_i^m = 1$ if the household won either lottery (small or large), and their interaction, $Win_i^m * D_i^m$, is the additional effect of winning the large lottery. Thus, in this specification, β_1 identifies the (total) effect of winning the small grant, β_2 identifies the additional effect of winning the large grant relative to the small grant, and $\beta_1 + \beta_2$ identifies the total effect of winning the large grant. We conduct F tests on the sum $\beta_1 + \beta_2$. Critically, β_1 and β_2 are estimated

²²Recall that both baseline and midline take place pre-intervention, as the midline survey concludes with the household's lottery choice. We use e to generally denote the endline survey waves, but results will be reported separately for each endline.

conditional on lottery selection, as captured in D_i^m . (In our regression tables, *won lottery* corresponds to Win_i^m , *won large lottery* corresponds to the interaction term $Win_i^m * D_i^m$, where large lottery winners are those who both won their lottery and selected the large lottery, and *risk loving* corresponds to D_i^m .)

In this regression model, λ_d are district fixed effects and X_i^m is a vector of household-level demographic controls: patience, household income at midline, age, age², gender, household head status, number of adult females, number of adult males, and number of children. We winsorize to the 5th and 95th percentiles to ensure that results are not driven by outliers — a concern given the high degree of skewness in the income and asset distributions. We examine impacts on endline levels conditional on baseline levels (as denoted by Y_i^b), consistent with the prescriptions in [Bruhn and McKenzie \(2009\)](#).²³ We apply multiple hypothesis corrections in Appendix 2.E.

2.5.1 How Lottery Winnings are Spent: Impacts on Investment

We estimate the effect of winning the lottery on small and large agricultural and business assets, as well as land and home durables. We categorize these investments into consumptive assets, productive assets that are relatively divisible, and productive assets that are relatively indivisible: *consumptive assets*: (1) home durables (e.g., a radio); *divisible investments*: (2) small livestock and agricultural tools (e.g., a pig or a spade), (3) business inventory (e.g., items for retail such as soap or salt); *indivisible investments*: (4) large livestock (e.g., cattle) and ploughs,²⁴ (5) business durables (e.g., a sewing machine or mirror for a salon) and (6) land.²⁵ Appendix 2.F includes descriptive statistics summarizing the purchases that respondents reported making with the grants.

In our discussion of results, we describe point estimates in levels of Ugandan shillings (UGX), with the percent change relative to the control group following each estimate in brackets, where the small (large) grant control group consists of non-winners who also chose the small (large) lottery. We will start with the treatment effects in the short run (first endline, 4 months after grant receipt) and continue with a discussion of treatment effects in the medium run (second endline, 18 months after grant receipt).

²³In cases for which pre-intervention levels of the outcome were collected at midline rather than baseline, the ρ term instead controls for the midline level of the outcome (where the midline also took place prior to the intervention). Specifically, business inventory, (non-inventory) durable business assets, and thus the aggregated measures of divisible and indivisible assets, are constructed at midline rather than baseline, since inventory and business durables were not disaggregated at baseline.

²⁴The plough is the only large agricultural machine reported by anyone in our sample.

²⁵In principle, land is divisible. In practice, we rarely see purchases of land below 500,000 UGX — only about 5% of land transactions made recently, in June 2013 or after, are valued at less than 500,000 UGX — suggesting that land is typically sold in larger units. Small, disbursed, parcels of land are unlikely to be as productive as large, connected plots; as a result potential land buyers must either restrict themselves to purchasing from neighbors or purchasing a large plot on its own. The matching process between buyers and sellers can create inefficiencies due to the thinness of particular land markets as shown by [Bryan et al. \(2022\)](#).

Short-Term Results (First Endline at 4 Months)

In the short run (3-4 months), we find that small grant winners make statistically significant investments in small livestock, large livestock, and business inventory, while large lottery winners make statistically significant investments in large livestock, business inventory, and land (Table 2.6). Those who win the small grant invest 191,000 UGX [31%] more in divisible goods than the control group ($p < 0.01$). Decomposing this divisible investment effect, we find that small grant winners have 84,000 UGX [31%] greater levels of small livestock and agricultural tools ($p < 0.01$) and 128,000 UGX [45%] greater levels of business inventory ($p < 0.01$) than the control group at first endline. We find no statistically significant effect of the small grant on total indivisible investment, though in decomposing indivisible investments, we see that small grant winners have 90,000 UGX [44%] higher levels of large livestock and ploughs ($p < 0.10$). We find no statistically significant effect of the small grant on home durables.

Winning the large grant, in contrast, leads to statistically significant effects on both total divisible and indivisible investment. Large grant winners have 319,000 UGX [40%] greater investment in divisible goods ($p < 0.05$) and 4,300,000 UGX [22%] greater investment in indivisible goods ($p < 0.05$), all relative to non-winners who also selected the large lottery. As in the case of small grant winners, we find no statistically significant effect of the large grant on home durables. We find that the divisible investment effect is driven by a 302,000 UGX [94%] increase in business inventory ($p < 0.01$), as we find no effect of the large grant on small livestock. The indivisible investment effect for large grant winners is driven by land, where the large grant leads to a 4,454,000 UGX [26%] increase in land values ($p < 0.05$). Large grant winners also have 215,000 UGX [70%] greater investment in large livestock and ploughs ($p < 0.05$). At the first endline, we find no statistically significant effect of winning the large grant on business durables, the other indivisible investment category.

The divisible investment effect for large winners is 67% larger than the effect for small winners, but the difference is not statistically significantly different. Large lottery winners do differentially increase their investment in indivisible goods relative to small lottery winners ($p < 0.05$). This is driven by land: large lottery winners' investment in land is differentially larger than that of both small lottery winners and the control group ($p < 0.05$), while their investment in large livestock and business durables is not statistically significantly larger than small lottery winners.

Medium-Term Results (Second Endline at 18 Months)

In the medium run (18 months), neither small nor large grant winners sustain statistically significantly higher levels of investment in divisible goods, and only large grant winners sustain a statistically significantly greater

level of total indivisible investment, including statistically significant increases in both business durables and land (Table 2.7). Small grant winners only sustain statistically significant increases in large livestock and ploughs, where their level of investment is 91,000 UGX [50%] higher than that of non-winners who also selected the small lottery ($p < 0.10$). The short-term effect on total divisible investment — small livestock and agricultural tools and business inventory — observed for small lottery winners at first endline is now smaller and statistically indistinguishable from zero. The short-term effects on total divisible investment, and on business inventory in particular, also fade for large lottery winners. Large lottery winners' investment in divisible goods is not statistically significantly different from small lottery winners, in aggregate or in the disaggregated divisible investment categories. Large lottery winners do, however, retain differentially higher levels of indivisible investment, in aggregate ($p < 0.10$) and for land ($p < 0.05$) in particular, relative to small lottery winners at second endline. In total, their indivisible investments are 7,638,000 UGX [31%] higher ($p < 0.05$) at second endline than the non-winners who also chose the large lottery. In decomposing the indivisible investment effect for large lottery winners at second endline, we see that the effect on investment in large livestock and ploughs is nearly zero (not statistically significant), but large lottery winners report 514,000 UGX [56%] greater investment in business durables ($p < 0.10$) and 7,488,000 UGX [32%] higher land values ($p < 0.05$) than non-winners who also selected the large lottery.

In sum, both small and large lottery winners make initial investments in divisible goods, which fade for both groups in the medium run. Only large lottery winners increase total investment in indivisible goods, and they sustain elevated indivisible investments in the medium run. In the short run, large lottery winners make indivisible investments in large livestock and land, while in the medium run, their indivisible investments consist of business durables and land. These results are consistent with the model's predictions, where small and large lottery winners experience short-lived increases in divisible investments that mirror savings (such as business inventory and small livestock). Only large grant winners are able to accumulate sufficient funds to sustain a higher overall level of indivisible investment in the medium run.

We note that the effect of the large grant on indivisible investment, in total and for land in particular, is quite large. At first endline, the effect of the large grant on total indivisible investment and on land values is more than 4,000,000 UGX, which is more than twice the size of the large grant itself. By second endline, the large grant effect on indivisible investment and land values is more than 7,000,000 UGX, or more than four times the size of the large grant itself. This may suggest mis-measurement, co-financing through other wealth or income, joint investments with others not included in the measurements, or differential increases in value for lottery winners. In the next section, we consider impacts on non-investment components of the household budget, which provides estimates on potential sources of co-financing.

2.5.2 Impact on the Full Household Budget

We now consider the effect of winning the lottery on all components of a household’s budget, namely consumption, savings, investment, income, and borrowing. We begin by estimating OLS regressions, following the same specification as laid out in Equation 2.6.

Our OLS approach, standard in the treatment effects literature, considers each outcome as an independent regression. However, the basic adding up of a household budget implies that the estimating equations are not independent but governed by a common budget constraint. That is, the following simple budget constraint should hold:

$$\Delta C + \Delta S + \Delta K = x * \Delta Y + \Delta B + Grant \quad (2.7)$$

where C is consumption (including purchases of home durables), S is savings, K is capital (both divisible and indivisible), such that the change in capital reflects investment, Y is income, and B is net credit. The treatment difference, i.e., treatment minus control, is reflected in each Δ term.²⁶ We use $x \in \{4, 18\}$ to denote the number of months between midline (the time of grant allocation) and the relevant endline, since income is measured as a monthly flow. We measure all other components of the budget constraint in levels, including consumption, which is unusual, but allows us to combine spending on home durables (a stock) with the flow of consumption, which we aggregate into a total level of spending between midline and each endline.

We complement our OLS analysis with a seemingly unrelated regression (SUR), which has three advantages. First, it allows us to impose a cross-equation restriction on the coefficients associated with the household’s budget constraint. Second, it allows us to account for the correlation in error terms across equations that the constraint implies. Both of these improve efficiency, if they hold. Third, the cross-equation restriction yields coefficients that necessarily sum to the household’s total available resources, and thus an increase in spending across categories does not exceed the amount of the grant (plus any increase in income or decline in spending in other categories that can be attributed to the grant).

The SUR uses the same basic specification as in Equation 2.6, with the constraint as a cross-equation restriction. Specifically, we apply this restriction across the coefficients on $Win_{i,m}$:

$$\beta_{1.consumption} + \beta_{1.savings} + \beta_{1.divInvestment} + \beta_{1.indivInvestment} - x\beta_{1.income} - \beta_{1.credit} = 350,000 \quad (2.8)$$

²⁶This can be derived by our theoretical model by defining $Y_t = e_t + z_t k_t^\alpha - \delta k_t$ and recursively iterating on the budget constraint, Equation 2.1 from midline to endline. In our empirical setting this yields $\sum_{t=m}^e C_t + A_e + K_e = \sum_{t=m}^e Y_t + A_m + K_m + Grant$. We further break out net assets into gross savings and gross debt, i.e, $A_e = S_e - B_e$. Then differencing this between treatment (winners) and control (non-winners), and using the fact that $E[A_m + K_m]$ is equal for winners and non-winners because of randomization, yields Equation 2.7.

And this restriction across the coefficients on $Win_{i,m} * D_{i,m}$:

$$\beta_{2.consumption} + \beta_{2.savings} + \beta_{2.divInvestment} + \beta_{2.indivInvestment} - x\beta_{2.income} - \beta_{2.credit} = 1,350,000 \quad (2.9)$$

where 350,000 UGX is the amount of the small grant, 1,350,000 UGX is the additional amount of the large grant, and $x \in \{4, 18\}$ denotes the number of months between midline and the respective endline, which is applicable to the coefficient on monthly income. All other outcome variables are stocks.²⁷ Divisible and indivisible investments are constructed as an aggregate of their respective subcategories, as in Section 2.5.1, but indivisible investments are adjusted downward by the capital gains common to land values in their district, through the process described in Section 2.3.2 and in greater detail in Appendix 2.A.²⁸ This is a simple adjustment to reflect real value at the time of purchase, net of any appreciation that may have taken place between the grant allocation and endline surveys, and is necessary if the budget constraint is to hold. Further, assuming that respondents' reported quantities reflect their "true" values, winsorization should also cause the constraint to be rejected. To eliminate the chance that winsorization leads the constraint to be rejected while still estimating the SUR on a sample that is comparable to that in the OLS, we estimate the SUR on a truncated sample: those whose indivisible investments fall within the 5th and 95th percentiles at second endline. Because the application of the budget constraint means that equations are not estimated independently, we choose one outcome on which to truncate our sample, and indivisible investments have the highest degree of skewness.

The adding up condition allows us to precisely model the effect of the grant on each budget category: consumption, savings, divisible investment, indivisible investment, and net borrowing. If, for example, self-financing (through own savings, sale of existing assets for purchase of new assets, or a decrease in consumption) enables households to increase the resources available for investment beyond the cash grant, then we should see this reflected in those budget categories, with the cross-equation constraint ensuring that total expenditures do not exceed the available resources.

Short-Term Results (First Endline at 4 Months)

The OLS results indicate that, in the short run, small lottery winners statistically significantly increase their savings and divisible investment, while large lottery winners statistically significantly increase their savings, divisible investment, and indivisible investment (Table 2.8). We find no significant effect on consumption, income, or borrowing, for small or large lottery winners. The effect of winning the large lottery is statistically

²⁷Recall from Section 2.3.2 that we construct consumption as a stock of total spending between midline and the respective endline, plus home durables.

²⁸Thus, point estimates of the grant effect on divisible investment will be identical between Sections 2.5.1 and 2.5.2 but will vary for indivisible investment.

indistinguishable from the effect of winning the small lottery in all budget categories except indivisible assets, where large lottery winners differentially invest ($p < 0.05$). Specifically, small lottery winners have 57,000 UGX [22%] higher savings ($p < 0.10$) and 191,000 UGX [31%] higher divisible investment ($p < 0.01$), relative to non-winners who also chose the small lottery. Winning the large lottery increases savings by 165,000 UGX [61%] ($p < 0.01$), divisible investment by 319,000 UGX [40%] ($p < 0.05$), and indivisible investment by 4,295,000 UGX [25%] ($p < 0.05$), relative to non-winners who also chose the large lottery. (Note that the divisible investment results are identical to those discussed in the previous section, and the indivisible investment results change only due to the capital gains adjustment that allows us to analyze land values net of appreciation over time. We include the investment effects again here as an essential component of the household budget constraint.)

Before comparing the SUR and OLS results, we first note that neither the cross-equation constraint on the small nor large grant is rejected in the SUR. This indicates that the constraints fit the data reasonably well. We see that point estimates and precision are qualitatively similar between the OLS and SUR (Table 2.9), except in the indivisible investment category. The lottery effects on indivisible investment are substantially smaller in the SUR relative to the OLS, for both small and large lottery winners. As in the OLS, the SUR highlights the differential effect of winning the large lottery on indivisible investment. In the short run, the SUR does not reveal an obvious story of co-financing larger purchases through grant-induced changes in other budget categories, such as borrowing, savings, or income. This may indicate measurement error in indivisible investment, and the SUR mechanically reduces the effect on indivisible investment to satisfy the requirement that the budget constraint must hold. Alternatively, this could indicate rapid appreciation in land values even beyond the capital gains calculated in Section 2.3.2, or joint financing of purchases with friends or family not measured in the household budget. We will return to this issue in Section 2.5.3.

Medium-Term Results (Second Endline at 18 Months)

Winning the large lottery leads to increases in savings ($p < 0.05$), indivisible investment ($p < 0.05$), and consumption ($p < 0.10$) in the medium run, where the consumption effect only emerges after 18 months (Table 2.10). The short-run effects of both the small and large lottery on divisible investment fade — now indistinguishable from zero — suggesting that these assets may have been a mechanism for short-term savings. There is no statistically significant effect of winning the small lottery on any budget category, including consumption, savings, divisible and indivisible investment, income, and borrowing, in the medium run. There is no statistically significant effect of winning the large lottery on income or borrowing in the medium run, but large lottery winners do have 681,000 UGX [13%] higher spending on consumption and home durables ($p < 0.10$), 134,000 UGX [48%] higher savings ($p < 0.05$), and 5,276,000 UGX [30%] higher

indivisible investment ($p < 0.05$), relative to non-winners who also selected the large lottery.

Comparing the OLS results (Table 2.10) to the SUR results (Table 2.11) at second endline, we again see that neither cross-equation constraint is rejected. Like in the short run, the point estimates on indivisible investment are substantially lower in the SUR than in the OLS, and the SUR suggests that only large lottery winners sustain elevated levels of indivisible investment into the medium run ($p < 0.10$). At the medium run, the SUR implies that large lottery winners spent 1,279,000 UGX, or about 75% of the grant, on indivisible investment goods. This result aligns with the model's prediction. Savings also increases for the large lottery winners ($p < 0.05$), with a point estimate of 136,000 (SUR), or approximately 8% of the grant amount. This suggests that there may still be large lottery winners saving for a large investment, or that some winners may be smoothing the windfall in income over time. We do not detect any change in income, for small or large lottery winners, in the medium run. In the longer run, however, we will find substantial income gains for large lottery winners, as depicted in Table 2.12 and discussed next, in Section 2.5.2.

Long-Term Results (Third Endline at 6 Years)

We returned to survey the participants after 6 years in order to observe the evolution of their investments and incomes. We estimate the effects of the large and small lotteries using OLS.²⁹ We find a large and statistically significant ($p < 0.05$) increase in monthly income for those winning the large lottery. Their monthly incomes are now 114,000 UGX, or 32%, higher than the large lottery control. Large lottery winners also have 12,700 UGX, or 16%, higher consumption than the control group, though weekly consumption is measured imprecisely ($p = 0.12$). This suggests that about 50% of the increase in income is consumed.

In Figure 2.3, we use distribution regressions (e.g., [Chernozhukov, Fernández-Val, and Melly \(2013\)](#); [Goodman-Bacon \(2021\)](#)) to show how the treatment effect emerges over the income distribution at each endline, for both small and large lottery winners. The vertical dotted lines correspond to the 25th, 50th, and 75th percentiles. At the first endline, large lottery winners are not statistically significantly more likely to have income in the higher (or lower) part of the distribution than the large lottery control group. By the second endline, large lottery winners are about 5 percentage points more likely than the control to have income above 380,000 UGX (corresponding to the 75th percentile), but this is imprecisely estimated. Large lottery winners may also be less likely to have income in the middle of the distribution — perhaps indicating that they are now more likely to compose the upper end of the income distribution — but again, these effects on income are still imprecise at the second endline. By the third endline, however, large lottery winners are statistically significantly more likely (specifically, about 12 percentage points more likely) to have monthly

²⁹Because of the long time horizon between the lottery and the 6 year follow-up survey, the budget constraint for the SUR is unlikely to hold, given common fluctuation in incomes, and so we estimate the effects for the third endline solely using OLS.

income that exceeds 600,000 UGX (corresponding to the 80th percentile of the distribution). This suggests that large lottery winners disproportionately compose the upper end of the income distribution, as a result of receiving the large grant. While not statistically significant until about 600,000 UGX, the effect of the large grant on the probability of having income greater than X at the third endline is positive for all values of X that exceed 260,000 UGX, which is just over median income. In other words, the large lottery moved its recipients into the upper half of the income distribution, and the distribution regressions at each endline underscore that this effect emerges over time and only for the large lottery winners. At the third endline, small lottery winners do not disproportionately compose any part of the income distribution relative to the control.

We do not detect statistically significantly higher investment levels in the long run. There are several potential explanations for this increase in long-term income without a commensurate long-term increase in business or agricultural assets. Many of the durable investments measured at 18 months may continue to generate income for the participants, but they are likely to have depreciated significantly over 6 years. While the market value of the durable depreciated assets is quite low, the assets may continue to be productive assets in the business. Improvements made to land may have similarly depreciated, increasing long-term productivity while having little impact on overall land values. Temporary improvements made to businesses may have also yielded long-term clients who remain customers and have increased profitability.

2.5.3 Understanding the Large Effect on Land Values

The effect of the large grant on indivisible investments, driven by very large effects from land, far exceeds the size of the grant itself. This cannot fully be explained by commensurate changes in other aspects of the household budget that might have enabled co-financing — while there is some increase in income, the principal result of the SUR is simply to adjust the point estimates on indivisible investment downward (as opposed to an indication that levels of investment were co-financed by an increase in borrowing, spending down of savings, or sale of older assets due to the grant). These large point estimates on indivisible investment, and on land in particular, could reflect a number of things:

1. Unmeasured transfers from others that enable purchases or investment on existing land that exceed the immediate household's available resources (or unmeasured transfers from own savings or own assets that enable purchases of new land or investment on existing land exceeding the size of the grant)
2. Differential appreciation in land values for large grant winners due to putting land to a more productive use

3. Differential appreciation in land values in the areas where large lottery winners tend to purchase land, beyond the level of appreciation captured in the district's overall capital gains (for which we adjust down land values in Section 2.5.2)
4. Large grant winners who purchase new land or invest in their land also differentially overestimate their land value, perhaps due to a higher private valuation of the land
5. Large grant winners differentially report biased land values, irrespective of land purchases, perhaps due to the large grant and its potential uses inflating their belief about returns to the household's wealth in general

These explanations differ in their implications for the expansion of financial services. For example, grant winners making already owned plots of land more productive has distinct implications for aggregate growth relative to grant winners simply speculating in land and driving up its price, especially since land is in fixed supply.

Our data do not allow us to entirely differentiate between these explanations. Using the information that we have to partially disentangle these explanations, however, we examine the extent to which large grant winners made new land purchases versus investments in improving land (both of which could increase a household's land value). First, descriptive statistics in which we simply asked respondents directly how they spent the grant reveal that land purchases accounted for the highest fraction of grant spending for 32% of large grant winners, by far the modal expenditure among large grant winners (Table 2.F.1).³⁰ About 6% of small grant winners also report that purchasing land was their largest expenditure with the grant.³¹ 14% of large grant winners and 9% of small grant winners report that land and building improvements accounted for the largest fraction of their grant spending, which could have increased the value of their land as well. In sum, the descriptive statistics suggest that land values likely reflect a combination of new land purchases and land improvements, with land purchases accounting for the modal largest expenditure made with the large grant.

In our long-term (6 year) follow-up survey, we did ask respondents retrospective questions about land purchases, investment, and sales. In Table 2.13, we show that winning the large lottery differentially increases the likelihood of purchasing land by 15 percentage points ($p < 0.05$) relative to small lottery winners and by 16 percentage points ($p < 0.05$), or 48%, relative to the large lottery control. We find no statistically

³⁰Table 2.F.1 reports the single item on which a grant recipient spent the largest fraction of their grant funds, such that the categories are mutually exclusive. Many grant recipients, for example, may have purchased both business inventory and business durables, but the category on which they spent the largest fraction of their funds appears in Tables 2.F.1.

³¹Small grant winners also purchasing land is consistent with the model's predictions that some households whose baseline wealth is sufficiently close to the threshold above which they can make the indivisible investment will choose the small lottery, if winning the small lottery provides enough funds to surpass the threshold.

significant impacts on the likelihood of investing in land or selling land, for either small or large lottery winners. These results increase our confidence that the large increase in land values that we detect in the regions where the experiment takes place are associated with high rates of land purchase (note that even the control means suggest relatively high frequency of land purchase), which could lead to substantial price effects for an asset like land that is not perfectly elastically supplied. We will analyze the policy implications of an inelastically supplied investment good using our calibrated model in the following section.

2.6 Calibrated Model Results

We now return to our model of Section 2.2, but we use the income and savings distributions observed in our midline data and the households' choice of lottery to calibrate a model of financially-constrained indivisible investment and savings dynamics. It is common in the macro development literature that links experimental data with macro models to use the experimental data to either calibrate or cross-validate the model. In Appendix 2.G, we do the latter, showing that analogous regressions on model simulated data can reproduce the large coefficients on indivisible investment that we find in Section 5.2. Having done so, we use the model to quantitatively assess the impact of financial intermediation.

2.6.1 Calibration

Mapping the model to data for aggregate analysis requires adding important elements of persistent heterogeneity in the productivity of labor, $\bar{e}$, and capital, $\bar{z}$, and disciplining the other parameters of the model using data moments.³² Specifically, we assume $\bar{e}_i$ and $\bar{z}_i$ are (independently) lognormally distributed, so that $\log \bar{e}_i \sim \mathcal{N}(0, \nu_e)$ and $\log \bar{z}_i \sim \mathcal{N}(\mu_z, \nu_e)$, respectively. Here, assuming the mean of the $\log \bar{e}_i$ to be zero is a normalization.

Our calibration results and fit are summarized in Table 2.14. We start by assigning four parameters. First we set the maximum loan-to-value ratio $\theta = 0$ in the benchmark. We do so for simplicity, but this broadly matches the limited access to credit that we observe. The interest rate r then effectively becomes the interest on savings, which we set to zero, consistent with the fact that our empirical sample all were provided with zero interest savings accounts. A period is a month, and we set the monthly maintenance cost of capital, δ , which is analogous to depreciation in an indivisible capital model, to 0.01. We set the constant survival probability, p , to $1 - 1/480$, implying an average adult life (or career) of 40 years.

We are left with eight parameters to calibrate: the discount factor, β ; the parameters disciplining the

³²We examined alternative models with additional parameters such as non-unitary risk aversion, Epstein-Zin preferences (with separate risk aversion and intertemporal substitution parameters) and both physical and pecuniary fixed costs to adjusting capital, but these extensions do not significantly improve the fit over our parsimonious model.

capital technology, $\underline{k}$ (size of a unit of indivisible capital), α (the capital elasticity in production), μ_z , ν_z , and σ_z (governing the (log) mean productivity, persistent dispersion, and productivity shock dispersion, respectively); and the parameters disciplining the labor income: ν_e and σ_e (disciplining the dispersion of the persistent income difference and productivity shocks, respectively).

Our calibration strategy is to choose these parameters to match the the moments in the data that are most of interest to this study: the observed lottery choices, the distribution and dynamics of income, and the distribution of wealth. For the lotteries, we target the percentage of households who choose each of the four lottery choices. For the distributions of both income and savings, we target the mean as well as five percentiles of the distribution (10th, 25th, 50th, 75th, and 90th), a total of 11 moments given the normalization of median income to one in the model and in the data.³³ Finally, for the dynamics of income, we target transitions across income terciles over the 16 months from baseline to midline. The 3x3 transition matrix constitutes nine additional moments. Together, these targets amount to 22 moments, substantially more than the eight free parameters we calibrate.

Given the parsimony of the model, we cannot match all moments, and instead minimize a weighted average of the sum of percentage squared deviations.³⁴ Table 2.14 summarizes the parameter values and model fit. Despite the over-identification, the model fit is reasonably close, especially for our purposes. The model is able to match the fact that a substantial share choose the large lottery (0.20+0.04=0.24 in the model vs. 0.20+0.07=0.27 in the data). The share choosing to delay is somewhat smaller (0.04 vs. 0.07, respectively), but overall the share of those delaying is also comparable (0.04+0.19=0.23 vs. 0.07+0.14=0.21, respectively). The savings distributions are comparable with median levels of savings well below median monthly income (0.13 vs. 0.19, respectively). Both distributions exhibit a leftward skew with the means exceeding the medians, and at least 25% of households in both distributions hold no positive savings. Of course, given the fact that our benchmark simulations have no credit, we cannot match the negative assets for those in the 10th percentile, but recalling Figure 2.2, this distinction is likely not crucial for our aggregate analysis. Although the right tail is lower in the model than in the data (e.g., 90th percentiles of 1.43 vs. 2.92), the mean is somewhat higher given the lack of debtors in the model.

We see that the income distributions are also broadly comparable. Though the mean income is somewhat higher in the model (1.97 vs. 1.73), due to somewhat higher weight in the tails, both distributions have a leftward skew with mean substantially higher than the median. Looking at the income transition matrix, both the model and data exhibit strong diagonal elements, showing a tendency to stay in one's income tercile, especially in the upper and lower terciles. Conditional on being in the lower income tercile, a household has

³³Savings includes only financial wealth and corresponds to S in the model.

³⁴The weights are equal to one, except for the lottery choice moments, which are critical in getting the indivisibility dynamics and so get a higher weight of 20.

a high chance of remaining there (67% vs. 63%), which could be evidence of a (stochastic) poverty trap. In the middle tercile, the numbers are lower (43% vs. 48%), reflecting, in the model, the transitional saving dynamics of moving to one of the stable steady states (recall Figure 2.1).

Looking at the calibrated parameter values, we see that $\underline{k} = 8.0$ implies that the indivisible investment is about 8 months of the median family income, or 40 times larger than the median level of liquid savings. The estimates of $\sigma_\epsilon = 0.53$ and $\sigma_z = 1.13$ indicate considerable uncertainty in labor and capital income. The first creates both variation and churning in the income distribution, while the latter simply creates variation. Finally, we note that the discount factor, $\beta = 0.84$, is quite low, much lower than typically calibrated for macroeconomic studies. Models with financial frictions often yield very low discount factors, however, which is needed to prevent people from simply saving out of the range of financial frictions (e.g., [Buera, Kaboski, and Shin \(2020\)](#)). Note that in this calibration, the presence of high-yield investment opportunities will generate patience among a world of impatient agents rather than impatience among the otherwise patient.

2.6.2 Counterfactual Assessment of Introduction of Financial Services

We now use the model to evaluate the impacts of introducing wide-scale financial services to the regions of study. We assess four different “policies” relative to our benchmark. For each policy, we assess these policies under two different scenarios about the supply of capital. The first is a scenario where capital is perfectly elastically supplied at a constant price, as is typically assumed in macro models of investment. In this case, only quantities will adjust to an increase in the demand for investment. Such an assumption is surely appropriate when the intervention is small relative to overall supply of capital, and so we label this “PE”. The second scenario is one in which capital is in fixed supply. Such a scenario is an extreme scenario but perhaps more appropriate for the case when land is the investment good. In this case, prices will adjust to changes in the demand for capital rather than aggregate quantities, and so we label this “GE”. We note, however, the purpose of both scenarios is to use the extreme cases to bracket possible impacts.

Table 2.15 presents the results for the various policies under both scenarios. We focus on the impacts of these policies on aggregates: income, consumption, net savings, capital, the value (price $\times$ quantity) of capital, the price of capital, and capital income. Given our interest in poverty traps and wealth dynamics, we also report the fraction of the poor in the economy (i.e., the fraction below a threshold defined by the lowest tercile of *wealth* in the benchmark economy) and the probability of staying poor. We report all of the counterfactual results relative to the benchmark values (which are given in the notes of Table 2.15 but not otherwise of particular interest). We point out that although savings and capital are positive, consumption approximately equals income in the steady state because capital does not depreciate and income

is already net of maintenance costs. That is, steady state investment is zero in the economy whether capital is elastically or inelastically supplied. The only difference between consumption and income comes from the small probability of death.

The first “policy” we evaluate is a ‘Divisible Capital’ case in which capital is perfectly divisible (though still bounded above) enabling a choice of any $k \in [0, k_1]$. This might be a viable policy if indivisibility in land, for example, were driven by the titling of plots or rental markets frictions. (However, our qualitative research suggests that titles are *not* the driving characteristic of land indivisibility.) Alternatively, this exercise can be understood as providing a benchmark comparison for how powerful indivisibility in investment can be in driving economic aggregates and poverty in order to answer the question in the paper’s title. Comparing the first and second columns, we see that indivisibility substantially lowers output. In the PE case, output would be more than twice as high without the indivisibility, and capital would be nearly three times as high, since even the poor-but-productive would be able to invest some, and gradually expand their capital over time. The GE column shows large differences in aggregate consumption and income as well, even though the aggregate capital stock is fixed, and these gains come from a redistribution toward households with a higher marginal product of capital on the margin.³⁵ The consequences are that no one is below the benchmark poverty line in this divisible capital world. In the GE case of divisible capital, the savings rate goes down substantially as the motive to self-finance the indivisible capital is eliminated.

We now turn to three financial policies which we compare to these two benchmarks. The first is an expansion of credit services that enables households to borrow up to (a relatively modest) 25% of the value of their capital, which we model as an increase of the collateral parameter, θ , from 0 to 0.25. The second is an expansion of interest-bearing savings facilities, which we model as an increase in the monthly interest rate, r , from 0 to 0.03, equal to the “delay” interest offered in our empirical experiment. The third is a combination of both, which enables borrowing together with saving but both at a positive interest rate.

The pure credit intervention is quite powerful when the price of investment is fixed. Under this scenario (‘Credit’ and ‘PE’ in the table), the 2.30 indicates that the capital stock is 130% larger while the 1.53 indicate that income is 53% larger. Capital income rises slightly more due to a better allocation of capital. The 0.91 indicates that net savings declines, a combination of both credit and decreased savings for self-financing reasons, but by only 9%. The fraction poor declines by 32%, although the chances of staying poor remain the same.

In striking contrast, in the case of credit under a fixed stock of capital, we see absolutely no impact on aggregate income, capital income, or poverty. Indeed, the policy is completely neutral even on the

³⁵Note that our calibrated elasticity of capital $\alpha = 0.35$ is consistent with capital’s share in other quantitative work, but implies an Inada condition, where the marginal products of capital are unbounded as capital approaches zero.

micro level. The change in θ is completely offset by a commensurate change in price, where $P = 1/(1 - \theta)$ or 1.33 in this case. Net savings declines as households simply keep their surplus savings above available credit, i.e., $\hat{a} = a' + \theta Pk'$, unchanged. Indeed, given our assumptions, one can easily show this neutrality result analytically.³⁶ The key point is that the power of credit to increase output and reduce poverty is substantially reduced in a world where the increased demand for investment manifests itself more in an increase in prices than through an increase in productive capital.

The results for the pure savings intervention demonstrate that when the investment price is fixed, interest-bearing savings can also be a powerful financial intervention. Under a fixed price of capital, savings responds considerably to the positive return on savings, increasing by 37%, a strong contrast to the impact of the credit policy. However, capital increases by 20%, capital income increases by 23%, and overall income rises by 11%. The impact of savings on poverty is perhaps as dramatic as in the credit intervention, as the fraction poor drops by 18% and the probability of staying poor is 14% lower in the savings distribution. The “poverty trap” is therefore weakened.

When we look at the case of capital in fixed supply, although the perfect neutrality result no longer holds, the aggregate impacts are nonetheless tempered. The price of capital increases but by only 6%, showing that the savings intervention doesn’t increase the demand for investment quite as much as the credit program. Income rises by only 2%, as does capital income. Here labor income is fixed, but interest becomes an important source of income. Savings rises by 20%, and both higher savings and a positive interest rate lead to interest income.³⁷ Hence, while the aggregate results are tempered, they are not completely neutral. Moreover, because it affects the accumulation of savings, the program still has important impacts on poverty. The fraction poor drops by 14% even under a fixed supply of land, and poverty is again much less persistent with the probability of staying poor also dropping by 14%.

Lastly, we see that the combination of interest-bearing savings and interest-paying credit can be an even stronger financial intervention when the price of capital is fixed. Indeed, the 74% increase in consumption and 72% in income represents more than half of the distance between the benchmark and the divisible capital case, and the 50% reduction in poverty is midway between the two as well. The persistence of poverty is

³⁶To see this, recall that θ only impacts the household’s problem through the borrowing constraint, $a' \geq -\theta Pk'$. Define $\hat{a} = a' + \theta Pk'$ as surplus savings above the lower bound. Then combining the recursive constraints (2.2) and (2.4) to eliminate l , substituting in for $\hat{a}$, and simplifying using $r = 0$ yields the following constraints:

$$c + \hat{a} + P(1 - \theta)(k' - k) = e + \hat{a} + zk^\alpha - \delta k$$

and

$$\hat{a} \geq 0.$$

It is clear that any change from θ to θ' is simply offset by a general equilibrium change in the price of capital from P to $P' = P(1 - \theta)/(1 - \theta')$, and no household’s problem is affected, except that while $\hat{a}$ remains constant, the 0.38 indicates that a' declines dramatically (by 62%) to offset the change in available credit according to the equation that holds by definition, $\hat{a} = a' + \theta Pk'$. The clarity of this analytical neutrality result is further justification for the $r = 0$ assumption.

³⁷Our neutrality result for credit hinged on zero interest loans. In the case where loans were positive interest there would be distributional consequences as well.

comparable to the case with only savings, however. In the case of capital in fixed supply, however, even the combined impacts of financial services remain muted. Net savings is actually lower than in the benchmark (although it is higher than the benchmark when the price of capital is fixed). The price of capital is higher when both credit and savings possibilities increase the price of capital, but poverty and its persistence is comparable to the savings only scenario.

In summary, the aggregate simulation results underscore that the indivisibility of capital has strong impacts on aggregate development and poverty. Financial interventions can be powerful forces for development and poverty reduction in this world of indivisible investments, but this hinges critically on the elasticity of the supply of capital. General equilibrium forces can undermine positive aggregate impacts measured in partial equilibrium experiments.

Given the importance of land as an indivisible investment in the empirics, a final question is whether we ought to think about peri-urban land in fixed or elastic supply. One reasonable answer is to simply assert that land captures simple area, and is in therefore in fixed supply. If the increased land value we capture reflects land purchases, so even the small-scale increases in land value among the treated come from the sales of land of the untreated, a quantitatively small but in this case economically important violation of the stable unit treatment value assumption (SUTVA) for our control. An alternative to consider is that land value, especially quality-adjusted land, is potentially elastic, however, reflecting either investments/improvements in land that increase the value and/or a true expansion of utilized land from around the perimeter of a town, for example. Following such an assumption, we can use the exogenous increase in demand of land from our experiment to calculate a back-of-the-envelope estimate of the elasticity of the supply of land. Specifically, using the results in Table 2.7 we estimate a percentage increase in the quantity of quality-adjusted land (calculated to be roughly 1% as a fraction of all land, participants and not). Recalling the price effects estimated in Appendix 2.B, which implied an increase in price of 20%, we estimate a very small elasticity of $1\%/20\% = 0.05$, quite close to the inelastic benchmark, indicating the GE results are most relevant.³⁸

2.7 Conclusions

We examine the importance of high-yield, indivisible investments in peri-urban and rural Uganda, where financial services are limited. Empirically, we have shown demand for a large-stakes-but-lower-expected-value lottery, especially among those who have self-financing motives for savings and investment motives for credit, even among a risk-averse population. The risk-loving behavior we observe is consistent with theories of high-yield indivisible investments, which, in principle, can lead to sizable gains to financial intermediation.

³⁸The specifics of these calculations are also included in Appendix B.

Winners of these large lotteries invest disproportionately in indivisible assets, land in particular. We find sizable income gains for winners of the large lottery which take time to develop — they are present at 6 years but not at 4 or 18 months (though there is some evidence of consumption increases for large lottery winners at 18 months). Our theory demonstrates that when indivisible investments are inelastically supplied, the impacts of expanding financial services, especially credit, on aggregate outcomes and poverty can be substantially lower. From a policy perspective, the results indicate that interest-bearing savings, perhaps along the lines of [Greaney, Kaboski, and Van Leemput \(2016\)](#), may be relatively more important when the investment good is in fixed supply.

Substantively, the findings about land elicit questions of their own for future research and highlight the importance of work in peri-urban environments which are growing quickly and often have lagging access to infrastructure. Poverty trap dynamics have been observed for smaller investments such as livestock, which are profitable and more elastically supplied (e.g., [Balboni et al. \(2022\)](#)), but the importance of land merits more consideration. The development literature has focused on the interaction of finance, land titling, and investment, and also the link between land plots, overall farm size, and misallocation (e.g., [Foster and Rosenzweig \(2022\)](#); [Gollin and Udry \(2021\)](#); [Acampora, Casaburi, and Willis \(2022\)](#)), but the issues we raise are unique. How common are land-driven poverty traps and to what extent does land reallocation increase productivity and income? Finally, our results indicate that, even outside of major urban areas, land may be an important investment for the poor. The lack of clear evidence on whether returns to land result from investment in the land by more productive households or simply as the result of holding a good that offers high capital gains is an important limitation of the findings. More research on these aspects of land markets is therefore encouraged.

2.8 Figures

Figure 2.1: Asset-Dependent Behavior and Poverty Traps

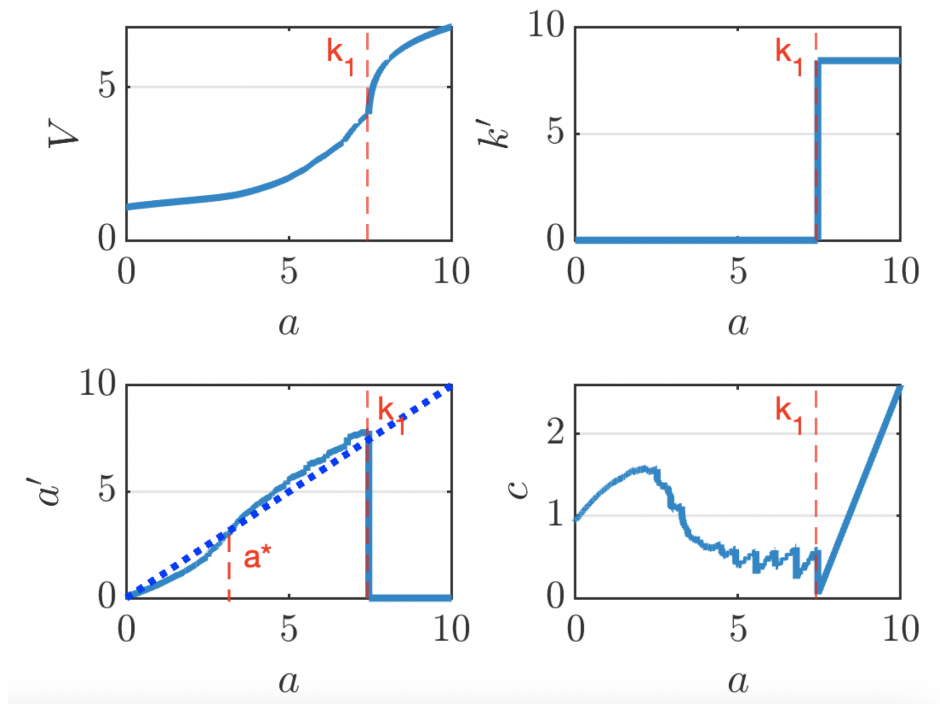


Figure 2.2: Asset-Dependence of Lottery Choice

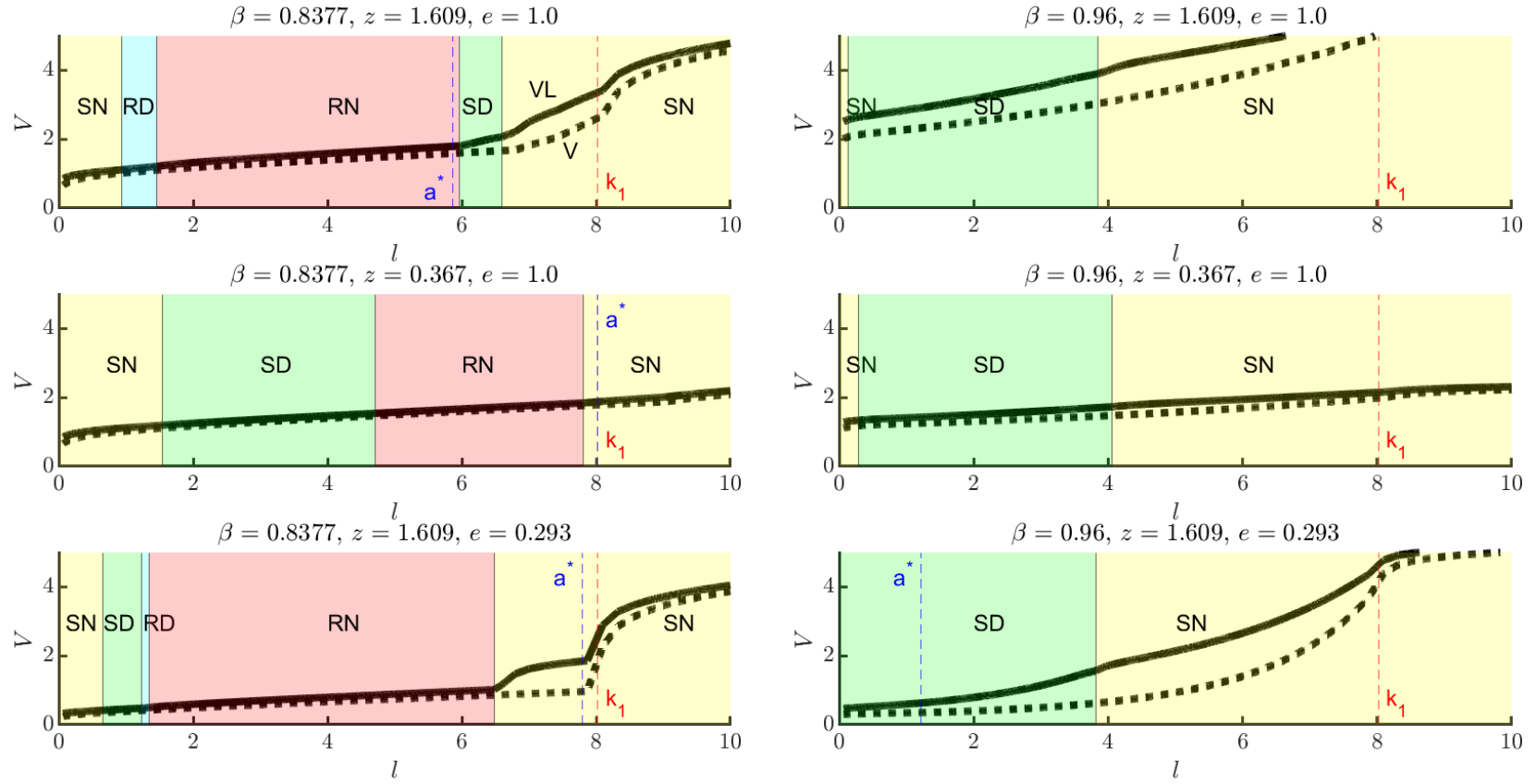
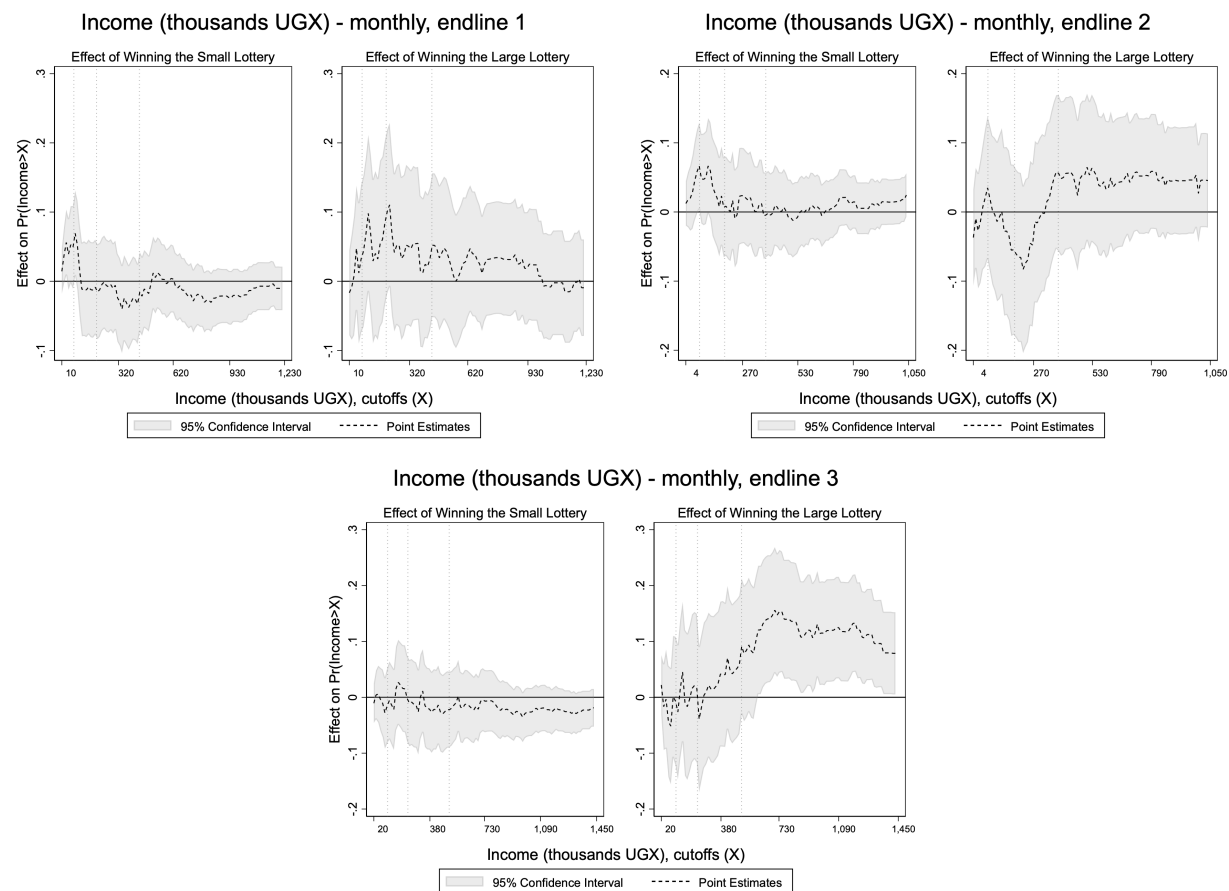


Figure 2.3: Distribution Regressions: Effect of Winning the Lottery on the Distributions of Income at Each Endline



Notes: Vertical dotted lines correspond to the 25th, 50th, and 75th percentiles of the outcome, respectively.

2.9 Tables

Table 2.1: Lottery choices

	Large	Small	Total
Patient	78 (7%)	144 (14%)	222 (21%)
Impatient	205 (20%)	621 (59%)	826 (79%)
Total	283 (27%)	765 (73%)	1,048 (100%)

Table 2.2: Lottery winners

	Large	Small	Total
Patient	19 of 78 = 24%	71 of 144 = 49%	90 of 222 = 40%
Impatient	66 of 205 = 32%	302 of 621 = 49%	368 of 826 = 45%
Total	85 of 283 = 30%	373 of 765 = 49%	458 of 1,048 = 44%

Table 2.3: Balance between small lottery winners and non-winners

	won small lottery	N	did not win small lottery	N	diff	p-value
<i>Household budget components</i>						
mtlly income _m	362,779	325	385,259	309	-22,480	0.44
wkly consumption _m	40,315	325	39,451	309	864	0.70
savings _m	289,452	325	268,359	309	21,093	0.51
credit outstanding _m	184,105	325	204,933	309	-20,828	0.51
home durable value _b	507,805	325	501,452	309	6,353	0.92
<i>Investment categories</i>						
total divisible investments _m	409,608	325	342,147	309	67,461	0.16
small livestock and ag. assets _b	176,231	325	161,210	309	15,020	0.41
bus inventory _m	170,068	325	150,846	309	19,222	0.53
total indivisible investments _m	11,184,306	325	11,166,890	309	17,416	0.99
large livestock and ag. assets _b	285,815	325	237,133	309	48,683	0.47
land value _b	9,920,154	325	10,091,586	309	-171,432	0.87
bus assets, no stock _m	405,198	325	286,228	309	118,970*	0.10
<i>Other financial indicators</i>						
operates non-farm business _m (0/1)	.53	325	.55	309	-.021	0.60
farmer _m (0/1)	.75	325	.77	309	-.019	0.57
work hours per week _m	77	325	78	309	-.98	0.58
had negative shock since baseline _m (0/1)	.64	325	.68	309	-.037	0.33
has formal savings _m (0/1)	.11	325	.15	309	-.035	0.19
acquired loans since baseline _m (0/1)	.32	325	.33	309	-.007	0.85
<i>Demographic characteristics</i>						
female (0/1)	.47	325	.52	309	-.05	0.20
household head (0/1)	.61	325	.6	309	.0042	0.91
respondent age	35	325	37	309	-1.2	0.18
education beyond primary school (0/1)	.27	325	.28	309	-.011	0.76
num people in household _b	5	325	5	309	.041	0.82
num adult females _b	1.1	325	1.2	309	-.076*	0.09
num adult males _b	1.4	325	1.4	309	.039	0.64
num children _b	2.5	325	2.4	309	.078	0.55
Observations	634					

All quantities in UGX; Outliers top/bottom coded to 95th/5th percentile; Midline denoted by **m** and baseline denoted by **b**
Sample consists of those retained through the second endline survey (those used in our analysis of treatment effects)

bus is an abbreviation for business; * $p < 0.1$, ** $p < .05$, *** $p < 0.01$

Table 2.4: Balance between large lottery winners and non-winners

	won large lottery	N	did not win large lottery	N	diff	p-value
<i>Household budget components</i>						
mtlhy income _m	423,464	74	378,105	159	45,359	0.40
wkly consumption _m	46,428	74	43,862	159	2,566	0.54
savings _m	353,284	74	281,918	159	71,366	0.19
credit outstanding _m	169,459	74	204,522	159	-35,063	0.54
home durable value _b	676,441	74	743,906	159	-67,465	0.62
<i>Investment categories</i>						
total divisible investments _m	642,345	74	458,409	159	183,936*	0.08
small livestock and ag. assets _b	199,385	74	197,148	159	2,237	0.95
bus inventory _m	301,149	74	201,572	159	99,576	0.13
total indivisible investments _m	18,325,135	74	14,950,233	159	3,374,902	0.17
large livestock and ag. assets _b	590,811	74	626,132	159	-35,321	0.85
land value _b	15,665,270	74	12,821,698	159	2,843,572	0.21
bus assets, no stock _m	540,676	74	498,157	159	42,518	0.78
<i>Other financial indicators</i>						
operates non-farm business _m (0/1)	.61	74	.61	159	-.002	0.98
farmer _m (0/1)	.73	74	.71	159	.019	0.77
work hours per week _m	.77	74	.77	159	-.42	0.90
had negative shock since baseline _m (0/1)	.64	74	.62	159	.012	0.86
has formal savings _m (0/1)	.054	74	.094	159	-.04	0.30
acquired loans since baseline _m (0/1)	.23	74	.33	159	-.097	0.13
<i>Demographic characteristics</i>						
female (0/1)	.38	74	.43	159	-.056	0.43
household head (0/1)	.73	74	.63	159	.1	0.13
respondent age	.38	74	.37	159	.1	0.53
education beyond primary school (0/1)	.31	74	.23	159	.078	0.21
num people in household _b	5.4	74	5.4	159	-.0097	0.98
num adult females _b	1.1	74	1.1	159	.042	0.64
num adult males _b	1.6	74	1.4	159	.19	0.26
num children _b	2.6	74	2.9	159	-.25	0.34
Observations	233					

All quantities in UGX; Outliers top/bottom coded to 95th/5th percentile; Midline denoted by **m** and baseline denoted by **b**
Sample consists of those retained through the second endline survey (those used in our analysis of treatment effects)

bus is an abbreviation for business; * $p < 0.1$, ** $p < .05$, *** $p < 0.01$

Table 2.5: Statistically significant differences between those choosing the large v. small lottery

	large	N	small	N	diff	p-value
<i>Income and consumption</i>						
$\Delta \ln \text{mthly income}_{m-b}^*$	.71	283	1.2	765	-.44	0.02
mthly crop income _m [*]	100,208	283	71,148	765	29,061	0.00
crop income/total income _m	.3	283	.25	765	.05	0.01
mthly crop income/adult equiv _m [*]	30,562	283	22,981	765	7,580	0.00
$\ln \text{mthly crop income}_m^*$	12	283	12	765	.21	0.00
$\Delta \ln \text{mthly bus income}_{m-b}^*$	6.7	283	7.6	765	-.87	0.02
wkly consumption _m	44,466	283	39,453	765	5,013	0.01
$\ln \text{wkly cons}_m$	10	283	10	765	.18	0.01
$\Delta \ln \text{wkly cons}_{m-b}^*$	.38	283	.24	765	.14	0.03
<i>Savings and wealth</i>						
savings _m	322,675	283	275,817	765	46,858	0.10
$\Delta \ln \text{savings}_{m-b}$	2.5	283	1.7	765	.86	0.03
bus assets _m	824,954	283	577,814	765	247,140	0.01
bus assets/wealth _m	.29	283	.24	765	.053	0.05
bus assets/adult equiv _m	295,958	283	218,673	765	77,285	0.07
$\ln \text{bus assets}_m$	5.3	283	4.3	765	.99	0.03
wealth (sav + bus assets) _m	1,245,155	283	920,185	765	324,970	0.01
wealth/adult equiv _m	431,561	283	337,778	765	93,784	0.07
$\ln \text{wealth}_m$	11	283	11	765	.65	0.05
$\Delta \ln \text{wealth}_{m-b}^*$	2.2	283	1.5	765	.7	0.05
net wealth (sav + bus assets - credit) _m	1,070,910	283	744,418	765	326,492	0.01
net wealth/adult equiv _m	377,813	283	277,547	765	100,265	0.05
$\ln \text{net wealth}_m$	16	283	16	765	.075	0.06
land value _b	13,345,159	283	9,852,680	765	3,492,479	0.00
land value/adult equiv _b	3,936,519	283	3,010,483	765	926,036	0.00
$\ln \text{land value}_b$	13	283	12	765	.91	0.04
<i>Desire to invest</i>						
wants credit to increase income _b (0/1) [*]	.84	283	.78	765	.062	0.03
would invest >\$100 _b (0/1) [*]	.95	283	.91	765	.038	0.04
would use credit for bus investment _b (0/1) [*]	.67	283	.59	765	.073	0.03
<i>Demographic characteristics</i>						
female (0/1) [*]	.42	283	.51	765	-.09	0.01
household head (0/1)	.66	283	.6	765	.058	0.09
respondent age	37	283	35	765	2.3	0.00
num people in household _b	5.5	283	5	765	.47	0.00
num adult males _b	1.5	283	1.4	765	.14	0.08
num children _b [*]	2.8	283	2.5	765	.34	0.00

*Denotes that variable was also selected by lasso. All quantities in UGX. Outliers top/bottom coded to 95th/5th percentile. Midline denoted by **m** and baseline denoted by **b**; **bus** is an abbreviation for business. Full list of covariates on which we test for differences between those selecting the small versus large lottery is in Table 2.C.1. Complete list of lasso-selected covariates is in Table 2.C.2.

Table 2.6: Grant effects on investment - First endline

	Productive divisible investments				Productive indivisible investments			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Home durables	Tot. div.	Small livestock and ag. tools	Bus. inventory	Tot. indiv.	Large livestock and ploughs	Bus. durables	Land
won lottery (0/1)	91,888 (62,914)	190,956*** (61,827)	84,271*** (27,645)	127,997*** (40,916)	-796,104 (954,012)	90,079* (49,229)	65,295 (70,324)	-861,540 (982,476)
won large lottery (0/1)	-67,114 (130,254)	127,754 (151,638)	-76,874 (63,475)	174,088* (102,373)	5,095,876** (2,075,135)	124,565 (118,938)	122,533 (172,279)	5,315,274** (2,127,186)
risk loving (0/1)	4,945 (78,014)	76,005 (75,969)	55,439 (34,643)	-1,218 (49,342)	-2,438,998** (1,041,186)	-41,153 (66,155)	27,890 (91,002)	-1,958,856* (1,055,963)
$\beta_1 + \beta_2$	24,774	318,710	7,397	302,085	4,299,771	214,644	187,828	4,453,734
P-value: $\beta_1 + \beta_2 = 0$	.83	.02	.9	.0012	.019	.048	.23	.017
Control mean if risk loving = 0	628,817	613,887	275,969	285,471	16,750,624	202,913	537,380	15,675,049
Control mean if risk loving = 1	748,134	787,251	369,204	320,723	19,021,050	308,805	748,535	17,415,000
R ²	.28	.41	.22	.46	.72	.31	.5	.69
Observations	867	867	867	867	867	867	867	867

All quantities in UGX; Outliers top/bottom coded to 95th/5th percentile; Heteroskedasticity-robust standard errors in parantheses

Controls include: pre-treatment levels of outcome, hh income, patience, gender, hh head, age, age², num ad females, num ad males, num children, district FE's

Table 2.7: Grant effects on investment - Second endline

	Productive divisible investments				Productive indivisible investments			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Home durables	Tot. div.	Small livestock and ag. tools	Bus. inventory	Tot. indiv.	Large livestock and ploughs	Bus. durables	Land
won lottery (0/1)	4,120 (71,454)	69,303 (60,111)	33,652 (27,418)	55,353 (42,651)	644,569 (1,487,327)	91,122* (51,760)	129,478 (113,988)	319,283 (1,494,017)
won large lottery (0/1)	13,266 (151,702)	-27,480 (138,689)	-20,836 (62,223)	16,520 (101,933)	6,993,895* (3,576,218)	-89,363 (121,789)	384,424 (294,705)	7,168,289** (3,493,766)
risk loving (0/1)	-34,722 (87,422)	86,905 (76,456)	36,522 (36,753)	20,325 (52,175)	-1,514,346 (1,743,776)	83,579 (77,256)	-14,138 (140,530)	-1,325,532 (1,722,529)
$\beta_1 + \beta_2$	17,386	41,822	12,816	71,873	7,638,464	1,759	513,902	7,487,571
P-value: $\beta_1 + \beta_2 = 0$	.9	.74	.82	.43	.019	.99	.057	.018
Control mean if risk loving = 0	725,261	596,244	280,831	286,676	21,386,603	181,553	703,200	20,327,217
Control mean if risk loving = 1	819,181	776,076	344,160	361,808	24,606,028	398,129	918,528	22,733,019
R ²	.26	.32	.13	.37	.52	.2	.36	.5
Observations	867	867	867	867	867	867	867	867

All quantities in UGX; Outliers top/bottom coded to 95th/5th percentile; Heteroskedasticity-robust standard errors in parantheses

Controls include: pre-treatment levels of outcome, hh income, patience, gender, hh head, age, age², num ad females, num ad males, num children, district FE's

Table 2.8: Grant effects on components of the household budget constraint - First endline

	(1) consum. (stock) wkly cons + home dur.	(2) savings	(3) div. investment	(4) indiv. investment*	(5) mthly income	(6) net credit
won lottery (0/1)	121,277 (90,290)	57,204* (30,408)	190,956*** (61,827)	-721,095 (906,426)	-8,344 (20,200)	-7,054 (18,092)
won large lottery (0/1)	-18,497 (176,415)	107,836 (69,336)	127,754 (151,638)	5,016,354** (1,965,752)	38,551 (44,173)	-23,513 (40,044)
risk loving (0/1)	110,406 (109,121)	12,933 (35,395)	76,005 (75,969)	-2,326,259** (966,582)	-4,074 (25,967)	-788 (22,591)
$\beta_1 + \beta_2$	102,781	165,040	318,710	4,295,259	30,207	-30,567
P-value: $\beta_1 + \beta_2 = 0$	.5	.0082	.02	.013	.44	.39
Control mean if risk loving = 0	1,608,566	259,468	613,887	15,232,971	327,076	73,180
Control mean if risk loving = 1	1,798,491	271,428	787,251	17,160,718	321,271	80,220
R ²	.29	.3	.41	.71	.42	.093
Observations	867	867	867	867	867	867

All quantities in UGX; Outliers top/bottom coded to 95th/5th percentile; Heteroskedasticity-robust standard errors in parantheses

*Total indivisible investment includes real land values, adjusted down to exclude appreciation in land values over time

Controls include: pre-treatment levels of outcome, hh income, patience, gender, hh head, age, age², num ad females, num ad males, num children, district FE's

Table 2.9: SUR with household budget constraint - First endline

	consum. (stock) wkly cons + home dur.	savings	div. investment	indiv. investment*	mtlly income	net credit
won lottery (0/1)	83,017 (93,281)	54,390* (32,223)	177,950*** (66,585)	-78,531 (132,562)	-26,015 (23,304)	-9,112 (19,249)
won large lottery (0/1)	-32,121 (195,226)	139,176** (67,492)	92,247 (139,403)	1,247,456*** (277,484)	29,663 (48,801)	-21,893 (40,313)
risk loving (0/1)	201,784* (118,592)	1,438 (40,557)	145,056* (84,223)	669,053 (1,189,239)	-4,107 (29,304)	-6,831 (24,071)
district fe's	Yes	Yes	Yes	Yes	Yes	Yes
demographic controls	Yes	Yes	Yes	Yes	Yes	Yes
<i>Within-equation tests:</i>						
$\beta_1 + \beta_2$	50,897	193,566	270,197	1,168,925	3,647	-31,004
P-value: $\beta_1 + \beta_2 = 0$	.77	.00	.03	.00	.93	.38
<i>Cross-equation tests:</i>						
P-value: Sm grant constraint	.95					
P-value: Lg grant constraint	.65					
Control mean if risk loving = 0	1,573,710	256,157	591,676	15,431,617	318,218	77,590
Control mean if risk loving = 1	1,839,073	265,070	818,792	17,691,141	328,638	70,739
R ²	.2	.25	.35	.42	.28	.073
Observations	783					

All quantities in UGX; Outliers top/bottom coded to 95th/5th percentile; Estimated with cross-equation budget constraint

Demographic controls include: pre-treatment levels of outcome, hh income, patience, gender, hh head, age, age², num ad females, num ad males, num children

Table 2.10: Grant effects on components of the household budget constraint - Second endline

	(1) consum. (stock) wkly cons + home dur.	(2) savings	(3) div. investment	(4) indiv. investment*	(5) mthly income	(6) net credit
won lottery (0/1)	294,543 (249,593)	48,839 (33,020)	69,303 (60,111)	506,277 (1,050,417)	16,240 (18,897)	-20,703 (23,852)
won large lottery (0/1)	386,636 (472,079)	84,891 (70,157)	-27,480 (138,689)	4,770,034* (2,493,362)	7,896 (41,709)	8,687 (51,953)
risk loving (0/1)	157,957 (294,962)	3,760 (38,289)	86,905 (76,456)	-1,002,692 (1,220,197)	27,946 (24,040)	-44,824 (30,255)
$\beta_1 + \beta_2$	681,180	133,730	41,822	5,276,311	24,136	-12,016
P-value: $\beta_1 + \beta_2 = 0$	.091	.033	.74	.02	.51	.79
Control mean if risk loving = 0	4,603,540	273,166	596,244	15,040,343	264,233	119,272
Control mean if risk loving = 1	5,036,368	279,987	776,076	17,326,575	289,215	76,352
R ²	.28	.23	.32	.51	.29	.069
Observations	867	867	867	867	867	867

All quantities in UGX; Outliers top/bottom coded to 95th/5th percentile; Heteroskedasticity-robust standard errors in parantheses

*Total indivisible investment includes real land values, adjusted down to exclude appreciation in land values over time

Controls include: pre-treatment levels of outcome, hh income, patience, gender, hh head, age, age², num ad females, num ad males, num children, district FE's

Table 2.11: SUR with household budget constraint - Second endline

	consum. (stock) wkly cons + home dur.	savings	div. investment	indiv. investment*	mothly income	net credit
won lottery (0/1)	21,291 (255,323)	46,674 (32,975)	24,199 (63,451)	333,714 (384,449)	4,314 (20,405)	-1,781 (24,813)
won large lottery (0/1)	342,344 (534,416)	89,967 (69,069)	-68,503 (132,841)	945,661 (804,880)	-824 (42,729)	-25,699 (51,965)
risk loving (0/1)	235,648 (328,117)	7,752 (41,662)	143,697* (80,082)	2,265,515 (1,545,600)	34,987 (25,483)	-22,269 (31,021)
district fe's	Yes	Yes	Yes	Yes	Yes	Yes
demographic controls	Yes	Yes	Yes	Yes	Yes	Yes
<i>Within-equation tests:</i>						
$\beta_1 + \beta_2$	363,635	136,641	-44,305	1,279,375	3,490	-27,480
P-value: $\beta_1 + \beta_2 = 0$	.44	.02	.70	.07	.93	.55
<i>Cross-equation tests:</i>						
P-value: Sm grant constraint	.52					
P-value: Lg grant constraint	.53					
Control mean if risk loving = 0	4,597,112	253,897	597,332	18,669,532	250,716	108,971
Control mean if risk loving = 1	5,049,384	274,655	822,299	22,488,440	293,904	83,239
R ²	.2	.2	.28	.29	.2	.058
Observations	783					

All quantities in UGX; Outliers top/bottom coded to 95th/5th percentile; Estimated with cross-equation budget constraint

Demographic controls include: pre-treatment levels of outcome, hh income, patience, gender, hh head, age, age², num ad females, num ad males, num children

Table 2.12: Grant effects on components of the household budget constraint - Third endline

	(1) weekly consumption	(2) home durables	(3) savings	(4) div. investment	(5) indiv. investment*	(6) mthly income
won lottery (0/1)	-818 (4,387)	-2,455 (113,772)	19,243 (54,942)	15,687 (117,392)	570,089 (1,080,048)	-15,468 (28,595)
won large lottery (0/1)	13,470 (9,130)	-80,403 (224,349)	32,624 (138,951)	16,592 (257,404)	-563,232 (2,357,928)	129,367** (62,467)
risk loving (0/1)	4,318 (5,536)	14,317 (141,031)	175,644** (76,641)	95,343 (147,478)	2,127,393 (1,391,724)	-34,305 (32,407)
$\beta_1 + \beta_2$	12,652	-82,858	51,867	32,280	6,857	113,900
P-value: $\beta_1 + \beta_2 = 0$	.12	.67	.69	.89	1	.042
Control mean if risk loving = 0	74,243	1,176,056	452,420	1,064,878	14,388,904	388,771
Control mean if risk loving = 1	81,032	1,304,103	633,123	1,258,952	18,360,600	355,319
R ²	.13	.18	.12	.18	.33	.18
Observations	838	838	838	838	838	838

All quantities in UGX; Outliers top/bottom coded to 95th/5th percentile; Heteroskedasticity-robust standard errors in parantheses

*Total indivisible investment includes real land values, adjusted down to exclude appreciation in land values over time

Controls include: pre-treatment levels of outcome, hh income, patience, gender, hh head, age, age², num ad females, num ad males, num children, district FE's

Table 2.13: Effect of winning the grant on likelihood of purchasing, improving, and selling land

	(1) purchased land between midline and endline3 = 1	(2) made improvements to land between midline and endline3 = 1	(3) sold land between midline and endline3 = 1
won lottery (0/1)	.011 (.038)	.032 (.039)	-.021 (.031)
won large lottery (0/1)	.15** (.078)	-.026 (.081)	-.04 (.065)
risk loving (0/1)	-.033 (.047)	.045 (.048)	.039 (.04)
$\beta_1 + \beta_2$	.16	.0057	-.061
P-value: $\beta_1 + \beta_2 = 0$	.017	.94	.29
Control mean if risk loving = 0	.37	.36	.18
Control mean if risk loving = 1	.33	.41	.22
R ²	.08	.043	.031
Observations	838	838	838

Heteroskedasticity-robust standard errors in parantheses; Controls include: pre-treatment land value, hh income, patience, gender, hh head, age, age², num ad females, num ad males, num children, district FE's

Table 2.14: Calibration Parameters and Moments

A. Assigned Parameters			
r	interest rate		0.00
θ	borrowing limit		0.00
δ	depreciation rate		0.01
p	survival probability		1 - 1/480
B. Calibrated Parameters			
β	discount factor		0.84
$\underline{k}$	size of capital unit		8.02
α	capital elasticity in production		0.35
μ_z	mean productivity		0.96
ν_z	persistent productivity dispersion		0.35
σ_z	productivity shock dispersion		1.13
ν_e	income dispersion		0.43
σ_e	income productivity shock dispersion		0.53
C. Moments			
		<u>Model</u>	<u>Data</u>
Income Distributions			
P10		0.39	0.15
P25		0.60	0.41
P50		1.00	1.00
P75		1.88	2.04
P90		4.18	3.72
Mean		1.98	1.73
Savings Distributions			
P10		0.00	-1.85
P25		0.00	0.00
P50		0.13	0.19
P75		0.47	1.16
P90		1.43	2.92
Mean		0.80	0.56
Lottery Choice			
Choose Risky Now		0.20	0.20
Choose Safe Now		0.56	0.59
Choose Risky Delay		0.04	0.07
Choose Safe Delay		0.19	0.14
Income Tercile Transition Probabilities			
		<u>Current</u> <i>Model</i>	
<u>Future</u> [†]		Low	High
Low	0.53	0.36	0.11
Medium	0.35	0.43	0.22
Hih	0.12	0.22	0.67
		<i>Data</i>	
		Low	High
Low	0.63	0.24	0.13
Medium	0.28	0.48	0.25
High	0.09	0.28	0.63

[†] Future measurements are taken after 16 months

Table 2.15: Counterfactual Aggregate Impacts of Financial Interventions

	Benchmark ($\theta = 0, r = 0$)	Divisible Capital		Credit ($\theta = 0.25, r = 0$)		Savings ($\theta = 0, r = 0.03$)		Combined ($\theta = 0.25, r = 0.03$)	
		PE	GE	PE	GE	PE	GE	PE	GE
Agg. Consumption	1.00	2.15	1.64	1.53	1.00	1.11	1.02	1.74	1.03
Agg. Income	1.00	2.15	1.64	1.53	1.00	1.11	1.02	1.72	1.02
Agg. Net Savings	1.00	1.57	0.19	0.91	0.38	1.37	1.20	1.44	0.58
Agg. Capital	1.00	2.98	<i>1.00</i>	2.30	<i>1.00</i>	1.20	<i>1.00</i>	2.76	<i>1.00</i>
Value of Capital	1.00	2.98	2.97	2.30	1.33	1.20	1.06	2.76	1.39
Price of Capital	1.00	<i>1.00</i>	2.95	<i>1.00</i>	1.33	<i>1.00</i>	1.06	<i>1.00</i>	1.39
Capital Income	1.00	4.26	3.09	2.32	1.00	1.23	1.02	2.75	1.03
Fraction Poor	1.00	0	0	0.68	1.00	0.82	0.86	0.50	0.86
Probability of Staying Poor	1.00	0	0	1.00	1.00	0.86	0.86	0.88	0.86

Notes: Values are reported relative to the benchmark, which has been normalized to one. Initial benchmark values are aggregate consumption = 2.10, aggregate income = 2.11, aggregate net savings = 0.86, aggregate capital = 1.59, price of capital = 1.00, capital income = 0.85, fraction poor = 0.33, and probability of staying poor = 0.42. Those people whose wealth is below the initial 33.33% of the wealth distribution are considered poor, and the probability of staying poor is calculated if they are poor 16 months later, conditional on being poor initially. Divisible capital allows households to purchase any value of $k \in [0, k_1]$. PE fixes the price of capital to 1. GE fixes the aggregate quantity of capital to that in the benchmark. In both cases, these values that are constant by construction are indicated by italics.

Appendix

Appendix 2.A Data Appendix

The baseline and all three endline surveys contained detailed questions on household consumption, savings, income, assets, borrowing and lending (mapping to the components of the household budget constraint, which we use to structure our empirical analysis). The midline survey, intended to be somewhat briefer as it took place immediately preceding the lottery, consisted of questions on consumption, savings, income, business assets (but not home durables, land, livestock, or other agricultural assets), and borrowing. We detail the measurement of key outcomes below.

In the first and second endline, we prime respondents with the consumption, savings, income, or asset level that they reported at baseline or midline, and we ask whether that level has increased, decreased, or stayed the same, to prevent drastically different interpretations of the same question across survey waves. If the level has changed, we then ask for the new level. In the third endline, we do not prime respondents with their earlier responses as six years have passed and earlier responses may not be a useful benchmark.

2.A.1 Consumption

Consumption is constructed from detailed questions on subcategories of regular spending over the period of a week plus questions on the frequency and amount of less regular expenses. Specifically, we ask about regular weekly spending on:

- Staple grains, beans, other (non-meat, non-fish) food that is prepared at home, and cooking supplies
- Meat
- Fish
- Milk
- Non-milk beverages, including tea, beer, liquor, coffee, soda, and juice
- Transportation, including fuel for transportation

- Airtime
- Electricity, gas, firewood, and charcoal

We ask whether the respondent has incurred any of these less regular expenses in the past year and, if so, then we follow-up with questions about how much:

- Visits to hospitals, doctors, or other healers, and medicine
- School fees
- Expenses associated with marriage and marital ceremonies

We combine spending on all of these categories and standardize the frequency of incurred expenses to generate a measure of weekly consumption.

2.A.2 Home Durables

To collect the value of home durables, which we group with consumption in our analyses, we first ask about ownership of home assets (lamps, radios, stoves, generators, cell phones, televisions, refrigerators, carpets, sofas, tables, bicycles, motorbikes, and any other household asset that we might have missed). If a respondent reports to own a particular asset, then we ask how many they own, and we ask for the average value of one unit. We then sum values across all home assets that a respondent owns to obtain our measure of home durables.

2.A.3 Savings

For savings, we first ask the respondent for an estimate of their total savings. We then ask where they hold their savings (formal bank, microfinance institution, a savings cooperative known as a SACCO, any other savings group, with another person, in a secret place, or in a mobile money account). For each place where they tell us they hold savings, we ask how much they hold in that place. We then sum these values over all of the respondent's recorded savings places and ask the respondent whether they believe that the aggregate sum they provided initially or this sum of components better represents their total savings. Their preferred measure becomes our measure of savings.

2.A.4 Income

We collect "typical monthly income" through detailed questions on subcategories of income: crop income, livestock income, non-farm business income, wage/salary income, and remittances. Like savings, we start

by asking the respondent for their best guess of their typical monthly income, and we then follow-up with detailed questions on each income component and produce our own calculation of total monthly income. Finally, we ask the respondent which measure they believe is more accurate, their initial aggregate estimate or our calculation from components.

To collect crop income, we ask the respondent which crops they harvest and how frequently. For each crop, we ask for the typical quantity that they produce with each harvest, how much they consume, and how much they sell. For the sold quantity, we ask the average price per unit sold, and we then calculate revenue per crop. We separately ask for the typical costs incurred to harvest all crops over the course of a year (including labor, fertilizer, and pesticides). We then construct crop income as revenue across all crops less costs across all crops.

To collect livestock income, we ask the respondent which animals the household has owned in the past 12 months. For each one that the household has owned, we ask how many they have sold in the past year and their earnings from these sales. We then ask which types of expenses they incur to maintain livestock (animal feed, labor, veterinary services, or other expenses) and the cost of each. We produce profit per animal and sum across all animals to construct total income from livestock.

To collect (non-farm) business income, we ask the respondent whether they own any businesses and how many. For each business, we ask about the number of months per year that the business operates, the typical sales per month, what types of expenses are incurred (inventory, labor, and any other costs), and the cost of each per month. From these questions, we construct monthly profit for each business. We ask the respondent whether our monthly calculation seems accurate, and if not, then we give them the opportunity to provide a corrected measure of monthly business profit.

We ask the respondent if they or any other household members earn income from wage or salaried jobs. If so, then we collect the typical monthly amount earned by the respondent and, separately, by other household members. Finally, we ask the respondent if they receive remittances from family within Uganda or abroad and if so, then we follow-up with questions about the typical monthly value of remittances.

2.A.5 Agricultural Assets

Agricultural assets include livestock as well as durables (pangas, axes, hammers, spades, sickles, and ploughs). For each animal or durable that a household owns, we ask how many they own and the average value of one unit. We collect livestock at an aggregate level (current total value of livestock) and through disaggregated categories for each animal, with a follow-up question about which measure of total livestock value is better.

2.A.6 Business Assets

To construct business assets, we separately ask about the current level of business inventory and other non-inventory business assets (machines or equipment, non-home buildings or land that are primarily for business use, and other capital assets) used in each non-farm business. We sum across all enterprises owned by the household to construct total business assets.

2.A.7 Land

We ask respondents to report the value of their land, including any dwellings on the land. As with prior categories, we prime the respondent with their previously reported land value and ask whether their land has since increased, decreased, or stayed the same in value. If it has changed, we ask the new value. At the third endline survey only, we add questions to separately capture land purchases, land sales, and land investment, including retrospectively over the entire experimental period.

As discussed in Section 2.3.2, we find substantial appreciation in land values over time (i.e., across survey waves). To derive the total capital gain in land value over time, we sum land values across all control households in each district d at endline over the sum of the same set of control households' land values at baseline. We allow capital gains (ϕ) to vary by district. We do this for both the first and second endline and adjust the ratios to reflect appreciation solely between midline and each respective endline. Specifically, for control households in a given d district:

$$\phi_d = \left(\frac{\sum_i land_i^e}{\sum_i land_i^b} \right)^{\frac{x}{y}} \quad (2.10)$$

where $d \in \{Ntungamo, Ibanda, Kagadi\}$, i denotes household, x denotes time in months between midline and the relevant endline, y denotes time in months between baseline and the relevant endline, b specifies baseline, and e specifies the relevant endline (either the first or second). The calculated capital gains rates are listed in Table 2.A.1. We find that land values appreciated at a rate of approximately 2% per month.

Table 2.A.1: Estimated capital gains by region

	midline to endline 1 (4 months)	midline to endline 2 (18 months)	average monthly (baseline to endline 2)
Ntungamo	1.084	1.36	1.017
Ibanda	1.079	1.47	1.022
Kagadi	1.078	1.42	1.020
Overall	1.081	1.41	1.019

We then apply these capital gains rates to adjust land values downward, i.e., net of appreciation, such

that we can measure the effect of the grants on real land values and such that the cross-equation budget constraint used in the SUR will hold. We construct new capital gains-adjusted land^{*} values for all households in each district d at each endline:

$$land_{id}^* = \left(\frac{1}{\phi_d} \right) land_{id} \quad (2.11)$$

2.A.8 Net Credit

For credit, we ask the respondent if they have any loans outstanding and, if so, then the current amount owed. We also ask the same set of questions for loans that the respondent has made to others. We construct net credit as the current outstanding amount that the household owes to others less the current amount others owe to the household. Thus, a positive value of net credit reflects that the respondent owes money on net and a negative value reflects that others owe money to the respondent on net.

Appendix 2.B Evidence of GE Effects for Land

The model in Section 2.2 indicated that risk-loving behavior can be linked to large, indivisible investments, and Section 2.6 showed that if the investment good is not elastically supplied, the aggregate impacts of financial services can be limited through general equilibrium forces. Given that the empirical results in Section 2.5 demonstrated that winners of a large lottery have a large propensity to invest winnings in land, and that land price appreciation (2 percent per month) is sizable (Section 2.3.2), a natural question is whether general equilibrium forces are important for land. That is, does demand for land investment increase the prices of land?

Given the high propensity to purchase land out of large lottery winnings, our randomized experiment generated exogenous variation in the demand for land. We therefore test the impact of the local grant winnings on land values by estimating the impact of more grants being awarded within close proximity to a participant household, using 0.5 and 1 mile radii around the household as measures of proximity. That is, we run the following two-stage least squares model:

$$\Delta LocalLandValue_{id} = \beta_0 + \beta_1 GrantsWithinRadius_i + \gamma X_i + \lambda_d + \varepsilon_{id} \quad (2.12)$$

$$HouseholdOwnLandValue_{id} = \beta_0 + \beta_1 \widehat{\Delta LocalLandValue}_i + \gamma X_i + \lambda_d + \varepsilon_{id} \quad (2.13)$$

Where X_i controls for the (sample) number of households within the radius of interest (0.5 miles or 1 mile), the number of households choosing the large lottery within the radius, whether the household won a grant itself (*won lottery*), whether the household itself chose the large lottery (*risk loving*), the household's own land value at baseline, and the same set of demographic controls included in our main estimating equation (Equation 2.6). We also include λ_d , district fixed effects, as in our main estimating equation. In this specification, the independent variation in winning a grant is the result of the realization of random draws within the area. The number of grants disbursed within a given radius vary at the household-level, and include only those grants which were given to *surrounding* households within the given radius. We cluster standard errors by the 141 geographic “neighborhoods” used in our census survey (with an average of 7 households per neighborhood).

Table 2.B.1 presents the impact of local grants on local land values. At the first endline, we find that each additional grant within 0.5 miles increases local land values by approximately 6%. The effect is even larger in a 1 mile radius — each additional grant increases values by 7%.³⁹ While the first stage is a bit weak

³⁹The effects are large, but if one unit of land is sold above the status quo price, the values of *all* land may increase correspondingly. Indeed, this indirect impact is precisely our interest. Note also that the impact may also include any increase in local growth from people investing in their businesses.

in the 0.5 mile radius specification with an F -statistic of 3.5, the F -statistic for the one mile specification is 17 (reported in Table 2.B.2). Point estimates are similar at the second endline, and again the F -statistic is stronger in the one mile specification.

Table 2.B.1: Effect of grants disbursed nearby on others' land values nearby (first stage)

	(1) Ln sum of others' land values within 0.5 mi _{e1}	(2) Ln sum of others' land values within 0.5 mi _{e2}	(3) Ln sum of others' land values within 1 mi _{e1}	(4) Ln sum of others' land values within 1 mi _{e2}
num grants within 0.5 mi	.063* (.034)	.066** (.031)		
num grants within 1 mi			.072*** (.018)	.063*** (.017)
num risk lovers within 0.5 mi	.00076 (.04)	-.023 (.034)		
num houses within 0.5 mi	.1*** (.024)	.097*** (.022)		
num risk lovers within 1 mi			.0077 (.022)	.0071 (.02)
num houses within 1 mi			.041*** (.012)	.044*** (.011)
won lottery (0/1)	-.02 (.065)	-.02 (.057)	.017 (.048)	-.019 (.049)
risk loving (0/1)	-.18** (.078)	-.16** (.07)	-.045 (.057)	-.064 (.058)
district fe's	Yes	Yes	Yes	Yes
demographic controls	Yes	Yes	Yes	Yes
R ²	.45	.47	.55	.55
Control mean (level)	108,967,724	146,208,633	198,887,799	267,238,782
Observations	740	774	764	801

All quantities in UGX; Outliers top/bottom coded to 95th/5th percentile; Standard errors clustered at the neighborhood
Sample size changes due to restriction that land values must be positive, the likelihood of which does not vary by treatment
Controls include: pre-intervention levels of own land value, hh income, patience, gender, hh head, age, age², num ad females, num ad males, num children

In the second stage (Table 2.B.2), we find that a 1% increase in neighboring land values leads to a 0.42% increase in the household's own land values at the first endline and 0.47% increase at the second endline (using the 1 mi radius specification). Given an average of 6.67 grants disbursed within a mile of each house, a household's own land value is 19.7%(=6.67 grants*6.3% increase in local land values per grant*0.47% increase in own land value per 1% increase in local land values) higher at the second endline due to grants disbursed by the experiment, as reported in Section 2.3.2.

Table 2.B.2: Effect of others' land values nearby on own land value (second stage)

	(1) Ln land _{e1}	(2) Ln land _{e2}	(3) Ln land _{e1}	(4) Ln land _{e2}
Ln sum of others' land values within 0.5 mi _{e1}	.24 (.42)			
Ln sum of others' land values within 0.5 mi _{e2}		.28 (.39)		
Ln sum of others' land values within 1 mi _{e1}			.42 (.26)	
Ln sum of others' land values within 1 mi _{e2}				.47* (.28)
num risk lovers within 0.5 mi	-.019 (.029)	-.046 (.034)		
num houses within 0.5 mi	-.035 (.06)	-.032 (.055)		
num risk lovers within 1 mi			-.019 (.02)	-.035* (.019)
num houses within 1 mi			-.027 (.022)	-.025 (.023)
won lottery (0/1)	.034 (.059)	.14** (.058)	.0045 (.059)	.12* (.061)
risk loving (0/1)	.057 (.11)	.18* (.094)	.043 (.073)	.16** (.074)
district fe's	Yes	Yes	Yes	Yes
demographic controls	Yes	Yes	Yes	Yes
First stage F-stat	3.5	4.5	17	14
Control mean (level)	18,174,911	17,141,538	18,656,827	22,678,972
Observations	740	774	764	801

All quantities in UGX; Outliers top/bottom coded to 95th/5th percentile; Standard errors clustered at the neighborhood; Sample size changes due to restriction that land values must be positive, the likelihood of which does not vary by treatment; Controls include: pre-intervention levels of own land value, hh income, patience, gender, hh head, age, age², num ad females, num ad males, num children

Hence, we have direct evidence of price increases resulting from the demand for land, indicating that land is not perfectly elastically supplied even in the underbanked, peri- and semi-urban small cities that we study. In the context of our model, this limits the impacts that financial services can have in promoting development and the escape from poverty. Moreover, the emphasis on land as an investment good disproportionately favors savings services relative to credit services toward these ends, since credit had no impact when the investment capital was in fixed supply.

Assuming increases in demand are purely partial equilibrium and so induce true increases in land (i.e.,

land is not in absolutely fixed supply), we use the estimates in Table 2.7 to estimate the increase in quantity demanded. We can use this to yield the back-of-the-envelope calculation of an elasticity of the supply of land of 0.05 in the paper. Specifically, we calculate a percentage increase in demand, where the total increase in demand for land is the product of the number of participants (1048), their lottery-choice probabilities (0.27 for the large lottery and 0.73 for the small lottery), the respective probabilities of winning (0.3 and 0.49, respectively), and the increase in land demand as a result of winning, per Column 8 of Table 2.7 (7.5 million and 300 thousand UGX, respectively). This yields an increase in demand for land of roughly 750 million UGX. Total land demand is the number of households in our original census (3734) times average land holdings in the control of our sample (roughly 21 million UGX), which yields a baseline demand of land of 78 billion. The percentage increase in land demand is therefore close to 1 percent. Dividing this 1% increase in land demand by the 20% yields the reported elasticity of 0.05.

Appendix 2.C Predicting Lottery Choice

The following table presents the full set of covariates on which we test for differences between those selecting the large lottery and those selecting the small lottery:

Table 2.C.1: Characteristics of those choosing the large v. small lottery

	large	N	small	N	diff	p-value
<i>Income and consumption</i>						
mtlhy income _m	398,608	283	371,090	765	27,518	0.29
mtlhy income/adult equiv _m	122,872	283	126,665	765	-3,793	0.68
total income per hour worked	911	283	910	765	.94	0.99
ln mtlhy income _m	12	283	12	765	-.043	0.73
Δ ln mtlhy income _{m-b}	.71	283	1.2	765	-.44**	0.02
mtlhy crop income _m	100,208	283	71,148	765	29,061***	0.00
crop income/total income _m	.3	283	.25	765	.05***	0.01
mtlhy crop income/adult equiv _m	30,562	283	22,981	765	7,580***	0.00
crop income per hour worked	860	283	696	765	164	0.14
ln mtlhy crop income _m	12	283	12	765	.21***	0.00
Δ ln mtlhy crop income _{m-b}	-1.5	283	-1.4	765	-.051	0.50
mtlhy bus income _m	125,247	283	119,725	765	5,523	0.68
bus income/total income _m	.28	283	.29	765	-.0087	0.79
mtlhy bus income/adult equiv _m	41,641	283	41,745	765	-103	0.98
bus income per hour worked	916	283	1,092	765	-176	0.49
ln mtlhy bus income _m	12	283	12	765	.033	0.54
Δ ln mtlhy bus income _{m-b}	6.7	283	7.6	765	-.87**	0.02
wkly consumption _m	44,466	283	39,453	765	5,013**	0.01
wkly cons/adult equiv _m	14,034	283	13,798	765	236	0.77
ln wkly cons _m	10	283	10	765	.18**	0.01
Δ ln wkly cons _{m-b}	.38	283	.24	765	.14**	0.03
<i>Savings and wealth</i>						
savings _m	322,675	283	275,817	765	46,858*	0.10
savings/adult equiv _m	103,375	283	96,393	765	6,982	0.51
ln savings _m	10	283	9.8	765	.31	0.35
Δ ln savings _{m-b}	2.5	283	1.7	765	.86**	0.03
bus assets _m	824,954	283	577,814	765	247,140**	0.01
bus assets/wealth _m	.29	283	.24	765	.053**	0.05
bus assets/adult equiv _m	295,958	283	218,673	765	77,285*	0.07
ln bus assets _m	5.3	283	4.3	765	.99**	0.03
Δ ln bus assets _{m-b}	.6	283	.23	765	.37	0.22
wealth (sav + bus assets) _m	1,245,155	283	920,185	765	324,970***	0.01
wealth/adult equiv _m	431,561	283	337,778	765	93,784*	0.07
ln wealth _m	11	283	11	765	.65**	0.05
Δ ln wealth _{m-b}	2.2	283	1.5	765	.7**	0.05
net wealth (sav + bus assets - credit) _m	1,070,910	283	744,418	765	326,492***	0.01
net wealth/adult equiv _m	377,813	283	277,547	765	100,265**	0.05
ln net wealth _m	16	283	16	765	.075*	0.06
Δ ln net wealth _{m-b}	-.22	283	-.25	765	.029	0.57
land value _b	13,345,159	283	9,852,680	765	3,492,479***	0.00

land value/adult equiv _b	3,936,519	283	3,010,483	765	926,036***	0.00
ln land value _b	13	283	12	765	.91**	0.04
Other financial indicators						
operates non-farm business _m (0/1)	.59	283	.54	765	.049	0.15
farmer _m (0/1)	.71	283	.75	765	-.041	0.18
work hours per week _m	78	283	77	765	.52	0.74
had negative shock since baseline _m (0/1)	.63	283	.67	765	-.046	0.16
has formal savings _m (0/1)	.088	283	.12	765	-.035	0.12
acquired loans since baseline _m (0/1)	.29	283	.32	765	-.035	0.27
credit outstanding _m	193,779	283	192,949	765	830	0.98
Desire to invest						
wants credit to increase income _b (0/1)	.84	283	.78	765	.062**	0.03
would invest >\$100 _b (0/1)	.95	283	.91	765	.038**	0.04
would use credit for bus investment _b (0/1)	.67	283	.59	765	.073**	0.03
would use credit for ag investment _b (0/1)	.053	283	.08	765	-.027	0.14
Hypothetical preferences						
would invest for 53% exp gain _b (0/1)	.67	283	.64	765	.032	0.33
would invest for 105% exp gain _b (0/1)	.7	283	.67	765	.031	0.34
would invest for 1% mthly interest _b (0/1)	.24	283	.23	765	.0054	0.86
desired mthly interest to invest now _b	16	283	16	765	.22	0.88
Demographic characteristics						
female (0/1)	.42	283	.51	765	-.09***	0.01
household head (0/1)	.66	283	.6	765	.058*	0.09
respondent age	37	283	35	765	2.3***	0.00
education beyond primary school (0/1)	.26	283	.29	765	-.026	0.40
num people in household _b	5.5	283	5	765	.47***	0.00
num adult females _b	1.1	283	1.1	765	-.0095	0.82
num adult males _b	1.5	283	1.4	765	.14*	0.08
num children _b	2.8	283	2.5	765	.34***	0.00
Observations	1048					

All quantities in UGX; Outliers top/bottom coded to 95th/5th percentile; Midline denoted by **m** and baseline denoted by **b**
bus is an abbreviation for business; * $p < 0.1$, ** $p < .05$, *** $p < 0.01$

In table 2.C.2, we analyze the predictors of lottery choice using lasso. Among 160 baseline and midline covariates, the following are selected:

Table 2.C.2: Predicting those who chose the large lottery: Lasso-selected covariates

	Penalized coefficient:
total income per hour worked	-.0000192
mthly crop income _m	3.59e-07
bus income per hour worked	-3.15e-06
monthly livestock income _m	2.16e-08
monthly livestock income _b	2.39e-06
mthly wage income _m	-5.37e-08
wealth _b	1.46e-09
net savings _b	-6.09e-08

bus assets _b	4.06e-08
wkly consumption _b	1.18e-06
farm assets _b	2.63e-07
ag assets _b	4.37e-08
ln mthly income _m	-.0189
ln mthly crop income _m	.0255
ln savings _b	-.005
Δ ln mthly income _{m-b}	-.00199
Δ ln mthly bus income _{m-b}	-.00236
Δ ln wkly cons _{m-b}	.0387
Δ ln wealth _{m-b}	.00288
mthly crop income/adult equiv _b	2.35e-08
farmer _m (0/1)	-.0522
experienced bad event _b (0/1)	-.0534
has formal savings _m (0/1)	-.11
acquired loans since baseline _m (0/1)	-.0311
wants credit to increase income _b (0/1)	.0775
would invest >\$100 _b (0/1)	.0787
would use credit for bus investment _b (0/1)	.0169
would use credit for ag investment _b (0/1)	-.0654
female (0/1)	-.0339
respondent age ²	.0000217
num children _b	.0146
crop: Irish Potato (0/1)	-.0545
crop: Sweet Potato (0/1)	-.0644
crop: Yam (0/1)	-.451
new crops since baseline _m (0/1)	-.0295
savings place: SACCO (0/1)	-.058
savings place: ROSCA or other cooperative/ community group (0/1)	.104
savings place: In a secret place (0/1)	.0434
bad event: Loss of crop due to disease, etc (0/1)	-.061
bad event: Assets damaged or destroyed (0/1)	.253
bad event: Sickness or injury to family member (0/1)	-.0456
opened a new business since baseline _m (0/1)	-.0294
Observations	1048

The table depicts the unstandardized penalized coefficients of those covariates which were selected from among 160 baseline and midline variables given to lasso. We set the penalty parameter using adaptive lasso. All quantities in UGX. Outliers top/bottom coded to 95th/5th percentile. Midline denoted by **m** and baseline denoted by **b**; **bus** is an abbreviation for business. Full list of covariates which we give to lasso is available on request.

Looking only at demographic characteristics and those financial outcomes which are collected at both baseline and midline, we can compare the variables selected by lasso when predicting midline lottery choice relative to baseline (hypothetical) risk preferences:

Table 2.C.3: Lasso-selected predictors of midline lottery choice and baseline hypotheticals

	Midline lottery	Baseline: Greater risk	Baseline: Moderate risk
--	--------------------	---------------------------	----------------------------

	choice	preference	preference
mtlhy crop income	3.76e-07		
mtlhy livestock income	6.82e-07		
mtlhy wage income	-1.06e-07		
net wealth (sav + bus assets - credit)	6.61e-09		
ln mtlhy income	-.0245	.00852	.0102
ln mtlhy crop income	.0227	.000845	
ln wkly consumption _m	.019		
ln bus assets _b	.00311		
recent negative shock (0/1)	-.0348		
respondent age ²	.00003		-4.03e-06
num children _b	.0162		
gender	-.0428	-.0561	-.0779
net savings		3.61e-08	
savings		1.09e-07	1.35e-07
ln savings _m		.000706	
mtlhy income/adult equiv		1.23e-09	
wkly consumption/adult equiv		2.25e-06	5.54e-07
num adult females _b		-.00643	
Observations	1048	1048	1048

The table depicts the unstandardized penalized coefficients of those covariates which were selected by lasso from among 39 demographic characteristics and financial outcomes collected at both midline and baseline. We set the penalty parameter using adaptive lasso. All quantities in UGX. Outliers top/bottom coded to 95th/5th percentile. Midline denoted by **m** and baseline denoted by **b**; **bus** is an abbreviation for business. Full list of covariates which we give to lasso is available on request.

Appendix 2.D Attrition

We first compare balance on observable characteristics between the retained and attrited sample. We then compute Lee Bounds.

2.D.1 Balance Between Retained and Attrited Sample

Tables 2.D.1 and 2.D.2 show that observable characteristics are, in general, balanced between the retained versus attrited sample, suggesting that attrition was idiosyncratic.

Table 2.D.1: Balance between retained and attrited, among those choosing the small lottery

	chose small retained	N	chose small attrited	N	diff	p-value
<i>Household budget components</i>						
mtlly income _m	373,736	634	358,289	131	-15,446	0.66
wkly consumption _m	39,894	634	37,321	131	-2,573	0.33
savings _m	279,172	634	259,580	131	-19,592	0.61
credit outstanding _m	194,256	634	186,626	131	-7,630	0.84
home durable value _b	504,709	634	583,386	131	78,678	0.32
<i>Investment categories</i>						
total divisible investments _m	376,729	634	388,587	131	11,858	0.84
small livestock and ag. assets _b	168,910	634	156,534	131	-12,377	0.58
bus inventory _m	160,700	634	173,702	131	13,003	0.73
total indivisible investments _m	11,175,818	634	10,075,985	131	-1,099,833	0.44
large livestock and ag. assets _b	262,088	634	241,527	131	-20,562	0.80
land value _b	10,003,707	634	9,121,756	131	-881,951	0.50
bus assets, no stock _m	347,215	634	389,267	131	42,053	0.63
<i>Other financial indicators</i>						
operates non-farm business _m (0/1)	.54	634	.53	131	-.013	0.79
farmer _m (0/1)	.76	634	.73	131	-.032	0.44
work hours per week _m	77	634	77	131	-.063	0.98
had negative shock since baseline _m (0/1)	.66	634	.73	131	.064	0.15
has formal savings _m (0/1)	.13	634	.099	131	-.029	0.37
acquired loans since baseline _m (0/1)	.33	634	.3	131	-.029	0.52
<i>Demographic characteristics</i>						
female (0/1)	.5	634	.55	131	.051	0.29
household head (0/1)	.6	634	.6	131	-.0087	0.85
respondent age	36	634	32	131	-4.5***	0.00
education beyond primary school (0/1)	.28	634	.33	131	.049	0.26
num people in household _b	5	634	4.9	131	-.11	0.63
num adult females _b	1.1	634	1.2	131	.029	0.59
num adult males _b	1.4	634	1.3	131	-.13	0.20
num children _b	2.5	634	2.5	131	-.00072	1.00
Observations	765					

All quantities in UGX; Outliers top/bottom coded to 95th/5th percentile; Midline denoted by **m** and baseline denoted by **b**
bus is an abbreviation for business; * $p < 0.1$, ** $p < .05$, *** $p < 0.01$

Table 2.D.2: Balance between retained and attrited, among those choosing the large lottery

	chose large retained	N	chose large attrited	N	diff	p-value
<i>Household budget components</i>						
mthly income _m	392,511	233	427,022	50	34,511	0.56
wkly consumption _m	44,677	233	43,485	50	-1,192	0.80
savings _m	304,584	233	406,980	50	102,396	0.11
credit outstanding _m	193,386	233	195,612	50	2,226	0.97
home durable value _b	722,479	233	564,394	50	-158,085	0.28
<i>Investment categories</i>						
total divisible investments _m	516,826	233	486,620	50	-30,206	0.79
small livestock and ag. assets _b	197,858	233	161,140	50	-36,718	0.36
bus inventory _m	233,197	233	234,000	50	803	0.99
total indivisible investments _m	16,022,090	233	13,242,000	50	-2,780,090	0.30
large livestock and ag. assets _b	614,914	233	537,000	50	-77,914	0.71
land value _b	13,724,807	233	11,576,000	50	-2,148,807	0.38
bus assets, no stock _m	511,661	233	502,800	50	-8,861	0.96
<i>Other financial indicators</i>						
operates non-farm business _m (0/1)	.61	233	.48	50	-.13*	0.09
farmer _m (0/1)	.72	233	.68	50	-.037	0.60
work hours per week _m	77	233	80	50	2.9	0.44
had negative shock since baseline _m (0/1)	.63	233	.62	50	-.0066	0.93
has formal savings _m (0/1)	.082	233	.12	50	.038	0.39
acquired loans since baseline _m (0/1)	.3	233	.24	50	-.056	0.43
<i>Demographic characteristics</i>						
female (0/1)	.42	233	.42	50	.0037	0.96
household head (0/1)	.66	233	.66	50	-.00094	0.99
respondent age	38	233	36	50	-1.4	0.45
education beyond primary school (0/1)	.26	233	.28	50	.022	0.74
num people in household _b	5.4	233	5.6	50	.23	0.56
num adult females _b	1.1	233	1.2	50	.08	0.42
num adult males _b	1.5	233	1.6	50	.12	0.53
num children _b	2.8	233	2.8	50	.03	0.92
Observations	283					

All quantities in UGX; Outliers top/bottom coded to 95th/5th percentile; Midline denoted by **m** and baseline denoted by **b**
bus is an abbreviation for business; * $p < 0.1$, ** $p < .05$, *** $p < 0.01$

2.D.2 Lee Bounds

We compute Lee Bounds (Lee, 2009) around the estimates in Tables 2.8, 2.10, and 2.12 (which also include the aggregate investment categories from Tables 2.6 and 2.7).

Table 2.D.3: Grant effects on components of the household budget constraint - First endline - Lee Bounds

	(1) consum. (stock) wkly cons + home dur.	(2) savings	(3) div. investment	(4) indiv. investment*	(5) mthly income	(6) net credit
Lower bound:						
won lottery (0/1)	-148,151* (79,287)	-54,110** (24,149)	3,195 (52,218)	-2,375,215*** (827,145)	-62,680*** (18,032)	-70,405*** (14,329)
won large lottery (0/1)	47,319 (154,848)	141,160** (61,070)	187,413 (138,772)	5,111,855*** (1,811,346)	33,697 (35,939)	-19,604 (32,414)
risk loving (0/1)	115,939 (108,737)	6,530 (35,059)	90,432 (76,377)	-1,886,169* (963,842)	-3,501 (26,011)	2,400 (22,577)
$\beta_1 + \beta_2$	-100,832	87,050	190,607	2,736,640	-28,983	-90,009
P-value: $\beta_1 + \beta_2 = 0$	0.45	0.12	0.14	0.08	0.35	0.00
Observations	830	830	830	830	830	830
Upper bound:						
won lottery (0/1)	209,641** (93,427)	79,795** (31,801)	253,603*** (63,307)	-483,214 (942,183)	6,398 (20,738)	24,335 (18,058)
won large lottery (0/1)	-31,805 (183,444)	109,859 (72,806)	117,423 (158,359)	5,270,092** (2,072,590)	46,522 (45,486)	-28,324 (39,899)
risk loving (0/1)	112,830 (109,195)	12,760 (35,447)	79,779 (75,989)	-2,284,135** (967,924)	-4,461 (25,969)	-1,839 (22,542)
$\beta_1 + \beta_2$	177,836	189,654	371,026	4,786,878	52,920	-3,989
P-value: $\beta_1 + \beta_2 = 0$	0.26	0.00	0.01	0.01	0.19	0.91
Observations	830	830	830	830	830	830
Control mean if risk loving = 0	1,608,566	259,468	613,887	15,232,971	327,076	73,180
Control mean if risk loving = 1	1,798,491	271,428	787,251	17,160,718	321,271	80,220

This table constructs Lee Bounds around the point estimates reported in Table 2.8. All quantities in UGX; Outliers top/bottom coded to 95th/5th percentile; Heteroskedasticity-robust standard errors in parantheses. *Total indivisible investment includes real land values, adjusted down to exclude appreciation in land values over time. All specifications include the following controls: pre-treatment levels of outcome, hh income, patience, gender, hh head, age, age², num ad females, num ad males, num children, district FE's. * $p < 0.1$, ** $p < .05$, *** $p < 0.01$

Table 2.D.4: Grant effects on components of the household budget constraint - Second endline - Lee Bounds

	(1) consum. (stock) wkly cons + home dur.	(2) savings	(3) div. investment	(4) indiv. investment*	(5) mthly income	(6) net credit
Lower bound:						
won lottery (0/1)	-407,772* (223,057)	-67,797*** (25,823)	-117,762** (50,527)	-2,562,868*** (890,014)	-45,320*** (16,049)	-102,605*** (19,489)
won large lottery (0/1)	602,600 (428,585)	117,276* (61,291)	57,066 (133,424)	5,285,826** (2,190,759)	32,263 (38,779)	22,405 (42,613)
risk loving (0/1)	192,877 (294,575)	2,274 (38,317)	98,264 (76,822)	-531,663 (1,218,042)	24,512 (23,907)	-43,127 (30,008)
$\beta_1 + \beta_2$	194,827	49,479	-60,696	2,722,958	-13,057	-80,200
P-value: $\beta_1 + \beta_2 = 0$	0.59	0.38	0.62	0.17	0.71	0.03
Observations	830	830	830	830	830	830
Upper bound:						
won lottery (0/1)	572,033** (257,861)	76,330** (34,248)	113,912* (62,345)	1,041,810 (1,097,676)	30,941 (19,623)	13,321 (24,035)
won large lottery (0/1)	356,270 (489,459)	72,817 (73,488)	-12,360 (144,195)	4,969,695* (2,642,252)	15,901 (43,465)	12,530 (52,184)
risk loving (0/1)	161,890 (295,049)	2,165 (38,354)	88,106 (76,570)	-910,770 (1,221,176)	27,702 (24,130)	-45,812 (30,347)
$\beta_1 + \beta_2$	928,303	149,147	101,551	6,011,505	46,842	25,852
P-value: $\beta_1 + \beta_2 = 0$	0.03	0.02	0.43	0.01	0.22	0.58
Observations	830	830	830	830	830	830
Control mean if risk loving = 0	4,603,540	273,166	596,244	15,040,343	264,233	119,272
Control mean if risk loving = 1	5,036,368	279,987	776,076	17,326,575	289,215	76,352

This table constructs Lee Bounds around the point estimates reported in Table 2.10. All quantities in UGX; Outliers top/bottom coded to 95th/5th percentile; Heteroskedasticity-robust standard errors in parantheses. *Total indivisible investment includes real land values, adjusted down to exclude appreciation in land values over time. All specifications include the following controls: pre-treatment levels of outcome, hh income, patience, gender, hh head, age, age², num ad females, num ad males, num children, district FE's. * $p < 0.1$, ** $p < .05$, *** $p < 0.01$

Table 2.D.5: Grant effects on components of the household budget constraint - Third endline - Lee Bounds

	(1) weekly consumption	(2) home durables	(3) savings	(4) div. investment	(5) indiv. investment*	(6) mthly income
Lower bound:						
won lottery (0/1)	-12,834*** (3,851)	-337,951*** (103,943)	-165,523*** (43,192)	-296,537*** (100,895)	-2,954,864*** (924,806)	-100,943*** (25,422)
won large lottery (0/1)	16,089** (8,164)	-52,630 (204,292)	24,113 (116,581)	-23,768 (228,185)	156,257 (1,979,303)	129,368** (56,376)
risk loving (0/1)	5,015 (5,493)	61,534 (151,073)	182,096** (75,649)	170,976 (157,704)	3,100,868** (1,478,582)	-38,952 (33,238)
$\beta_1 + \beta_2$	3,255	-390,581	-141,410	-320,305	-2,798,607	28,425
P-value: $\beta_1 + \beta_2 = 0$	0.65	0.03	0.20	0.12	0.11	0.57
Observations	807	807	807	807	807	807
Upper bound:						
won lottery (0/1)	4,093 (4,467)	66,252 (122,693)	56,236 (58,217)	130,791 (126,066)	1,270,593 (1,186,374)	-2,419 (29,730)
won large lottery (0/1)	16,250* (9,022)	-91,180 (242,751)	44,394 (144,561)	88,252 (287,047)	771,684 (2,556,211)	150,179** (66,362)
risk loving (0/1)	4,294 (5,539)	88,558 (151,965)	172,680** (76,191)	186,672 (157,743)	2,974,328** (1,488,632)	-32,484 (33,246)
$\beta_1 + \beta_2$	20,343	-24,929	100,630	219,042	2,042,277	147,760
P-value: $\beta_1 + \beta_2 = 0$	0.01	0.91	0.45	0.40	0.37	0.01
Observations	807	807	807	807	807	807
Control mean if risk loving = 0	74,243	1,176,056	452,420	1,064,878	14,388,904	388,771
Control mean if risk loving = 1	81,032	1,304,103	633,123	1,258,952	18,360,600	355,319

This table constructs Lee Bounds around the point estimates reported in Table 2.12. All quantities in UGX; Outliers top/bottom coded to 95th/5th percentile; Heteroskedasticity-robust standard errors in parantheses. *Total indivisible investment includes real land values, adjusted down to exclude appreciation in land values over time. All specifications include the following controls: pre-treatment levels of outcome, hh income, patience, gender, hh head, age, age², num ad females, num ad males, num children, district FE's. * $p < 0.1$, ** $p < .05$, *** $p < 0.01$

Appendix 2.E Multiple Hypothesis Corrections

We apply multiple hypothesis corrections to the estimates in Tables 2.8, 2.10, and 2.12 (which also include the aggregate investment categories from Tables 2.6 and 2.7). Specifically, we estimate False Discovery Rate (FDR) sharpened q-values, per [Anderson \(2008\)](#), and apply one penalty at each endline across the six household budget categories on which we test for an effect of the small and large lottery. We separately penalize the set of hypotheses concerning the effect of winning the small lottery and the effect of the large lottery.

Table 2.E.1: Grant effects on components of the household budget constraint - First endline - Multiple hypothesis corrections

	(1) consum. (stock) wkly cons + home dur.	(2) savings	(3) div. investment	(4) indiv. investment*	(5) mthly income	(6) net credit
won lottery (0/1)	121,277 (0.18) [0.32]	57,204* (0.06) [0.18]	190,956*** (0.00) [0.00]	-721,095 (0.43) [0.56]	-8,344 (0.68) [0.56]	-7,054 (0.70) [0.70]
won large lottery (0/1)	-18,497 (0.92)	107,836 (0.12)	127,754 (0.40)	5,016,354** (0.01)	38,551 (0.38)	-23,513 (0.56)
risk loving (0/1)	110,406 (0.31)	12,933 (0.71)	76,005 (0.32)	-2,326,259** (0.02)	-4,074 (0.88)	-788 (0.97)
$\beta_1 + \beta_2$	102,781	165,040	318,710	4,295,259	30,207	-30,567
P-value: $\beta_1 + \beta_2 = 0$	.5	.0082	.02	.013	.44	.39
FDR sharpened q-value: $\beta_1 + \beta_2 = 0$	.33	.041	.041	.041	.33	.33
Control mean if risk loving = 0	1,608,566	259,468	613,887	15,232,971	327,076	73,180
Control mean if risk loving = 1	1,798,491	271,428	787,251	17,160,718	321,271	80,220
R ²	.29	.3	.41	.71	.42	.093
Observations	867	867	867	867	867	867

This table includes p-values in parentheses and FDR sharpened q-values in square brackets, corresponding to the point estimates in Table 2.8. All quantities in UGX; Outliers top/bottom coded to 95th/5th percentile; Heteroskedasticity-robust standard errors in parantheses. *Total indivisible investment includes real land values, adjusted down to exclude appreciation in land values over time. All specifications include the following controls: pre-treatment levels of outcome, hh income, patience, gender, hh head, age, age², num ad females, num ad males, num children, district FE's. * $p < 0.1$, ** $p < .05$, *** $p < 0.01$

Table 2.E.2: Grant effects on components of the household budget constraint - Second endline - Multiple hypothesis corrections

	(1) consum. (stock) wkly cons + home dur.	(2) savings	(3) div. investment	(4) indiv. investment*	(5) mthly income	(6) net credit
won lottery (0/1)	294,543 (0.24) [.88]	48,839 (0.14) [.88]	69,303 (0.25) [.88]	506,277 (0.63) [.88]	16,240 (0.39) [.88]	-20,703 (0.39) [.88]
won large lottery (0/1)	386,636 (0.41)	84,891 (0.23)	-27,480 (0.84)	4,770,034* (0.06)	7,896 (0.85)	8,687 (0.87)
risk loving (0/1)	157,957 (0.59)	3,760 (0.92)	86,905 (0.26)	-1,002,692 (0.41)	27,946 (0.25)	-44,824 (0.14)
$\beta_1 + \beta_2$	681,180	133,730	41,822	5,276,311	24,136	-12,016
P-value: $\beta_1 + \beta_2 = 0$	.091	.033	.74	.02	.51	.79
FDR sharpened q-value: $\beta_1 + \beta_2 = 0$	.14	.11	.65	.11	.62	.65
Control mean if risk loving = 0	4,603,540	273,166	596,244	15,040,343	264,233	119,272
Control mean if risk loving = 1	5,036,368	279,987	776,076	17,326,575	289,215	76,352
R ²	.28	.23	.32	.51	.29	.069
Observations	867	867	867	867	867	867

This table includes p-values in parentheses and FDR sharpened q-values in square brackets, corresponding to the point estimates in Table 2.10. All quantities in UGX; Outliers top/bottom coded to 95th/5th percentile; Heteroskedasticity-robust standard errors in parantheses. *Total indivisible investment includes real land values, adjusted down to exclude appreciation in land values over time. All specifications include the following controls: pre-treatment levels of outcome, hh income, patience, gender, hh head, age, age², num ad females, num ad males, num children, district FE's. * $p < 0.1$, ** $p < .05$, *** $p < 0.01$

Table 2.E.3: Grant effects on components of the household budget constraint - Third endline - Multiple hypothesis corrections

	(1) weekly consumption	(2) home durables	(3) savings	(4) div. investment	(5) indiv. investment*	(6) mthly income
won lottery (0/1)	-818 (0.85) [1]	-2,455 (0.98) [1]	19,243 (0.73) [1]	15,687 (0.89) [1]	570,089 (0.60) [1]	-15,468 (0.59) [1]
won large lottery (0/1)	13,470 (0.14)	-80,403 (0.72)	32,624 (0.81)	16,592 (0.95)	-563,232 (0.81)	129,367** (0.04)
risk loving (0/1)	4,318 (0.44)	14,317 (0.92)	175,644** (0.02)	95,343 (0.52)	2,127,393 (0.13)	-34,305 (0.29)
$\beta_1 + \beta_2$	12,652	-82,858	51,867	32,280	6,857	113,900
P-value: $\beta_1 + \beta_2 = 0$	.12	.67	.69	.89	1	.042
FDR sharpened q-value: $\beta_1 + \beta_2 = 0$	.43	1	1	1	1	.34
Control mean if risk loving = 0	74,243	1,176,056	452,420	1,064,878	14,388,904	388,771
Control mean if risk loving = 1	81,032	1,304,103	633,123	1,258,952	18,360,600	355,319
R ²	.13	.18	.12	.18	.33	.18
Observations	838	838	838	838	838	838

This table includes p-values in parentheses and FDR sharpened q-values in square brackets, corresponding to the point estimates in Table 2.12. All quantities in UGX; Outliers top/bottom coded to 95th/5th percentile; Heteroskedasticity-robust standard errors in parantheses. *Total indivisible investment includes real land values, adjusted down to exclude appreciation in land values over time. All specifications include the following controls: pre-treatment levels of outcome, hh income, patience, gender, hh head, age, age², num ad females, num ad males, num children, district FE's. * $p < 0.1$, ** $p < .05$, *** $p < 0.01$

Appendix 2.F Description of Grant Purchases

The following table combines first and second endline data to report grant use for as many respondents as we were able to re-survey: when respondents were surveyed at both endlines, we take the grant use reported at second endline, allowing for the most recent update on spending. Many respondents purchased multiple different items with the grants, but these tables reflect mutually exclusive categories, where we depict the single item on which they spent the greatest fraction of their grant funds:

Table 2.F.1: Lottery winner grant uses

	Percent among large lottery recipients	Percent among small lottery recipients
Purchased land	32	6
Business inventory	18	20
Land/building improvements (includes irrigation, solar, and iron roofs)	14	9
Business durables (includes vehicles for business use)	10	6
Small livestock	5	21
Cattle	5	2
School fees	4	9
Household durables (non-vehicle)	4	4
Savings	3	2
Hospital or funeral fees	3	2
Paid down debt	1	2
Farming inputs	1	6
Hired labor	1	0
Farming equipment	0	4
Regular consumption (food, transportation, precautionary health)	0	2
Rented land	0	2
Lost/stolen/did not receive	0	1
Vehicle (not for business)	0	1
Lended out	0	0
Total:	100	100

Besides land purchases, some commonly cited specific examples of purchases made with the grants (both large and small) include:

- Goats, pigs, and chickens
- Coffee seedlings / coffee plants
- Water tank / irrigation drum (for collecting rainwater)
- Iron sheets (as a roof material)
- Solar panels and batteries

- Motorcycle or bicycle, often for delivery
- Inventory for retail, such as soap, salt, and coffee

Appendix 2.G Model Validation Regressions

In this Appendix, we evaluate the model’s ability to rationalize and reproduce our key empirical finding on the impact of lottery winnings. Specifically, we ask whether, first, the model replicates the fact that winners of the large lottery make large, indivisible investments and even in excess of their winnings. Furthermore, we ask whether our use of the SUR cross-equation budget constraints can be justified through the lens of the model.

Clearly, there are other empirical findings along which the parsimonious model is limited. We note two important examples of limitations. First, with only one investment good, the indivisible good which we will interpret as land, the model cannot replicate the investment in agricultural and business assets that small lottery winnings yield. Second, in the data capital gains appear to be an important return to investment in land, but in the model the returns to the indivisible investment are through realized income. Realized income leads to higher levels of both savings and consumption. Hence, we focus on capital investment in the model and compare it to land investment in the data.

To do so, we generate 500 samples of individual simulations from the ergodic distribution in model. These samples are identical in size to our empirical sample, 867 agents, and we have the full series of simulated monthly data for consumption, savings, income, and capital investment, with 16 months between baseline and midline. At midline, we simulate lotteries with lottery choice proportions and lottery winning proportions matching the field experiment via construction. We then simulate 4 months between midline and first endline, and 18 months between midline and second endline to again match the empirics.

Using these simulated data, we run Monte Carlo regressions analagous to those in Equation 2.6, with slight necessary modifications given the model. First, with only one capital good, we have only a single investment outcome (rather than agricultural investment, business investment, and land as separate outcomes). We focus on land investment as the empirical comparison, since business investment and agricultural investment are less discrete.⁴⁰ Second, we have no demographic controls other than age and age², and no geographic controls. Third, in the application of the cross-equation restriction (Equation 2.7, for our SUR regressions with cross-equation restrictions), we omit borrowing, which is zero by assumption.

Finally, in constructing our income and consumption outcomes, we distinguish between the true measure in the budget constraint (which we call “True Values”), an approach that more closely mirrors our empirics (“Empirical Proxy”), and an approach which introduces true measurement bias into the data used (“Biased Measurement”). “True Values” constructs income and consumption as sums of the full series between midline

⁴⁰Other work has emphasized the indivisibility or minimum scale of livestock that can lead to poverty traps, e.g., [Balboni et al. \(2022\)](#); [Barrett, Carter, and Jean Paul Chavas \(2019\)](#).

and (first) endline. “Empirical Proxy” constructs them as we do in the empirics, using the endline value multiplied by the number of months between midline and endline.⁴¹ The point of “Empirical Proxy” is that our measurement at a point in time automatically introduces some deviation from the budget constraint.

Lastly, we account for two features of the empirical data not present in the model. First, the model has only one value for capital (land) investment and so, the precision of the estimates is quite high. In the empirical data, land purchases are of varied size and also likely reflect considerable measurement error. Second, expenditures may suffer not only from classical measurement error but actual bias. To mimic this, we consider multiplicative measurement error, multiplying land realizations in the simulated data by the random variable, $X = \tilde{X}\mu_e$, where $\tilde{X}$ is a log-normally distributed random variable with mean of one and log variance of ν_e^2 . Here $\nu_e = 1.95$ is calibrated to match the average standard error on the coefficient for winning the large lottery to our empirical standard error. The coefficient $\mu_e = 2.2$ captures the bias in measurement and its value is calibrated to equal the average ratio of expenditures to income in the pre-experimental data, much of which may reflect overstatement of land values as discussed in Section 2.5.3. We call this alternative set of simulated data “Measurement Error”.

Table 2.G.1 presents a summary of the results of these regressions on simulated data focusing on the investment outcome and comparing those to the empirical estimates for indivisible investment in Column (4) of Table 2.8. The top panel presents results for the estimates of the impacts of winning the (small) lottery. In the first two columns, we see that across all estimation techniques the mean coefficients average small positive numbers, but the mean standard errors dwarf the mean coefficients. Using the Measurement Error data, standard errors on the OLS coefficient is comparable to that in the empirics (847,510 vs. 906,426). Looking at the SUR budget constraint, we see that the mean p-value for rejecting the constraint on winning the small lottery are high, and it is rejected in roughly 5% of the samples.

The lower panel shows results for the incremental impacts of winning the large lottery on indivisible investment, which are of greater interest. The model is able to generate the surprisingly large coefficients on winning the large lottery that (based on point estimates) indicate that the additional expenditures on indivisible investment actually exceed the incremental grant winnings of 1,350,000 UGX. Using the True Values, the estimated coefficients average roughly 2,000,000 with very small standard errors (11,837).⁴² These values are actually somewhat above the confidence bands of the SUR estimate in Table 2.9 of 1,247,456 (standard error: 277,484). Moreover, the SUR estimation for the True Values yields high average p-values and the bud-

⁴¹For consumption, we multiply endline values times the number of months *minus one* and then add consumption from the month directly after the midline. The empirics combines the past week’s consumption with durable purchases since midline. Since these durables are likely purchased upon receiving winnings, we choose this as the closest analog because immediate consumption after winnings is generally highest.

⁴²For all estimations and simulated data, the realized 95% confidence intervals for the distributions of the estimated coefficients closely reflect those expected by the standard errors of the coefficients.

get constraint is rarely rejected. Using the SUR on the Empirical Proxy, the average coefficient is virtually identical.⁴³ However, the average p-value of the SUR budget constraint is 0.16 for those winning the large lottery, and the constraint is rejected in 54% of the regressions. Finally, we turn to the Measurement Error regressions. By construction, the average OLS standard error in the estimates equals the empirical value of roughly 1,950,000. The estimated coefficients are much larger with the biased measurement error, averaging roughly 4,100,000 in the simulations, which compares well with the roughly 5,000,000 in the empirics, especially given the large standard errors. Focusing now on the SUR estimates for the Measurement Error data, we see that applying the SUR budget constraint yields coefficients close to the true coefficient in the model of just over 2,100,000 and with a much smaller average standard error of roughly 210,000. These patterns mimic those in the actual data. Indeed, this coefficient in the empirics is statistically indistinguishable, roughly 1,200,000 with a standard error of about 300,000. The p-values on the large constraint average 0.29, but the constraint is again rejected more often (in 13% of the samples).

Together, these estimates show the usefulness of the SUR in a situation where biased measurement error can lead to unreasonably large estimates. As argued in the empirics, the SUR returns reasonable estimates (true to the actual estimates) and smaller standard errors.

⁴³We omit OLS for the Empirical Proxy, since it only changes consumption.

Table 2.G.1: Model-Simulated Monte Carlo Regressions

	Won Small Lottery			
	Mean Investment Coefficient	Mean Standard Error	SUR Budget Constraint	
			Mean p-Value	Rejected at 5%
Simulated Data, Estimation				
True Values, OLS	3,591	5,478	.	.
True Values, SUR	3,578	5,651	0.56	0.04
Empirical Proxy [†] SUR	3,617	5,639	0.50	0.05
Measurement Error ^{††} OLS	391,660	898,616	.	.
Measurement Error ^{††} SUR	-2,836	97,935	0.48	0.04
	Won Large Lottery			
True Values, OLS	2,030,588	11,837	.	.
True Values, SUR	2,030,840	12,208	0.73	0.01
Empirical Proxy [†] SUR	2,030,997	12,185	0.16	0.54
Measurement Error ^{††} OLS	4,073,000	<i>1,944,900</i>	.	.
Measurement Error ^{††} SUR	2,100,600	211,784	0.29	0.13

The natural empirical analogs are the estimates from indivisible investment equation in Tables 2.8 (OLS) and 2.9 (SUR). The coefficients coefficients for winning the (small) lottery are -721,095 (standard error: 906,426) and -78,531 (132,562) for the OLS and SUR estimations, respectively. The coefficients for additionally winning the large lottery are 5,016,354 (1,965,752) and 1,247,456 (277,484) for the OLS and SUR estimations, respectively.

[†] Empirical Proxy data constructs total consumption between endline and midline by multiplying the value of endline consumption times the number of months minus one and adding consumption from the month directly after the midline as we do to construct consumption in the empirics. OLS estimates are omitted since, for investment, they are identical to True Values, OLS.

^{††} Measurement Error data multiplies simulated land realizations by the random variable, $X = \tilde{X}\mu_e$, where $\tilde{X}$ is a log-normally distributed random variable with mean of one and log variance of ν_e^2 . Here $\nu_e = 1.95$ is calibrated to match the average standard error on the coefficient for winning the large lottery to our empirical standard error in Table 2.8 (i.e., 1,965,752). This matching value is italicized. The coefficient $\mu_e = 2.2$ captures the bias in measurement and its value is calibrated to equal the average ratio of expenditures to income in the pre-experimental data.

Chapter 3

Big Fish in Thin Markets: Competing with the Middlemen to Increase Market Access in the Amazon

Viva Ona Bartkus, Wyatt Brooks, Joseph P. Kaboski, and Carolyn Pelnik

3.1 Introduction

The geographic concentration of both economic activity and poverty are well known facts of economic development. Remoteness itself, together with limited access to markets, can magnify poverty and even entrap those in rural areas in poverty. This motivates investment in transportation infrastructure, but large scale infrastructure investments are often prohibitively expensive, fraught with political economy issues, or infeasible for developing countries.¹ Are there alternative, cost-effective ways to spur economic development in rural areas where public infrastructure is limited?

We study exactly such an intervention in one of the most remote settings in the world, the impoverished fishing communities of the Brazilian Amazon. Here boats themselves, including floating stores that visit periodically, are the only access to outside markets. Fishing itself is straightforward; while labor-intensive it does not involve much uncertainty. Instead, the challenge is delivering the good to market. With a

¹For the benefits of large scale infrastructure investments, see, for example, [Donaldson \(2018\)](#); [Lu \(2020\)](#); [Asturias, Manuel Garcia-Santana, and Ramos \(2019\)](#); [Asher and Novosad \(2020\)](#). For the challenges, see [Alder and Kondo \(2020\)](#); [Burgess et al. \(2015\)](#); [Jedwab and Storeygard \(2020\)](#).

seasonally-harvested and highly-perishable good – a huge, pre-historic, air breathing fish called a pirarucu – these communities’ only direct access to the final market would involve multi-day ferry rides to Manaus, the capital city of Amazonas. Instead, they sell to middlemen, part of a cartelized supply chain, who arrive infrequently and with acute market power. The intervention we study is an NGO-initiated but community-financed procurement of boats (together with ice and fuel, the necessary complementary inputs), which enables communities to cut out the middlemen and bring the goods to market themselves. We posit that the intervention may yield direct effects, where communities sell directly to the end market rather than to intermediaries, indirect effects, where communities continue to sell to middlemen but with greater bargaining power, or both.

To evaluate the intervention, we conduct surveys in affected communities combined with other administrative data sources. For the purposes of evaluating the intervention, we benefit from plausibly exogenous variation in which communities received boats: although boats were ordered at the same time, shipwrights were unable to fulfill all orders properly for the 2018 fishing season. Remarkably, we find that the intervention increased income substantially, by 27% over income in the control communities. There is also evidence that it increased consumption, including food expenditures. We find no significant increase in fishing, however – important in a situation where overfishing could plausibly be a concern – but instead an increase in the price received for the fish themselves. Not only do we find direct evidence of welfare improvements, but our cost-benefit analysis shows the boat investment to be cost-effective as well. Extrapolating the additional revenues over time, the boats would potentially pay for themselves in under 3 years, well within the life of the typical boat.

Our impact regressions evaluate the effect of receiving a boat in the endline data, after controlling for baseline and time-invariant characteristics, such as distance to market. Identification of unbiased causal impact estimates rests on the conditional exogeneity of which communities received boats. In balance tests, we find only one statistically significant difference ($p < 0.05$) in our baseline surveys, at a rate that is not out of line with type II errors. This difference, however, is in baseline fishing levels, and so we control for these in our baseline estimation. Moreover, using historical administrative data we find no differential trends in fishing.

We study impacts on household-level income, consumption, travel, and time use. The impact on monthly income of R\$184 is significant at the 5% level. The impact on income is of similar magnitude to the overall impact on monthly consumption of R\$148, although the latter is not significant at conventional levels. Looking in more detail, we find roughly half of this increase is on food expenditures (significant at a 5% level) and an important component of the increase is on gas, although this is only marginally significant. We also find that households are more likely to travel to a distant city. Time use is unaffected.

To discover more about the increase in income, we evaluate the community level impacts based on data gathered from a community-level key informant survey. We find no significant impact on quantities sold, although our estimates are not precise. We do, however, find an increase in price of R\$0.77/kilogram, a 20 percent increase relative to the control (significant at a 5% level). This estimate is consistent with 60 percent of the community increase in pirarucu income coming mechanically from higher prices. Moreover, though not significant, our point estimate suggests that prices paid by middlemen may have increased as well. Indeed, our results are consistent with the intervention yielding an indirect effect in its first year: Communities do not appear to substitute away from selling to middlemen, but they may earn a higher price when selling to middlemen. We suspect that this may reflect a risk mitigation strategy in which communities continue to sell a substantial portion of their pirarucu to known buyers.

If the intervention is cost effective, why would the community members not have implemented such a strategy on their own? We posit that, in the context of their economic environment, the lack of a boat constituted a poverty trap, especially in the presence of a cartel with market power in pricing. Theories of poverty traps often involve fixed costs. To fix ideas, we develop such a theory in which middlemen are capable of financing the large fixed cost of a boat, but poor communities are not. Given market power, they can pay below market prices (or equivalently charge a delivery price substantially above the true cost of transporting goods) that can create a man-made poverty trap, even when the natural cost structure does not lead to a poverty trap. We show how cooperative investment in the boats can allow people to escape this poverty trap both directly, the income of those who purchase the boat increases through lower transportation costs, and indirectly, by lowering the market power of middlemen, thereby improving the terms that remote producers receive.

The rest of this paper is organized as follows. After placing our contribution in the literature, Section 2 gives background to the fishing communities we study, their supply chain, and the intervention. Section 3 presents our empirical results, including balance tests, impact estimates, and cost-benefit analysis. Section 4 presents a simple theoretical explanation of our findings, while Section 5 concludes.

Related Literature

This paper is related to the literature that studies how access to outside markets affects people living in remote regions of developing countries. Often this is studied in the context of transportation infrastructure, such as roads or bridges. These papers often emphasize how market access affects labor markets ([Asher and Novosad, 2020](#); [Brooks and Donovan, 2020](#)), or reduces dispersion in prices across space ([Sotelo, 2020](#); [Van Leemput, 2016](#); [Aker, 2010](#)). Our emphasis on decreasing the market power of buyers in poorly connected

markets is similar to the study of how roads affect competition in [Asturias, Manuel Garcia-Santana, and Ramos \(2019\)](#), who study how a major highway system affected market power in manufacturing firms. Our intervention is similar to roads in that it connects markets and involves high fixed costs, however it differs in a key way: roads allow local agents to access outside markets and for outside participants to more easily reach local markets, but community boats only increase market access for the local agents.

We join an important literature in development that studies the role of middlemen, and supply chains more broadly (e.g., [Burke, Bergquist, and Miguel, 2019](#); [Grant and Startz, 2021](#)). Here we follow an experimental approach like [Bergquist and Dinerstein \(2020\)](#), which finds that Kenyan maize traders face very large start-up costs and capture a great deal of surplus. While their experiment focused on changing the behavior on the part of the middlemen, we focus on giving the communities an alternative to the middlemen.

Our paper also has similarities to [Jensen \(2007\)](#), which studies fisheries in the development context. In both that project and ours, price dispersion had welfare costs that could be mitigated with technology, though the nature of the friction in that case (information) is quite different than in ours (physical transportation). In each case we consider technologies that alleviate those frictions (mobile phones in that case and new boats in ours). We view the key difference between these interventions to be their cost. Once available and known, mobile phones can be widely adopted quickly with minimal need for outside intervention. However, due to the very high cost of boats relative to village income, it may be very difficult for communities to make the investment, even if everyone knows about the technology and knows that it has positive net present value. Like [Jensen and Miller \(2018\)](#), we study fishing boats, but they focus on the shipbuilding industry while we focus on the use of boats by fishermen.

On the theory side, our model extends the literature on poverty traps (e.g., [Banerjee and Newman, 1993b](#); [Piketty, 1997](#); [Aghion and Bolton, 1997](#); [Ghatak and Jiang, 2002](#)) that assumed competitive markets. Using the case of intermediaries, we show how market power can create micro poverty traps, when a technological poverty trap would otherwise not exist.²

Finally, our paper joins a literature that studies questions related to the Amazon rainforest itself, though ours is unusual in that we focus on the economic behavior of people living deep in the jungle. Much of the existing literature focuses on the effects of human activity on deforestation, often in the context of farming (e.g., [Assunção et al., 2017](#); [Pfaff, 1999](#)). Presumably, the literature on economies deep in the jungle is limited by the difficulty of data collection.

²In that sense, it relates the earlier work of [Murphy, Shleifer, and Vishny \(1989\)](#), but investments are not complementary in our model, so escaping the poverty trap does not require “big push” investment.

3.2 Background on the Brazil Amazon, Supply Chain, and Experiment

We study whether or not boats may offer a cost-effective means of increasing access to market for fishermen in 2 remote reserves, Mamirauá and Uacari, in the Brazilian Amazon. Partnering with Fundacao Amazonas Sustentavel (FAS), the leading environmental non-governmental organization in Brazil, we exploit plausibly exogenous variation in the timing in which communities gained access to boats.

3.2.1 Setting

The Amazon communities we study are remote and poor, and yet they have some advantages and remain where they are because of their longstanding lifestyle. Calling themselves “children of the forest,” the river communities living deep in the Amazon are among the most remote in the world. Although the State of Amazonas designated vast areas as protected reserves, such as the Mamirauá and Uacari Reserves, poor homesteading squatters are permitted to build their homes on any river front areas that they carve out of the forest. With a church, one room school, and meeting house, each community numbers from 4-40 families with an average of 17-18 households. Even though by income level these river communities are categorized as extremely poor, the plentiful fish and forest fruits combined with chickens and vegetables adds nutrition into their diet.

For river travel, these communities think not in terms of distance, but rather in terms of travel time or diesel cost to power their boats. As a frame of reference, the main commercial transport ferry travels upriver from the capital city of Manaus, Amazonas (population 2 million) to Fonte Boa (population 37,000) in Mamirauá in 2 days 21 hours. Traveling further upriver to Carauari (population 26,000) in Uacari takes the ferry an additional 3 days 13 hours, making the whole journey from Manaus 6 days 10 hours one way.

The annual pirarucu fishing season represents the main source of income for most river communities. Pirarucu (*arapaima giga*) is an ancient, air-breathing, giant fish native to the lakes in the Amazon River basin. One of the largest freshwater fish in the world, pirarucu regularly grow to 3 meters (10 feet) and 220 kilograms (480 pounds), although most are caught around 2 meters and 75-100 kilograms. The lakes themselves are formed when low water levels of the dry season separate them from the main river, entrapping the pirarucu. Once over-fished and endangered, careful fish management over the last decade has coaxed pirarucu numbers up to such an extent that most communities catch pirarucu annually.

Though technically challenging, hunting pirarucu is a relatively straightforward process for the experienced fishermen, involving little of the yield uncertainty we often associate with fishing. The annual dry

season and water level drop of the Amazon River by 15 meters (50 feet) signals the beginning of the pirarucu fishing season, usually at the end of August to November. Entire communities move inland to their designated lakes for a few weeks to engage in this collective effort. “Hunting” rather than “fishing” better describes the activity. Pirarucu must surface every 20-25 minutes to breathe air, and young men harpoon pirarucu as the fish come up to the surface to breathe while standing in their fragile, flat-bottom kayaks. Once the pirarucu are dragged to shore, all community members work together to clean and prepare the pirarucu and then carry them the 2-3 miles from their lake to the Amazon River.

3.2.2 Existing Supply Chain

While hunting is straightforward, the current pirarucu supply chain and market structure limits the value captured by fishermen through the sale of their fish. The boats of the regional middlemen, with ice-stocked holds of 20-30 ton capacities, arrive at the communities’ river docks to pick-up the catch. These middlemen have relationships with buyers in the regional market towns such as Fonte Boa and Carauri, or alternatively, transport the pirarucu to Manaus for sale to the eight fish processors (*frigorificos*). The ultimate consumer, the Amazonian public, considers pirarucu a delicacy and part of a traditional family meal to celebrate Holy Week between Palm Sunday and Easter every spring.

Figure 3.1 displays the prices received at each step of this pirarucu value chain in Mamirauá in terms of reais/kilogram (R\$/kg). Our primary focus in this paper is getting the product from fishermen to “regional buyers” or “local fish processors” (i.e., frigorificos) located in Manaus. First, note that without a large boat, fishermen have two options to sell fish: middlemen and local buyers (in either Tefe or Fonte Boa).³ Without a larger boat the latter option involves the risky and/or costly work of bringing the large, perishable fish into market towns using public ferries or the fishermen’s own smaller boats or kayaks. The price received is similar, but delivery to local buyers is costly. Delivery does not benefit from returns to scale, and ice and fuel are expensive in remote areas. Without ice there is significant risk of spoilage. Consequently, this option is less profitable and is used when middlemen do not visit. It is more common for smaller communities that middlemen visit less frequently. We will find (see Table 3.4) that just over half of the fish in the control are sold to middlemen. Second, note the seven-fold difference between the price received by fishermen in the communities and the price received by “local fish processors.”

The communities in Mamirauá and Uacari capture only about one sixth of the overall value created by the pirarucu they catch and only about 30% of the fish’s pre-processed value. The sizable gaps between upstream and downstream prices in Figure 1 are not themselves evidence of market failure, however, and

³A third option is to consume the fish themselves, but this is not common. Pirarucu are treated as a cash product, and the local diet consists more of eggs, vegetables, fruits, and other river fish (like tambaqui, which are plentiful). Selling directly to regional buyers in Manaus is prohibitive without the larger boats.

accusations against middlemen are not always justified. Indeed, middlemen can play vital roles in economies, especially in remote areas, by connecting distribution channels, making risky investments, taking on price risk, and establishing new markets. Although gaps between purchase and sales price may be substantial, they may reflect high costs and risk, and competition and free entry can bid rents down to zero. We therefore review contextual evidence of collusive behavior as well as the cost structure and profitability of middlemen in the market.

We begin with contextual evidence of collusive behavior. The market is organized annually at the Tefe Pirarucu Business Meeting (held annually in June/July, i.e., directly before the pirarucu fishing season begins in late August) brings together the eight *frigorificos* located in Manaus, and 8-10 viable middlemen operators with appropriate cold storage transportation boats in Mamirauá and 6 middlemen for Uacari. (Tefe itself is a small city located intermediate between the closer markets and Manaus.) The middlemen do not live in the river communities, but rather urban areas, including Manaus.

During this annual meeting, through the negotiation between the *frigorificos* and the middlemen, target prices per pirarucu kilogram are established for the fishing season, both the prices middlemen will pay fishing communities and the price they will receive from frigorificos in Manaus.⁴ Based on Tefe municipal records, these meetings have been held annually since 2004.

Hundreds of river communities in Mamirauá and Uacari sell their pirarucu to this limited number of middlemen. The river is divided up amongst the middlemen by custom (and distance so middlemen take a cluster of communities) so that two middlemen rarely go to the same river community's dock. At present, no contract or other type of guarantee governs the communities' sale of pirarucu to middlemen. The middlemen pay the communities either in cash or leave behind a promissory note. The middlemen only earn their money after they land the pirarucu in Manaus and complete the sale to the frigorificos. Many river communities relay experiences of how the middlemen do not always return to the river communities later in the year to settle their debts for the previous fishing season.

Next we review the profitability of middlemen. Table 3.A.1 in Appendix A includes an estimated basic income statement for the middleman carrying pirarucu from Mamirauá reserve to markets in larger cities like Manaus.⁵ The total cost per kilogram of pirarucu delivered is R\$1.03, inclusive of labor (captain and crew), in-kind labor compensation (i.e., meals), fuel, ice, and cost of capital. Figure 3.1 and our data indicate the price paid by middlemen at the community dock was R\$3.5-\$3.6 per kilogram of pirarucu. Thus, their total costs are roughly R\$4.63 per kilogram. The middlemen can sell the pirarucu to regional buyers (in Tefe or

⁴To the knowledge of the staff of the local NGO, Foundation for Amazonas Sustainability, no representatives of the fishing communities have been invited or been able to participate in this annual pre-fishing season meeting.

⁵The analysis is based on research of the consultants from the University of Notre Dame's Mendoza College of Business as part of the collaboration explained below.

Fonte Boa) for R\$8.5 per kilogram and to buyers in Manaus for R\$12. This price range is therefore enough to cover their costs two to three times over. There are of course logistical risks, but for pirarucu in the Manaus markets, there remains very little downside pricing risk for the middlemen, as the price for kilogram has remained constant at R\$11-13 for the fishing seasons of 2016-2021.

A natural question is why entry does not drive profits down. Perhaps middlemen require a combination of intensely local knowledge and capital that is simply rare. Moreover, potential contracting issues may prevent separate parties from coming together with local knowledge and capital, respectively. We view the results of the intervention below as evidence that entry is indeed profitable.

3.2.3 Intervention

Our NGO is partner, Fundacao Amazonas Sustentavel (FAS)—the leading environmental non-governmental organization in Brazil—focuses on the sustainable economic development of both the Amazon’s natural resources and its river communities. Through its health, education, infrastructure, environmental conservation, and income generation programs, FAS serves 9,430 families (39,460 people) living in 16 protected reserves in the State of Amazonas. Since 2008, its flagship La Bolsa Floresta program, sponsored by the Fundo Amazonia, Bradesco Bank, and the State of Amazonas, pays a modest subsidy of R\$50/month to households in river communities who promise to care for the forest in an attempt to “make the forest worth more standing than clearing.”

In 2017-2018, FAS collaborated with the University of Notre Dame’s Mendoza College of Business to conduct a thorough examination of the opportunities presented by the existing pirarucu market structure. Based on this analysis, FAS decided to support further income generation of river communities through investment in larger transport boats. These investments in boats would improve access for pirarucu fishermen to regional markets. Starting in early 2018, through a collaborative process based on workshops with the communities, FAS first organized Mamirauá and Uacari fishermen into associations of 8-10 communities and then supported them in the acquisition of boats. These FAS boats would carry the communities’ pirarucu to regional market towns such as Fonte Boa (in Mamirauá) and Carauari (in Uacari) rather than depending on middlemen. La Bolsa Floresta Program financed the purchase of the FAS boats and FAS staff time and expenses. FAS Supervisor, Edvaldo Correa de Oliveira described the challenge:

Before the FAS boats, the communities were held ‘hostage’ to the middlemen. There have been times when a community has agreed with the middlemen to arrive on their docks on a particular day in November when they pre-negotiated a price for R\$4.5 or R\$5 per kilogram of pirarucu. When the middleman arrives, he only offers them R\$3.5 per kilogram. The communities have no

ice, no storage, no transportation, and thus, no choice but to sell at the lower dictated price.

Improving the Amazon fishermen's access to local and regional markets was FAS's main objective, with the most significant investment in larger transport boats. However, reducing dependence on the middlemen's cartel also required two further initiatives: first, coordination with municipal leaders for permissions to create outdoor markets for the sale of pirarucu during the fishing season (August-November, 2018); and second, investment in cold storage facilities in Fonte Boa and Manaus to enable pirarucu sales to customers during peak demand 4-5 months after the fishing season during Holy Week in the Spring of 2019. As outlined in Figure 3.1 of the Pirarucu Supply Chain, the fishermen could expect higher prices per kilogram of pirarucu as they moved down the value chain beyond the middlemen, to sell to local buyers and regional buyers.

However, this overall intervention to improve market access in a single fishing season also presented risks. These risks included the uncertainties of new channels (their own boats), selling to new customers in regional markets with whom they had no relationships yet (rather than existing middlemen); and new delayed timing of payment (upon sale of pirarucu 4-5 months after the fishing season). Given these combined risks, one could expect poor fishing communities to engage in risk mitigating strategies, such as continuing to sell a portion of their pirarucu through the existing channels of known middlemen. During the first year of the intervention, fishing communities could also be expected to build relationships with local and regional buyers of pirarucu and develop approaches for pirarucu sales in March/April during Holy Week. And thus, as communities gain experience of these novel changes in market access, it could be expected that fishing communities would shift more of their pirarucu sales toward high priced channels and customers in subsequent years. The first fishing season of these combined changes in market access could reasonably be thought of as the floor of the possible impact of the intervention, with increases in prices and household income in future fishing seasons.

The intervention has an element of a one-year natural experiment in the way that some communities received boats and others were delayed by one year. The Mamirauá and Uacari community associations, supported by FAS, contracted with sellers in Manaus, Tefe, and Fonte Boa for 8 boats for the 8 Mamirauá associations and 1 boat for the Uacari association. All 9 boats were ordered at the same time. However, as of the start of the pirarucu fishing season in August 2018, only 6 of the 8 boats were delivered to the Mamirauá associations.⁶ Both of the remaining 2 Mamirauá sectors did receive their boats in time for the following year's (2019) pirarucu fishing season. The failure to deliver these boats on time in effect created a source of novel variation in Mamirauá. In our sample, 18 communities could transport their catch on their FAS boats

⁶More specifically, when the boat was delivered to one of the Mamirauá associations, its electrical and other systems on board were not functioning properly. The communities refused to take receipt of their boat until those systems were fixed. The seller agreed to the repairs. Furthermore, a second Mamirauá association did not receive its ordered boat due to a backlog at the boat builder.

to sell in regional markets, while 14 communities with no access to a FAS boat for the 2018 pirarucu fishing season were forced to continue their existing practices of selling to the middlemen. In Uacari, the late arrival of the FAS boat from the shipwright seller in the 2018 pirarucu fishing season created a circumstance in which the 8 Uacari communities in our sample transported their pirarucu to regional markets with the FAS boat in 2018, while the other 4 communities did not.⁷ Like in Mamirauá, all sectors in Uacari did receive a boat in time for the 2019 fishing season. We evaluate the extent to which there were pre-existing differences or trends in our balance tests of the next section.

3.3 Empirical Analysis

To evaluate the effects of the boats, we combine our own household and key informant surveys collected in 2018 with secondary data from FAS, the Mamirauá Institute, and Associacao dos Produtores Rurais de Carauai (ASPROC). Both the Mamirauá Institute and ASPROC are environmental research institutes in the Amazon. After checking for balance on community baseline characteristics, we proceed with the assumption of conditional exogeneity of boat assignment, and we estimate impact regressions to evaluate the effect of the boat on household income and consumption as well as the pirarucu prices, quantity, and revenue attained at the community level.

3.3.1 Data

Our sample consists of 308 households across 44 communities, of which 26 receive a boat in time for the 2018 fishing season (the treatment group). Of the 308 households surveyed, 194 live in treated communities. We rely on a natural rather than controlled experiment, and so we did not conceive this experiment *ex ante*. We therefore lack baseline survey data, an important limitation. However, we combine several data sources *ex post*, following the allocation of the boat treatment during the 2018 season, to establish the conditional exogeneity of boat receipt and to estimate the impact of the boat treatment on household financial outcomes and community pirarucu sales. Our empirical analysis relies on a household and key informant survey (an interview with the community fishing leader) conducted in late 2018, following the 2018 fishing season when some communities received boats on time while others did not; phone calls to fishing leaders to retrospectively collect information about the 2017 fishing season; operational records from our NGO partner, FAS; a community-level survey collected by FAS prior to the 2018 fishing season; and historical fishing levels documented by environmental research institutes in the Amazon between 2013 and

⁷In Uacari, the Amazonas state government arbitrarily decided that only communities in the state (non-federal) reserves should have a FAS boat in 2018, while all communities in both state and federal reserves would have the FAS boat in 2019.

2017 for some communities in our sample. Because the natural experiment created unexpected variation that we utilize *ex post*, these data on historical fishing levels for some communities in our sample are the only true panel data that we are able to access for this study. The decision to collect primary data via a household and key informant survey of the community fishing leader takes place after we realize the variation in the timing of boat delivery, essentially enabling us to collect primary data in a single endline survey wave while establishing pre-intervention balance with secondary data sources. We provide a detailed description of each data source and its use in the Data Appendix (Appendix B).

Together, the household and key informant surveys, both reflecting the post-2018 harvest season, provide information on our outcomes of interest. The household surveys compose our only data source at the individual-level and contain information on family income, consumption, time use, and travel. The key informant survey of the community fishing leader contains information on the community’s 2018 pirarucu catch and revenue. We construct household income in 2018 as the total annual income recorded by households in the 2018 household survey plus the average income per household earned if the community sold pirarucu to FAS, which FAS then sold during the 2019 Holy Week, a time of high demand for pirarucu.⁸ From the key informant surveys, we construct the pirarucu price accessed by the community, based on the reported pirarucu quantity and revenue that the community achieved during the 2018 harvest.

The rest of our data sources are used to assess pre-intervention balance and test for any preexisting differences in fishing trends. In late 2019 and early 2020, we contact the community fishing leaders via phone to collect information about the 2017 pirarucu catch and revenue.⁹ We use FAS’ operational records to collect information about community population, remoteness in terms of distance to the closest market, and eligibility for the Bolsa Floresta social assistance program. We use a community-level survey conducted by FAS prior to 2018 on access to health and education infrastructure, such as whether a community has a school and employs health workers, to understand pre-intervention sociodemographic characteristics in our sample. Finally, we use longitudinal data on historical fishing levels, from the Mamirauá Institute and ASPROC, environmental research institutes in the Amazon, to test for preexisting differences in fishing trends. These historical fishing levels are documented at the sector-level and are available for four of the

⁸We collect incremental income from the FAS 2019 Holy Week Fair, which reflects income earned from fish caught during the 2018 pirarucu season. The pirarucu harvest occurs once per year and so we calculate annual income, inclusive of the incremental income from the fair, and then divide by 12 to produce average monthly income. However, we were not able to repeat this data collection for the 2019 pirarucu harvest. As all communities had received boats for the 2019 fishing season, we planned non-experimental data collection as a follow-up. Because much of the income of the 2019 season is realized in the Holy Week of 2020, data collection unfortunately had to be abandoned when Manaus was one of the hardest hit cities in Brazil with the COVID-19 pandemic.

⁹We learn from FAS that many of the fishing leaders retain strong records or recollections of recent years’ pirarucu harvests, so we initiate this data collection exercise in the hope of acquiring data on pirarucu quantity and revenue in 2017, the pre-intervention harvest season. Unfortunately, this effort was also cut short by the COVID-19 pandemic, which hit the Brazilian Amazon especially hard. We are able to reach 28 of the 44 community fishing leaders.

eight sectors in our sample.¹⁰

Table 3.A.2 in Appendix A contains summary statistics on our household sample. The respondents are predominantly male (65%) and middle-aged (average age is about 40), with low education levels. Only 31% of the respondents have completed primary school. They are also heavily engaged in fishing (85% of the sample) and farming (78% of the sample). Two-thirds are eligible for FAS' flagship social assistance program, La Bolsa Floresta, indicating that they tend to be quite poor and engaged in livelihoods that FAS deems important for the Amazon's environmental sustainability. Because our household survey data consist of a single cross-section collected after the intervention, we summarize any potentially time-variant characteristics across the control group only ($N = 114$). Among the control group, the average monthly income is R\$670 (or about 180 USD).¹¹ Average monthly spending on food amounts to 65% of household income, or R\$435. Only 35% of control respondents own a radio, and only 54% own a cell phone. Households allocate about 30 hours per week to fishing and 20 hours per week to farming. They typically make 17 trips to a nearby city per year and just under one-half of the sample report regular travel to a distant city.

3.3.2 Balance

The variation in our data does not come from a randomized experiment. Hence, it is particularly important to assess pre-intervention balance, comparing time-invariant and pre-2018 characteristics between communities that received boats as expected in 2018 and those that did not. In general, balance is good: we detect no statistically significant differences between treatment and control communities on the number of households per community, eligibility for Bolsa Floresta, distance to market, presence of a school, number of teachers, presence of health agents, number of health agents, presence of a water ambulance, or access to water treatment (Table 3.1).

Given the balance, we can use these data as descriptive statistics of our communities. On average, communities in our sample are home to 17-18 households, and about 80% of the communities contain households who are eligible for La Bolsa Floresta. Communities are quite remote – the closest market is typically about 10 hours away by boat (in many communities, travel time is a better indicator of effective remoteness than distance as it reflects differences in currents and transportation infrastructure). About 80-90% of communities have their own school, and there are typically 2-5 teachers per community (0.11-0.26 teachers per household) and 1-2 health agents per community (0.04-0.09 health agents per household). About 20% of the communities in our sample have access to a water ambulance, and about 80% have access to water treatment.¹²

¹⁰Sectors represent the associations of communities through which boats were purchased.

¹¹This assumes an exchange rate of R\$3.7 = 1 USD at the time of the survey in October/November 2018.

¹²We sometimes have missing values in community characteristics, if FAS did not collect data from a given community in

We do, however, find one important difference in our data, though this is not out of line with what one would expect to find even if treatment were randomized. Among the community fishing leaders from whom we were able to collect information retrospectively about the 2017 harvest season, we find that treatment communities harvested a greater quantity of pirarucu in 2017 relative to control communities (significant at the 5% level). Given this finding, we control for pre-treatment differences in pirarucu levels. Because pirarucu quantity is an outcome of interest itself, we also examine whether there is a difference in trend between treatment and control communities using historical fishing levels between 2013 and 2017. We find no evidence that this is the case.¹³ We show that the pirarucu harvest in sectors that receive a boat does not change differentially over time relative to sectors that do not receive a boat (Table 3.A.3 in Appendix A).

In sum, given the results on balance, we proceed with an assumption of exogeneity of treatment, conditional on a set of household and community controls explained below.

3.3.3 Impact on Households

We estimate OLS regressions to measure the impact of the boat on household income, consumption, travel to market centers, and time allocation. We specify the following equation for each household i in community c in reserve r :

$$Y_{icr} = \beta boat_c + \delta X_i + \gamma X_c + \lambda_r + \varepsilon_{icr} \quad (3.1)$$

In this model, β is the coefficient of interest, measuring the impact of the boat, allocated at the community level, on household outcome Y_{icr} . We control for household characteristics X_i , which include the number of household members, and community characteristics X_c , which include travel time (in hours) to the closest market and its square (“remoteness controls”), and the interactions of these remoteness controls with the reserve $r \in \{\text{Mamirauá, Uacari}\}$. Following our results on balance, we include the quantity of pirarucu harvested by the community in 2017 as a control.¹⁴ We also include reserve fixed effects, λ_r . Additionally, we winsorize household data to the 5th and 95th percentiles and cluster standard errors at the community

their health and education survey. We separately test whether the rate of missing data varies differentially between treatment and control communities and find no evidence that this is the case.

¹³The average fishing level in 2017 reflected in this time series analysis does not perfectly correspond with the average fishing level in 2017 reflected in Table 3.1, our balance table, because each relies on a different data source. The balance table reports fishing levels in 2017 from retrospective phone calls to fishing leaders in 28 of 44 communities (reflecting both sectors in Uacari and 7 of 8 sectors in Mamirauá) while the time series on fishing levels comes from environmental records, which were only available for 4 of the 8 sectors in Mamirauá. While the incompleteness of these data is not ideal, it reflects the challenge of collecting information in a region as remote as the Brazilian Amazon, especially during the COVID-19 pandemic, when the region was hit especially hard. To our knowledge, we are the first people to access and compile data to this extent on a supply chain in the Amazonas.

¹⁴To maintain a constant sample, we equate missing values in a community’s 2017 pirarucu quantity to the mean, and we include a flag to indicate that the value was missing in a given community. The quantity of pirarucu harvested in 2017 is not missing differentially between treatment and control.

level. We also report randomization inference p-values in brackets below clustered standard errors, given that the optimal level of clustering is unclear. Decisions about fishing and sales are made at the community level, but boats were purchased at the sector level. In contrast to standard errors clustered at the community level, the randomization inference p-values are based on the 10 community groups that ordered boats (6 sectors in the Mamirauá reserve and 1 sector in the Uacari reserve that received boats on time, and 2 sectors in Mamirauá and 1 sector in Uacari that did not). Our discussion of results below assumes the precision indicated by clustered standard errors.

We find that the boat results in a statistically significant and sizable increase in household income (Table 3.2). Households in communities with a boat experience a 27% increase in income, an additional R\$184 per month on the control mean of R\$670 per month (significant at a 5% level). We find an effect of similar magnitude on monthly consumption, R\$148, a 17% increase, though this estimate is not significant at conventional levels. We further disaggregate the components of consumption into spending on food, gas, diesel, and transport. We find that the boat results in higher monthly spending on food and gas. Households in communities with a boat spend an additional R\$68 on food each month, relative to the control mean of R\$435, a 16% increase (significant at a 5% level). This increased food expenditure accounts for roughly half of the observed increase in total consumption. Households with a boat also spend an additional R\$42 on gas, relative to the control mean of R\$200, a 21% increase (significant at a 10% level). Gas is used both for cooking and as a boat lubricant. This suggests that households in communities with a boat spend a substantial portion of the incremental income on direct food consumption and its associated costs, perhaps substituting from less preferred to more preferred food goods.¹⁵

We find a small, negative and imprecise effect of the boat on diesel spending, which is predominantly used as a fuel for motor boats. We see suggestive evidence that households spend more on transport to market centers (above fuel costs). We estimate that households in communities with a boat spend an additional R\$38 per month on transport, relative to a control mean of R\$152, a 25% increase (not significant at conventional levels). This is consistent with our finding that households in a community with a boat are 16 percentage points more likely to travel to distant cities (significant at a 5% level). Travel to distant cities may encompass both personal and business travel, so a positive change on this dimension is not necessarily indicative of a direct effect in which households are more likely to travel to market themselves and sell pirarucu directly to the consumer, bypassing the middlemen. The increase in travel to distant cities, thus, is best interpreted as a measure of welfare in which households gain broad access to goods, services, and markets to which they did not have prior connection.

¹⁵We find that households in communities with a boat are more likely to purchase fish and canned food, which may be preferred goods.

Importantly, we find no statistically significant effect of the boat on households' allocation of labor hours to fishing and farming. The estimated effects on hours allocated to fishing and farming are not only imprecise but also small and slightly negative (Table 3.A.4 in Appendix A). Assuming that the chief impact of the boat is on market access and not labor productivity, this may suggest that higher incomes are not realized through a quantity channel, which is consistent with our community level analysis of the pirarucu market in the next section.

3.3.4 Impact on the Pirarucu Market

To measure the effect of the boat on prices attained, quantity sold, and revenue earned by the community, we estimate the following OLS regression for each community c in reserve r :

$$Y_{cr} = \beta boat_c + \gamma X_c + \lambda_r + \varepsilon_{cr} \quad (3.2)$$

Similar to Equation 3.1, β is the coefficient of interest, measuring the impact of the boat on community outcome Y_{cr} . We control for community characteristics X_c , which include travel time (in hours) to the closest market and its square ("remoteness controls"), the interactions of these remoteness controls with the reserve $r \in \{\text{Mamirauá, Uacari}\}$, the quantity of pirarucu harvested by the community in 2017, and the number of households in the community, if the outcome is not already reported per household.¹⁶ We also include reserve fixed effects, λ_r . When estimating the effect of the boat on 2018 prices, we additionally control for 2017 prices. Like in the household results, we report randomization inference p-values based on 10 community groups in brackets below heteroskedasticity-robust standard errors.

We find a striking impact of the boats on the price received for pirarucu (Table 3.3). Communities with a boat obtain nearly 20% higher prices for their pirarucu, an additional R\$0.77 per kilogram of pirarucu on the control mean of R\$4 per kilogram (significant at the 5% level). This may be driven in part by greater bargaining power with the middlemen: we find that communities are able to negotiate 25% higher prices from the middlemen, raising the cartel price from 3.6 R\$/kg to 4.5 R\$/kg (not significant at conventional levels; p-value = 0.13). Consistent with the narrative put forth by FAS, this suggests that communities are no longer "hostage" to the middlemen's prices, as they now have the ability to store and transport their own product to market and thus a higher outside option. Given the final pre-processing price of R\$12/kg paid by *frigoríficos*, this amounts to the portion of pre-processing value captured by the fishermen rising from approximately 30% to 40%. We find an imprecise negative impact of the boat on the number of middlemen

¹⁶Like in the household level specification, we equate missing values in a community's 2017 pirarucu quantity to the mean, and we include a flag to indicate that the value was missing in a given community.

with whom communities deal, perhaps indicating that communities either sell to other types of buyers or deal with fewer middlemen because they are able to negotiate a higher price more quickly.

In Table 3.4, we show the effect of the boat on quantity and revenue from pirarucu sales. We find a near zero but imprecise effect of the boat on the total quantity of pirarucu per household: 45 kg/household on a control mean of 1,100 kg/household (not statistically significant).¹⁷ Assuming a single pirarucu weighs 75 kg, this amounts to an effect of less than 1 additional fish per household. The lack of an effect on total quantity is consistent with the lack of labor reallocation towards fishing. This may also be consistent with a community view that leaving the once endangered pirarucu in the Amazon is an investment of sorts – savings for the future. We see a larger, but still imprecise, effect on the quantity of pirarucu sold specifically to middlemen. We estimate that communities with a boat sell 32% more pirarucu to middlemen than communities without a boat, an effect of 200 kg/household on a control mean of 620 kg/household (not statistically significant). Given the much smaller effect on total quantity, this may suggest that communities are actually substituting sales to non-middlemen with sales to middlemen. This is consistent with an indirect effect: communities with a boat have greater bargaining power and so they increase relative sales to middlemen, versus other types of buyers, now that they can insist on a higher price. From the middlemen’s perspective, an increase in the quantity purchased by middlemen when prices increase exogenously from new outside option (i.e., a move up the monopsonist’s perceived upward sloping supply curve) would be at least theoretically consistent with the evidence of monopsony power among middlemen discussed in Section 2. From the fishing community’s perspective, continued use of middlemen may also reflect an attempt to hedge risk by maintaining relationships with pre-existing buyers. As communities build relationships with other customers in market towns, they may shift more pirarucu sales towards these higher priced channels in future years.

We find larger, but still imprecise, effects of the boat on revenue from pirarucu, both in total and from sales to middlemen (Table 3.4). Communities increase revenue per household by R\$1,500 on a control mean of R\$4,300, a 35% increase (not statistically significant). They increase revenue from middlemen by R\$1,700 per household on a control mean of R\$2,100, an 80% increase (not statistically significant). These point estimates are imprecisely measured, but the magnitudes are in line with the significant income estimate in the household estimates: the point estimates on pirarucu revenue at the community level account for about two-thirds of the estimated increase in monthly household income.¹⁸ These point estimates also suggest that the increase in prices of R\$0.77 per kilogram mechanically accounts for 60% of the R\$1,500 increase in

¹⁷Pirarucu is harvested at the community level, but we normalize per household to adjust for any differences in community size.

¹⁸The R\$1,500 estimate of community fishing revenue in Table 3.4 is annual, whereas the household income effect of R\$184 in Table 3.2 is monthly, yielding an annual increase of roughly R\$2,200. A 95% confidence interval around the revenue point estimate easily encompasses R\$2,200, and so we interpret any difference between the increase in annual household income and the increase in annual pirarucu revenue per household as statistical noise.

pirarucu revenue.¹⁹

Given our imprecision, we cannot rule out sizable positive impacts on quantities. However, together with the household estimates on time allocation, the results paint a picture in which the effect on total revenue is driven by higher prices rather than increases in the total quantity sold. Meanwhile, the effect on revenue from middlemen seems likely driven by a combination of an ability to negotiate a higher price from the middlemen and therefore selling more to the middlemen.

In sum, the presence of the boat enables communities to earn a higher price on their pirarucu. Communities do not appear to fish more, either in terms of time allocated towards fishing or quantity of pirarucu harvested, and so increases in revenue seem to result predominantly from higher prices. This translates to an increase in household income, which communities primarily spend on food and gas, and greater travel to distant cities.

3.3.5 Cost-Benefit Analysis

The question of the return on investment associated with economic development programs can frequently prove thorny. In this particular experience, however, the diligent data gathering of FAS staff can provide insights. Let us examine the return on investment for an average FAS boat of approximately 10-15 tons, transporting pirarucu in Mamirauá. It cost about R\$150,000 and serves an association of 8-10 communities. Our research on the 2018 pirarucu fishing season indicates that each household with access to a FAS boat earned R\$184 in incremental monthly income as compared to control communities with no access to FAS boats. Since communities average 17-18 households per community, the total annual gain in income of R\$38,000+ per community, for an annual increase in revenue that exceeds the capital expense. However, operational expenses are considerable, and subtracting away operational costs for the FAS boats in terms of gas, diesel, and ice, the payback period on the FAS boat investment solely from pirarucu fishing is estimated just under three years. It is not unprecedented for boats on the Amazon to continue working twenty years. Moreover, FAS boats are used during the offseason to transport basic goods and people which provides additional benefits to the communities.

Given the increased income for communities fishing pirarucu, this development policy could productively be extended to the Amana and Piagaucu-Purus Reserves which would impact thousands of additional fishing households. Both reserves have extensive lakes with pirarucu and experienced fishermen. Potential

¹⁹Based on results from Table 3, this is produced by multiplying R\$0.77/kg (the price premium that communities with a boat earn) and 1,119 kg (the mean quantity of pirarucu sold in the control, i.e., assuming the intervention has no effect on the total quantity of pirarucu fished). Of this, the indirect effect (i.e., the increased revenues from middlemen sales due to a rise in their price) is sizable. Using control sales to middlemen (622 kg/household) and multiplying by the middleman price increase of R\$0.9/kg yields R\$560, while using the higher treatment sales (826 kg/household) yields R\$743. Thus, the indirect effect mechanically would account for 37%-50% of the total increase in revenue. These calculations should be taken suggestively, given the lack of precision of our point estimates.

acquisition of FAS boats would enable these communities to bypass regional middlemen and transport their pirarucu from Amana to the regional center of Tefe and from Piagauçu-Purus to Manaus markets.

3.4 Model

In the previous section we showed strong effects on income from a cooperative intervention to bypass middlemen. Now we rationalize this finding in a model where the market power of middlemen endogenously determines the prices received by communities for their fish. Our dynamic model has simple mechanics that build on earlier work by [Banerjee and Newman \(1993b\)](#) and [Ghatak and Jiang \(2002\)](#). We show that the market power of middlemen can create a poverty trap, which would not exist if markets were competitive. Moreover, we show that a cooperative intervention that resembles the one studied in this paper can eliminate the poverty trap. Furthermore, we show that the response of the middlemen may be critical to escaping the poverty trap.

3.4.1 Fishing Communities

We begin by formally stating the problem of a representative community that sells fish to outside markets. Communities produce a fixed quantity of fish every period. They make decisions over how much of the fish they consume within the community and how much they sell outside the community. Selling fish allows communities to purchase other goods, and to save money for the subsequent period. They choose where to sell the fish. They may either sell to middlemen or they may make a discrete investment to operate a boat. Operating a boat is costly, but allows them to bypass the middlemen and sell at a higher price directly in outside markets.

The community begins each period t with wealth w_t . Each period, communities catch Q units of perishable fish. They can sell their fish to a middleman at a price p_t , or they can transport the fish directly to the market where they will earn a price p^* . They may only sell directly in the market if they pay a fixed cost I to operate a boat, which must be paid at the beginning of the period. Communities do not have access to external financing, so they can only operate a boat if their wealth w_t is greater than I . The number of units of fish that may be brought to market by a purchased boat in each period is q . For simplicity, we model the boat as only lasting a single period.

Communities have preferences over consumption c_t and assets transferred to next period, b_t .²⁰ Assets include both savings, which earns a constant interest rate r , and the value of boats, which may be used to

²⁰Modeling utility in each period as a function of consumption and savings follows [Banerjee and Newman \(1993b\)](#). Relative to modeling utility over an infinite sequence of consumption choices, this greatly simplifies the analysis and allows us to easily characterize dynamics.

transport fish. Their utility function takes the form:

$$u_t = (1 - s) \ln c_t + s \ln b_t.$$

The consumption good c_t is an aggregate of the quantity of fish consumed f_t and the quantity of other goods purchased using the proceeds from the sale of fish, o_t . The aggregation of these goods into the consumption good is given by:

$$c_t = (1 - \phi) f_t^\alpha + \phi o_t,$$

where $\alpha < 1$ indicates diminishing returns to fish.²¹ The extent to which communities value other goods relative to fish is captured by the parameter ϕ .

We denote the discrete choice to operate a boat as $\chi_t \in \{0, 1\}$. We choose o_t to be the numeraire good, so the community's problem is written as:

$$\max_{\{\chi_t, c_t, b_t, f_t, o_t\}} (1 - s) \ln c_t + s \ln b_t$$

subject to

$$o_t + b_t + p_t f_t \leq (1 + r)w_t + \bar{x} + \min\{q, Q\}(\chi_t p^* + (1 - \chi_t)p_t) + (Q - \min\{q, Q\})p_t - \chi_t rI \equiv y_t, \quad (3.3)$$

$$\chi_t I \leq w_t, \quad (3.4)$$

$$c_t = (1 - \phi) f_t^\alpha + \phi o_t, \quad (3.5)$$

$$\chi_t \in \{0, 1\}. \quad (3.6)$$

The first inequality is the budget constraint, which requires that expenditures not exceed their funds. Their funds come from wealth w_t , a lump sum payment the household receives each period $\bar{x}$, and income from fishing. If they decide to employ the boat, that means the assets used in the acquisition of the boat were not saved, so the return on those assets (equal to rI) are foregone by choosing to employ the boat.²² Therefore, the boat is only profitable if the interested earnings from the boat exceed the return on savings those funds.²³

The second inequality requires that investment in the boat must be financed from existing wealth. The third

²¹In Footnote 3, we noted that the fish consumed and fish sold are different species of fish, but since both are the output of fishing activities, we ignore that distinction here for simplicity.

²²In our notation, the timing is as follows. First, households enter the period with w units of assets and first decide whether or not to operate a boat using I units, and earn r on the $w - I$ remaining assets. Second, all production occurs and income from both assets and fishing are realized by households. Third, all purchases are made and consumption occurs.

²³For simplicity, we assume there is a market for boats, so a community can always buy or sell their boat for I at the beginning of each period.

constraint is the definition of the consumption good, and the fourth specifies that the decision to operate a boat is discrete. We denote available resources, which depend on endogenous decisions, wealth and prices, as y_t . Then the first-order conditions of this problem imply consumption and savings given by:

$$o_t + p_t f_t = (1 - s) y_t, \quad (3.7)$$

$$b_t = s y_t. \quad (3.8)$$

Moreover, as long as $o_t > 0$, the demand for fish is given by:

$$f_t = \left(\frac{(1 - \phi) \alpha}{\phi p_t} \right)^{1/(1-\alpha)},$$

which is conveniently independent of wealth.²⁴ Dynamics in the model follow from the fact that asset savings in period t become the assets in period $t + 1$:

$$w_{t+1} = b_t.$$

We assume that $s(1 + r) < 1$, so that the dynamic model has a steady-state wealth level.

3.4.2 Wealth Dynamics

We now characterize the dynamics of the wealth process and show that monopsonistic middlemen pricing gives rise to a poverty trap. If a community has sufficiently high initial wealth, they are able to operate a boat and receive greater earnings in each period. Those with lower initial wealth are unable to operate a boat, have lower earnings and save less. We provide conditions under which this gives rise to two different steady state wealth levels, and a cutoff in initial wealth that determines which steady state the community eventually realizes. Those communities below the initial wealth cutoff and transitioning toward the low steady state level of wealth are those experiencing the poverty trap.

Consider two cases for the relationship between the market price p^* and the price offered by the middlemen p_t , which are both taken as given by the community. Here we focus on the case where parameters are such that communities would like to sell more than one boat full of fish.²⁵

If $p_t \geq p^* - \frac{I}{q}$, purchasing and operating their own boat is not better for the community than selling to the

²⁴This is a result of our assumption of quasi-linear preferences. For low levels of income, $o_t = 0$ and $f_t = (1 - s)y_t/p_t$. This occurs when:

$$y_t < \frac{p_t^{-\alpha/(1-\alpha)}}{1 - s} \left(\frac{(1 - \phi) \alpha}{\phi} \right)^{1/(1-\alpha)}$$

²⁵The assumption allows for the community to always face the middlemen price of selling on the margin. Formally, here we

middlemen. This would arise, for example, if middlemen have a (weakly) better transportation technology than that available to the community. In this case, no boat is operated by the community ($\chi_t = 0$ and $y_t = w_t + \bar{x} + p_t Q$), so that $w_{t+1} = s((1+r)w_t + \bar{x} + p_t Q)$. In this case there is no poverty trap as there is only one steady state level of wealth, which is given by:

$$\bar{w}(p_t) = \frac{s(\bar{x} + p_t Q)}{1 - s(1+r)}. \quad (3.9)$$

If $p_t < p^* - \frac{I}{q}$, the boat offers greater net income to the community, so they will purchase it if their wealth is great enough. In this case, $y_t = w_t + (p^* - p_t)q - I + p_t Q$ if the community purchases the boat, and $y_t = w_t + p_t Q$ otherwise. The law of motion for wealth is therefore:

$$w_{t+1} = \begin{cases} s((1+r)w_t + (p^* - p_t)q - I + p_t Q) & \text{if } w_t > I \\ s((1+r)w_t + p_t Q) & \text{otherwise.} \end{cases}$$

The lower and upper branches of the law of motion lead to different possible steady states $\bar{w}_0(p_t)$ and $\bar{w}_1(p_t)$, with:

$$\bar{w}_0(p_t) = \frac{s(\bar{x} + p_t Q)}{1 - s(1+r)}, \quad (3.10)$$

$$\bar{w}_1(p_t) = \frac{s(\bar{x} + (p_t^* - p_t)q - I + p_t Q)}{1 - s(1+r)}. \quad (3.11)$$

Clearly, $\bar{w}_1(p_t) > \bar{w}_0(p_t)$. Which of these potential steady states is realized in the long run depends on I . If $I > \bar{w}_1(p_t) > \bar{w}_0(p_t)$, the boat will ultimately be unaffordable, since, for any initial wealth level, eventually $w_t < I$. The only steady state wealth level is $\bar{w}_0(p_t)$. Likewise, if $\bar{w}_1(p_t) > \bar{w}_0(p_t) > I$, the boat will be ultimately affordable, since eventually $w_t > I$ for any initial wealth level. In that case, the only steady state is $\bar{w}_1(p_t)$. In these cases, in the long run communities never operate boats or always operate boats, respectively.

There is a third possibility, which we are most interested in. If $\bar{w}_1(p_t) > I > \bar{w}_0(p_t)$ then poverty traps are possible. In this case, initial wealth determines which steady state is achieved. If a community has initial wealth of at least I they transition to the higher steady state level $\bar{w}_1(p_t)$. If their initial wealth is below I they transition to the lower wealth level $\bar{w}_0(p_t)$. The existence of these two separate steady states depends on the prices offered by middlemen being sufficiently low. We plot an example of this in the first panel of Figure 3.2. In the next subsection we show how market power by middlemen can depress the prices they offer.

are assuming that:

$$Q - q > \left(\frac{(1-\phi)\alpha}{\phi p_t} \right)^{1/(1-\alpha)}.$$

3.4.3 Middlemen

The previous section showed prices offered by middlemen are crucial for wealth dynamics. In order to understand the relationship between the boats and the prices offered by middlemen, we now consider the problem of the middlemen directly. We consider the case where the middlemen live for one period and operate as a cartel. The middlemen take the final market price p^* as given and use the same boat technology as the communities to bring the fish to market. Hence, they face a marginal cost of c to transport the fish. We assume that $c = rI/q$ so that both middlemen and communities have available the same transportation technology.

The cartel of middlemen acts as a monopsonist in purchasing fish from the community. That is, the cartel offers a take-it-or-leave-it price to the community, and the community then decides how much fish to sell to the middlemen at that price. The supply of fish that the middlemen face from a community S_χ depends on whether the community operates a boat or not. Community supply without a boat is:

$$S_0(p) = Q - \left(\frac{(1-\phi)\alpha}{\phi p} \right)^{1/(1-\alpha)}$$

The supply for a community with a boat is:

$$S_1(p) = Q - q - \left(\frac{(1-\phi)\alpha}{\phi p} \right)^{1/(1-\alpha)}$$

The profit maximization problem for a middleman:

$$\max_{p_t} S_\chi(p) (p^* - p - c)$$

which yields the following first-order condition:

$$\frac{\frac{1}{1-\alpha} \left(\frac{(1-\phi)\alpha}{\phi} \right)^{1/(1-\alpha)} p^{(\alpha-2)/(1-\alpha)}}{S_\chi(p)} = \frac{1}{(p^* - p - c)}.$$

The left-hand side (the percentage loss in supply from a marginal increase in price) is decreasing in the price, since both the exponent on price is negative and supply of fish in the denominator is increasing in price. In contrast, the right-hand side (the percentage increase in markup) is increasing in price, approaching infinity as the price approaches the competitive price of $p^* - c$. A set of appropriate parameter restrictions assures an interior solution satisfying our assumptions.²⁶

²⁶The right-hand side defines an upper bound on the middlemen price, since it goes to infinity as the middlemen price

The optimal middlemen price p_χ depends on whether or not the community has a boat because it depends on the supply function, $S_\chi(p)$. Notice that the supply of fish to middlemen is always smaller with the boat at any given price, $S_1(p) < S_0(p)$. Consequently, $p_1 > p_0$. Hence, an additional indirect benefit of the boat is a higher middlemen price, which leads to the following wealth dynamics:

$$w_{t+1} = \begin{cases} s((1+r)w_t + \bar{x} + (p^* - p_1)q - rI + p_1Q) & \text{if } w_t > I \\ s((1+r)w_t + \bar{x} + p_0Q) & \text{otherwise.} \end{cases} \quad (3.12)$$

The upper and lower branches of the law of motion again lead to two different steady states,

$\bar{w}_1 = \frac{s(\bar{x} + (p^* - p_1)q - rI + p_1Q)}{1 - s(1+r)} > \frac{s(\bar{x} + p_0Q)}{1 - s(1+r)} = \bar{w}_0$, but now the difference between them is greater than if they faced a common price. The dynamics of this poverty trap is illustrated in the second panel of Figure 3.2.

To understand the role that competition plays in this analysis, it is useful to compare this to the case where the cartel is eliminated and middlemen are competitive. In the competitive case, middlemen are only able to charge a price that allows them to break even. At a constant marginal cost rI/q , the competitive middlemen price would be $p = p^* - rI/q$ for all communities. As discussed in the previous subsection, at those prices the poverty trap is eliminated. Hence, the poverty trap in this model can only exist because of the market power of the middlemen.

3.4.4 Interventions to Escape Poverty Trap

Here we consider two types of interventions in the model that are motivated by the intervention considered in previous sections. The intervention implemented by FAS in Brazil had effects that can be interpreted in the model.

First, the FAS intervention can be interpreted in the model as implicit credit. By agreeing to participate in the boat program, communities agree to forego the Bolsa Floresta payments they usually receive from FAS for the period of time needed to pay for the boat (approximately three years). FAS is the sole holder of the source of repayment (the future payments), and, in any case, these remote communities would not have

approaches the competitive price, $p_t = p^* - c$. Recall that we assumed that the supply of fish was bounded below by $S(p_t) = Q - f_t > q$, so that some fish are sold to middlemen regardless of whether a boat is purchased. In order to ensure this, we need to set parameters to have a lower bound on prices that ensures that. The possible prices are therefore bounded above and below

$$\left(\frac{(1-\phi)\alpha}{\phi(Q-q)} \right) < p_t < p^* - c$$

We can find the parameters necessary to ensure this price, by plugging in for equality into the above optimality condition and setting the left hand side greater than the right-hand side. Manipulating gives us the following necessary assumption on parameters:

$$\frac{\alpha}{\alpha-1} \left(\frac{\phi}{(1-\phi)} \right) (Q-q)^{(2-\alpha)/(1-\alpha)} \left(p^* - c - \left(\frac{(1-\phi)\alpha}{\phi(Q-q)} \right) \right) > q.$$

Since the left-hand side is a product of positive terms and p_t^* can be arbitrarily large in principle, the existence of parameter values satisfying this assumption is assured.

been able to receive such a large loan from other sources. Just as cooperation allows communities to share the high cost of the boat across communities, credit allows individual communities to break the cost over time, thus eliminating the poverty trap.

This aspect is interpreted in the model as simply moving money from the time when income is realized to the time when it is needed for investment by allowing units to be shifted from $\bar{x}$ to w_t . Since w_t is required to finance the boat in equation (3.4), if enough funds can be moved forward in time this could eliminate the poverty trap entirely. This allows for communities initially in a poverty trap with $\bar{w}_0 < I$ to escape if $\bar{w}_0 + \bar{x} \geq I$. However, for an individual community it could be that their own income $\bar{x}$ over the duration of the loan is simply not large enough.

The second way to interpret the FAS intervention in the model is as encouraging *cooperation* among communities, which allows them to share a boat. Although a single community may lack the assets needed to acquire a boat (or return needed to justify it), it may be feasible (and profitable) for a group of communities to jointly purchase one. However, in the absence of means of committing to agreements among communities they may be unwilling to do so. Because FAS operates programs that help communities, including providing them with the Bolsa Floresta payments as part of their participation in their programs, FAS is able to enforce these agreements. Indeed, each group of communities that shares a boat under this program from FAS signs a contract regarding the use and maintenance of each boat, which is overseen by FAS. It is unclear how this type of agreement could be reached or enforced in the absence of this intervention.²⁷

In the model, we implement this aspect of the intervention by allowing the case where two communities can share a single boat (each using half of its capacity and bearing half the cost), and show that this can also break the poverty trap. We use the results from the previous subsection, and show that even partial use of the boat has a direct benefit of higher achieved prices on part of the fish harvest, and also an important indirect benefit of achieving higher prices from the middlemen. This helps them to achieve a higher steady state wealth level and may even cause them to achieve the highest steady state wealth level where each of the two communities is eventually wealthy enough to afford their own boat.

Suppose two communities both initially have wealth $\bar{w}_0$, the steady state level from the poverty trap in the previous case. Now suppose that $\bar{w}_0 > I/2$, and they now have the option of sharing the costs and benefits of operating a boat. That is, each spends $I/2$ to operate the boat each period and can use the boat to sell $q/2$ units of fish in the market. Now the cartel of middlemen faces a supply function from these communities of:

$$S_{0.5}(p) = Q - \frac{q}{2} - \left(\frac{(1-\phi)\alpha}{\phi p} \right)^{1/(1-\alpha)}.$$

²⁷In addition, FAS assists communities find customers for pirarucu sales through fairs.

As in the previous subsection, facing this supply function the cartel then solves for an optimal price $p_{0.5}$, where $p_0 < p_{0.5} < p_1$. Now a complete description of wealth dynamics is given by:

$$w_{t+1} = \begin{cases} s((1+r)w_t + \bar{x} + (p^* - p_1)q - I + p_1Q) & \text{if } w_t > I \\ s((1+r)w_t + \bar{x} + (p^* - p_{0.5})\frac{q}{2} - \frac{I}{2} + p_{0.5}Q) & \text{if } I > w_t > I/2 \\ s((1+r)w_t + \bar{x} + p_0Q) & \text{otherwise.} \end{cases} \quad (3.13)$$

We can now define a potential steady state based on the middle branch given by:

$$\bar{w}_{0.5} = \frac{s((p^* - p_{0.5})\frac{q}{2} - \frac{I}{2} + p_{0.5}Q + \bar{x})}{1 - s(1+r)}. \quad (3.14)$$

Clearly, this is an improvement on the initial wealth level for two reasons. First, $q/2$ units of fish are sold at market prices, and $p^* > p_0$. Second, middlemen respond to the change in supply that the boat causes so that $p_{0.5} > p_0$, even for the fish not shipped on the boats. Therefore, this in itself is an improvement that would not be available in the absence of cooperation. Moreover, if $\bar{w}_{0.5} > I$ then in a finite number of periods, wealth exceeds I and each community can operate its own boat. This means that the steady state achieved by the communities will not be $\bar{w}_{0.5}$ but $\bar{w}_1$, the highest steady state wealth achieved in the previous case. Again, because of both of the direct gains (better price at the market for shipped goods) and indirect gains (better middlemen prices), there is a range of parameter values such that $\bar{w}_0 < I < \bar{w}_{0.5}$. This implies that the cooperative strategy alone could help an otherwise poor community escape the poverty trap entirely. Moreover, for parameter values such that I is only slightly less than $\bar{w}_{0.5}$, the indirect pricing *response* of the middlemen to the boat may be crucial to escaping the trap. This is illustrated in the third panel of Figure 3.2.

Taken together, these interventions allow for communities initially in a poverty trap with $\bar{w}_0 < I$ to overcome it either by sharing the cost of the boat with cooperation, or by increasing available assets with credit. Given the high cost of these boats and relatively small community sizes, we believe that both aspects played an important role.

3.4.5 Is there a poverty trap?

Further mapping the model to the Amazonas environment, a poverty trap could exist if the cost of a boat were prohibitively high in a single year for a single community. Although in the model two communities share a boat, we have already noted that in reality the boats are shared by 8-10 communities. So, what if one community attempted to purchase it, without credit?

From Table 3.2, the average income per household is about R\$670 per household per month. With an average of 18 households per community, the overall annual income of a river community is approximately R\$145,000. The average cost of a sufficiently large boat, with the necessary hold constructed to carry ice, is approximately R\$150,000. Without access to credit, a community would need to dedicate more than a full year's income for a single year to a boat purchase, essentially implying that the community would not consume food, gas, or diesel during that time. The high level of consumption relative to income in Table 3.2 implies that communities live close to subsistence, further underscoring the implausibility of such a purchase.

Would the additional income in a single year make this possible? Table 3.3 shows the effects on market pricing, including an increase from R\$4/kilogram to R\$4.77 for those communities with access to boats, nearly a 20% price improvement. If we assume that communities were somehow able to access this higher price, that is, that they were not subject to the lower price implied by the middlemen's market power, then average community income would increase by R\$15,500 ($1,119 \text{ kg/household} \times 18 \text{ households/community} \times \text{R\$0.77/kg}$) to R\$160,000. The river community would be closer to the purchase of a boat, but it would still not be financially possible at current fishing levels. Without a huge expansion in fishing, which is limited by the conservation needs of preserving the pirarucu numbers, one would need a community many times larger than the average to enable such an investment. The groups that purchased boats consisted of 8-10 communities — implying 144-180 households on average. The largest community in our sample is just 46 households, much smaller than this.

3.5 Conclusion

We have presented results in which an NGO-initiated and community-financed boat purchase enables households to attain higher prices on harvested fish and subsequently increase their incomes. Using variation that came from the inability to supply boats to all communities simultaneously, we identified the causal impact of the intervention. The intervention led to a substantial increase in income for these poor communities and directly impacted welfare through an increase in consumption. Our strongest results showed that these benefits came from higher prices received for their fish from the end market, but we also showed results of effects from the upward response of prices from middlemen. We find that the boats are cost-effective, paying for themselves within 3 years. Working alone, however, the cost of the boat would be prohibitive for any individual community. We presented a model that showed how both higher market prices and the responses of middlemen could be pivotal in enabling communities to escape man-made poverty traps that the lack of wealth for boats otherwise constitutes.

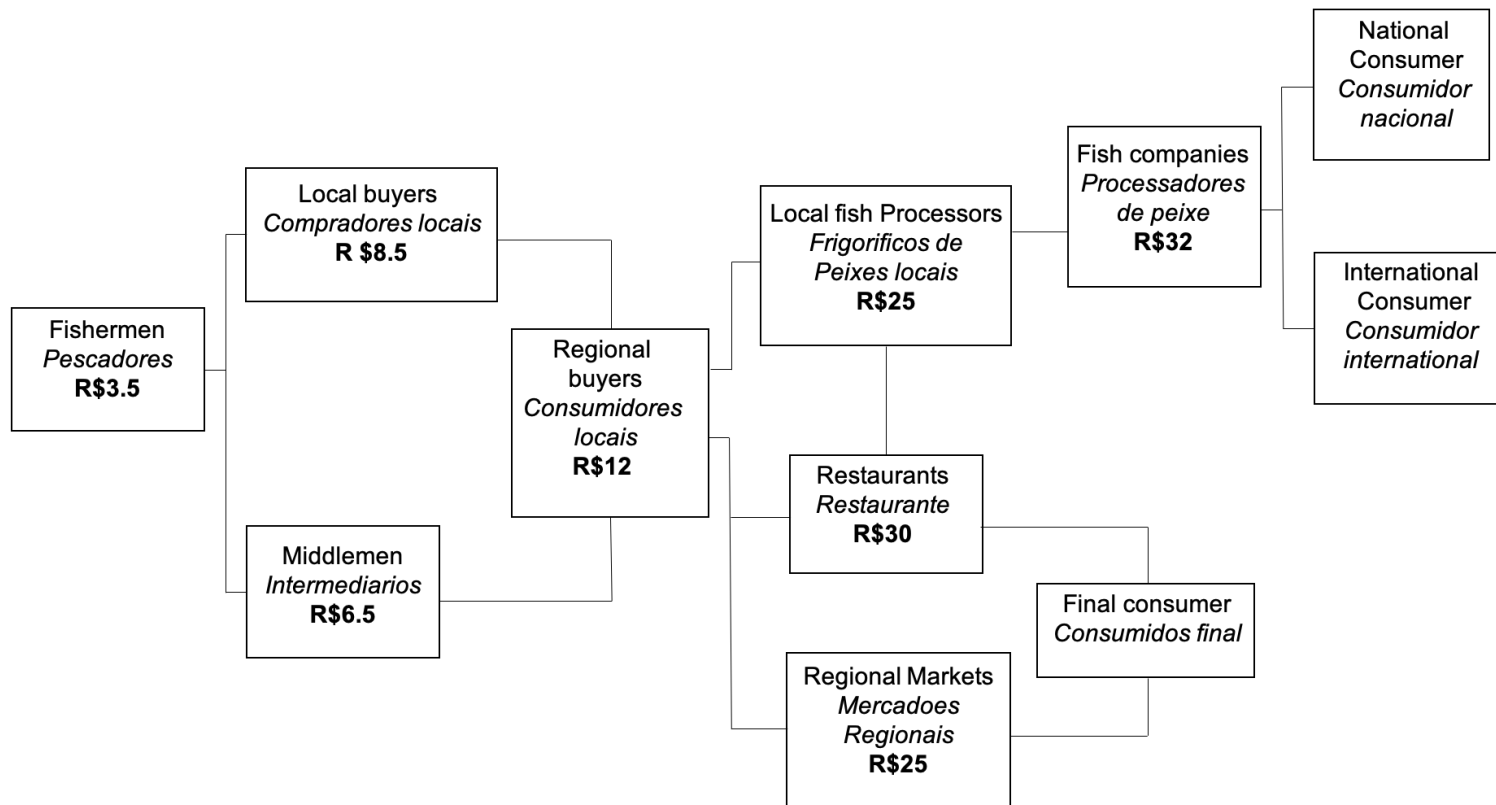
The results are important. Typical approaches for reaching remote communities (large scale infras-

tructure) are expensive, so alternatives such as boats and bridges ([Brooks and Donovan, 2020](#)) may be an important tool for policymakers. This is especially true in remote areas of developing countries where it is difficult to implement larger scale projects.

Our paper suggests several avenues for future research. First, in our setting we plan to conduct a long-term follow-up to evaluate how the use of the boats changes over time, how additional moneys are spent and invested, and how increased integration affects community engagement with outside areas, such as external labor markets and educational opportunities. Second, the results suggest that cooperative investments were profitable. This may indicate that other investments for reaching communities may also be profitable, and more research is needed to understand why investments may not be made. Further investigating the collusive nature of supply chains, such as the one we examined, is a first step towards designing the policies that will improve the welfare of remote communities.

3.6 Figures

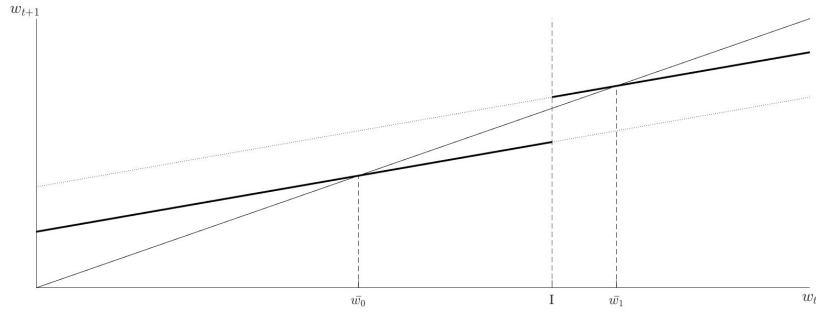
Figure 3.1: Pirarucu Supply Chain and Market (Reais/kg) in Mamiraua, Amazonas, 2018



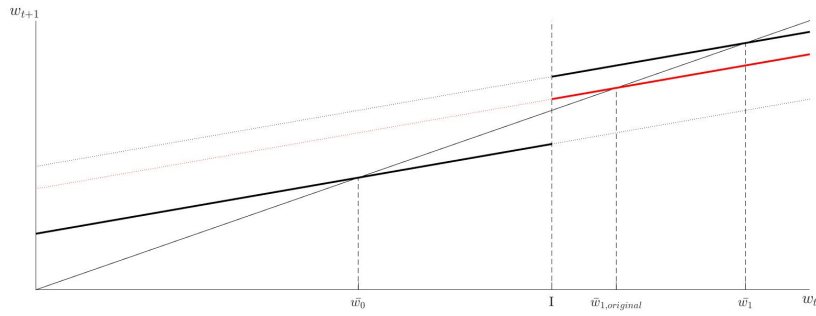
Source: Mamiraua Institute with own analysis

Figure 3.2: Wealth Dynamics Under Monopsonistic Middlemen Pricing

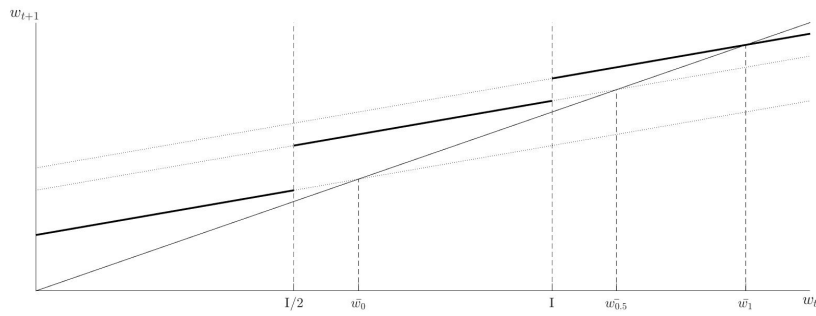
Panel 1: Poverty trap at low middleman price



Panel 2: Boat enabling higher middleman price and higher steady state wealth



Panel 3: Cooperative investment eliminating poverty trap



3.7 Tables

Table 3.1: Balance on Baseline Community Characteristics

	Boat in 2018	N	No boat in 2018	N	Diff	(s.e.)	P-value
<i>Panel A: Collected from NGO's operational records</i>							
num households	18	26	17	18	1	(3.6)	0.78
eligible for Bolsa Floresta = 1	.81	26	.78	18	.03	(.13)	0.81
hrs to closest market	11	26	9.6	18	1.8	(4.8)	0.71
<i>Panel B: Collected retrospectively from fishing leaders</i>							
num pirarucu 2017	229	18	65	10	164**	(78)	0.05
<i>Panel C: Collected from NGO's community survey</i>							
has school = 1	.95	21	.82	17	.13	(.1)	0.21
num teachers per hh	.11	21	.26	17	-.16	(.11)	0.18
has health agents = 1	.57	21	.53	17	.042	(.17)	0.80
num health agents per hh	.035	21	.092	17	-.056	(.054)	0.30
has ambulance = 1	.19	21	.18	17	.014	(.13)	0.91
has water treatment = 1	.75	20	.82	17	-.074	(.14)	0.60

Notes: Standard errors estimated in a t test on the equality of means between the communities receiving a boat in 2018 and the communities without a boat in 2018. * $p < 0.1$, ** $p < .05$, *** $p < 0.01$

Table 3.2: Effects on Households

	Components of consumption:						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	income	consum	food	gas	diesel	transport	distant city travel = 1
FAS boat 2018	184**	148	68**	42*	-6.5	38	.16**
	(72)	(92)	(32)	(22)	(14)	(28)	(.069)
	[.22]	[.32]	[.19]	[.28]	[.85]	[.22]	[.38]
Control Mean	670	892	435	200	42	152	.47
R ²	.044	.17	.14	.37	.12	.092	.27
Observations	308	308	308	308	308	308	308

Notes: In parentheses, we report standard errors clustered at the community level. In square brackets, we report randomization inference p-values based on 10 community groups. Outcomes (1)-(6) are monthly in 2018 Brazilian Reais, R\$. Consumption is the sum of food, gas, diesel, and transport spending. Outliers winsorized to the 5th and 95th percentile. All specifications include reserve fixed effects and the following controls: the number of pirarucu harvested in 2017, household size, time to market and its square, the interaction of time to market with reserve, and the interaction of time to market squared with reserve. Missing values in control variables set to mean, and missing value flag included as covariate. Asterisks denote statistical significance under clustered standard errors: ***1%, **5%, *10%.

Table 3.3: Effects on Market Pricing and Channel (Community Level)

	(1) R\$/kg all	(2) R\$/kg mid	(3) num middlemen
FAS boat 2018	.77** (.34) [.28]	.9 (.56) [.24]	-2.1 (2.8) [.43]
Control Mean	4	3.6	5.6
R ²	.38	.47	.12
Observations	41	32	43

Notes: In parentheses, we report heteroskedasticity robust standard errors. In square brackets, we report randomization inference p-values based on 10 community groups. **All** reflects annual sales to any buyer. **Mid** reflects annual sales solely to middlemen. Sample in columns (1) and (2) consists of those communities reporting positive quantity sold to any buyer (1) or to middlemen (2). All specifications include reserve fixed effects and the following controls: the number of pirarucu harvested in 2017, the number of households residing in the community, time to market and its square, the interaction of time to market with reserve, and the interaction of time to market squared with reserve. Columns (1)-(2) additionally including the pirarucu price in 2017 as a control. Missing values in control variables set to mean, and missing value flag included as covariate. Asterisks denote statistical significance under robust standard errors: ***1%, **5%, *10%.

Table 3.4: Effects on Quantity Sold and Revenue Earned by Communities

	(1) kg/hh all	(2) kg/hh mid	(3) R\$/hh all	(4) R\$/hh mid
FAS boat 2018	45 (275) [.86]	204 (307) [.4]	1,486 (1,491) [.55]	1,734 (1,531) [.22]
Control Mean	1,119	622	4,314	2,145
R ²	.32	.29	.25	.22
Observations	41	41	41	41

Notes: In parentheses, we report heteroskedasticity robust standard errors. In square brackets, we report randomization inference p-values based on 10 community groups. **All** reflects annual sales to any buyer. **Mid** reflects annual sales solely to middlemen. All specifications include reserve fixed effects and the following controls: the number of pirarucu harvested in 2017, time to market and its square, the interaction of time to market with reserve, and the interaction of time to market squared with reserve. Missing values in control variables set to mean, and missing value flag included as covariate. Asterisks denote statistical significance under robust standard errors: ***1%, **5%, *10%.

Appendix

Appendix 3.A Main Appendix

Table 3.A.1: Sailing Boats from Mamirauá to Market
Variable Transportation Costs

	Unit Cost (R\$)	Unit	Quantity	Cost (R\$)
Fuel	5	Per liter	4,000	20,000
Ice	0.30	Per kilogram	30,000	9,000
Meals	20	Per day	12	240
Deck hand labor	254.50	Per person	2	509
Captain's labor	509	Per person	1	509
Boat depreciation	556	Per trip	1	556
Total cost per trip:				R\$30,814
Average kg of fish transported per trip :				30,000 kg
Average variable transportation cost per kg of fish:				R\$1.03/kg

Source: Qualitative interviews conducted by the University of Notre Dame's Mendoza College of Business with middlemen.

Table 3.A.2: Summary Statistics for the 2018 Household Survey

	mean	sd
<i>Characteristics fixed prior to treatment, full sample (N = 308)</i>		
female = 1	.35	.48
age	39	13
hh size	5.4	2.5
born in community = 1	.42	.49
completed primary school = 1	.31	.46
hh receives Bolsa Floresta = 1	.67	.47
fisherman = 1	.85	.36
farmer = 1	.78	.42
livestock herder = 1	.16	.36
<i>Potentially time-variant characteristics, control group (N=114)</i>		
monthly income, R\$	670	482
monthly food spending, R\$	435	251
monthly transport spending, R\$	152	164
owns radio = 1	.35	.48
owns cell phone = 1	.54	.5
fishing hours/week	29	25
farming hours/week	18	15
annual trips to nearby city	17	15
travels to a distant city = 1	.47	.5

Table 3.A.3: Testing for Differential Pre-Trends in Fishing Levels
2013-2017, Mamirauá

	(1) ln fish	(2) ln fish
year*boat	.034 (.82)	.034 (.84)
year	.014 (.73)	
Year FE's	No	Yes
R ²	.032	.079
Observations	18	18

Notes: In parentheses, we report wild cluster bootstrap p-values, given the small number of sectors composing the sample. The unit of observation is the sector-year. **Year** denotes a yearly time trend. All specifications include sector FE's. The outcome is the natural log of the sector's total fish harvest in a given year.

Table 3.A.4: Effects on Household Time Allocation

	(1) fishing hrs/week	(2) farming hrs/week
FAS boat 2018	-.6 (4.1) [.92]	-1.8 (2.3) [.23]
Control Mean	29	18
R ²	.041	.1
Observations	308	308

Notes: In parentheses, we report standard errors clustered at the community level. In square brackets, we report randomization inference p-values based on 10 community groups. All specifications include reserve fixed effects and the following controls: the number of pirarucu harvested in 2017, household size, time to market and its square, the interaction of time to market with reserve, and the interaction of time to market squared with reserve. Missing values in control variables set to mean, and missing value flag included as covariate. Asterisks denote statistical significance under clustered standard errors: *** 1%, ** 5%, * 10%.

Appendix 3.B Data Appendix

Our empirical analysis utilizes a combination of several primary and secondary data sources:

Primary data sources:

1. A household survey conducted between October and November 2018, following the 2018 pirarucu harvest, that collects household-level information on demographics, finances, labor allocation, and travel.
2. A key informant survey with the fishing leader in each community, collected at the same time as the household survey, which documents community-level pirarucu quantity, revenue, and access to intermediaries during the 2018 fishing season.
3. Phone calls to fishing leaders to retrospectively collect information about pirarucu quantity and revenue during the 2017 fishing season.

Secondary data sources:

4. Operational records from our NGO partner, FAS, to collect community-level information on the size, remoteness, and Bolsa Floresta eligibility of the communities in our sample at the time of the intervention. FAS operational records also allow us to attribute any income earned from the 2019 Holy Week pirarucu fairs, which sold pirarucu that were harvested during the 2018 season, back to the respective communities that contributed fish to be sold at the fair.
5. A community-level survey collected by FAS prior to 2018, which surveyed the leader of most communities in our sample about access to education and health infrastructure.
6. Historical records from the Mamirauá Institute and ASPROC, environmental research organizations in the Amazon, on fishing levels leading up to 2018. The Mamirauá Institute and ASPROC document this information at a sector-level, i.e., at the level of the associations formed between communities to purchase boats.

Our qualitative analysis and description of the setting in Section 3.2 relies on qualitative interviews conducted by the University of Notre Dame's Mendoza College of Business with fishermen, municipal leaders, pirarucu customers, middlemen, frigorificos, and staff at our NGO partner, Foundation for Amazonas Sustainability, across Manaus and the Mamiraua and Uacari reserves.

Bibliography

- Abadie, Alberto. 2002. “Bootstrap tests for distributional treatment effects in instrumental variable models.” *Journal of the American Statistical Association* 97 (457):284–292.
- Abebe, Girum, A Stefano Caria, Marcel Fafchamps, Paolo Falco, Simon Franklin, and Simon Quinn. 2021. “Anonymity or distance? Job search and labour market exclusion in a growing African city.” *The Review of Economic Studies* 88 (3):1279–1310.
- Acampora, Michelle, Lorenzo Casaburi, and Jack Willis. 2022. “Land rental markets: Experimental evidence from Kenya.” Tech. rep., National Bureau of Economic Research.
- Aghion, Philippe and Patrick Bolton. 1997. “A Theory of Trickle-Down Growth and Development.” *Review of Economic Studies* 64 (2):151–172.
- Ahlfeldt, Gabriel M, Stephen J Redding, Daniel M Sturm, and Nikolaus Wolf. 2015. “The economics of density: Evidence from the Berlin Wall.” *Econometrica* 83 (6):2127–2189.
- Aker, Jenny C. 2010. “Information from markets near and far: Mobile phones and agricultural markets in Niger.” *American Economic Journal: Applied Economics* 2 (3):46–59.
- Alder, Simon and Illenin Kondo. 2020. “Political Distortions and Infrastructure Networks in China: A Quantitative Spatial Equilibrium Analysis.” *Working Paper* .
- Anderson, Michael L. 2008. “Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects.” *Journal of the American statistical Association* 103 (484):1481–1495.
- Angelucci, Manuela, Dean Karlan, and Jonathan Zinman. 2015. “Microcredit impacts: Evidence from a randomized microcredit program placement experiment by Compartamos Banco.” *American Economic Journal: Applied Economics* 7 (1):151–182.
- Angrist, Joshua, Peter Hull, and Christopher Walters. 2023. “Methods for measuring school effectiveness.” *Handbook of the Economics of Education* 7:1–60.
- Asher, Sam and Paul Novosad. 2020. “Rural Roads and Local Economic Development.” *American Economic Review* 110 (3):797–823.
- Assunção, Juliano, Molly Lipscomb, Ahmed Mushfiq Mobarak, and Dmitri Szerman. 2017. “Agricultural Productivity and Deforestation in Brazil.” Tech. rep. Working Paper.
- Asturias, Jose, Manuel Garcia-Santana, and Roberto Ramos. 2019. “Competition and the Welfare Gains from Transportation Infrastructure: Evidence from the Golden Quadrilateral in India.” *Journal of the European Economic Association* 17 (6):1881–1940.
- Attanasio, Orazio, Britta Augsburg, Ralph De Haas, Emla Fitzsimons, and Heike Harmgart. 2015. “The impacts of microfinance: Evidence from joint-liability lending in Mongolia.” *American Economic Journal: Applied Economics* 7 (1):90–122.

- Balboni, Clare, Oriana Bandiera, Robin Burgess, Maitreesh Ghatak, and Anton Heil. 2022. "Why do people stay poor?" *The Quarterly Journal of Economics* 137 (2):785–844.
- Balboni, Clare, Gharad Bryan, Melanie Morten, and Bilal Siddiqi. 2020. "Transportation, Gentrification, and Urban Mobility: The Inequality Effects of Place-Based Policies." Tech. rep.
- Banerjee, Abhijit, Emily Breza, Esther Duflo, and Cynthia Kinnan. 2024. "Can microfinance unlock a poverty trap for some entrepreneurs?" Tech. rep., National Bureau of Economic Research.
- Banerjee, Abhijit, Esther Duflo, Rachel Glennerster, and Cynthia Kinnan. 2015. "The miracle of micro-finance? Evidence from a randomized evaluation." *American Economic Journal: Applied Economics* 7 (1):22–53.
- Banerjee, Abhijit, Michael Faye, Alan Krueger, Paul Niehaus, and Tavneet Suri. 2023. "Universal Basic Income: Short-Term Results from a Long-Term Experiment in Kenya." Tech. rep., UC San Diego.
- Banerjee, Abhijit V and Andrew F Newman. 1993a. "Occupational choice and the process of development." *Journal of Political Economy* 101 (2):274–298.
- Banerjee, Abhijit V. and Andrew F. Newman. 1993b. "Occupational Choice and the Process of Development." *Journal of Political Economy* 101 (2):274–298.
- Barrett, Christopher B., Michael R. Carter, and Jean-Paul Jean Paul Chavas. 2019. *The Economics of Poverty Traps*. Cambridge, MA: National Bureau of Economic Research, Inc.
- Barrett, Christopher B, Paswel Phiri Marenja, John McPeak, Bart Minten, Festus Murithi, Willis Oluoch-Kosura, Frank Place, Jean Claude Randrianarisoa, Jhon Rasambainarivo, and Justine Wangila. 2006. "Welfare dynamics in rural Kenya and Madagascar." *The Journal of Development Studies* 42 (2):248–277.
- Baseler, Travis. 2023. "Hidden income and the perceived returns to migration." *American Economic Journal: Applied Economics* 15 (4):321–352.
- Bassi, Vittorio, Matthew E Kahn, Nancy Lozano Gracia, Tommaso Porzio, and Jeanne Sorin. 2024. "Jobs in the Smog: Firm Location and Workers' Exposure to Pollution in African Cities." Tech. rep., National Bureau of Economic Research.
- Bassi, Vittorio, Raffaella Muoio, Tommaso Porzio, Ritwika Sen, and Esau Tugume. 2022. "Achieving scale collectively." *Econometrica* 90 (6):2937–2978.
- Beaman, Lori, Dean Karlan, Bram Thuysbaert, and Christopher Udry. 2023. "Selection into credit markets: Evidence from agriculture in Mali." *Econometrica* 91 (5):1595–1627.
- Bergquist, Lauren, Benjamin Faber, Thibault Fally, Matthias Hoelzlein, Edward Miguel, and Andres Rodriguez-Clare. 2019. "Scaling Agricultural Policy Interventions: Theory and Evidence from Uganda." Unpublished manuscript, University of California at Berkeley.
- Bergquist, Lauren Falcao and Michael Dinerstein. 2020. "Competition and Entry in Agricultural Markets: Experimental Evidence from Kenya." *American Economic Review* 110 (12):3705–3747.
- Blattman, Christopher, Nathan Fiala, and Sebastian Martinez. 2014. "Generating skilled self-employment in developing countries: Experimental evidence from Uganda." *The Quarterly Journal of Economics* 129 (2):697–752.
- Blattman, Christopher, Eric P Green, Julian Jamison, M Christian Lehmann, and Jeannie Annan. 2016. "The returns to microenterprise support among the ultrapoor: A field experiment in postwar Uganda." *American economic journal: Applied economics* 8 (2):35–64.
- Borjas, George J. 1987. "Self-selection and the earnings of immigrants." *American Economic Review* 77 (4):531–553.

- Brooks, Wyatt and Kevin Donovan. 2020. "Eliminating uncertainty in market access: The impact of new bridges in rural Nicaragua." *Econometrica* 88 (5):1965–1997.
- Bruhn, Miriam and David McKenzie. 2009. "In Pursuit of Balance: Randomization in Practice in Development Field Experiments." *American Economic Journal: Applied Economics* 1 (4):200–232. URL <http://www.jstor.org/stable/25760187>.
- Bryan, Gharad, Shyamal Chowdhury, and Ahmed Mushfiq Mobarak. 2014. "Underinvestment in a profitable technology: The case of seasonal migration in Bangladesh." *Econometrica* 82 (5):1671–1748.
- Bryan, Gharad, Jonathan de Quidt, Mariajose Silva-Vargas, Tom Wilkening, and Nitin Yadav. 2022. "Market Design for Land Trade: Evidence from Uganda and Kenya." *London School of Economics, London*.
- Bryan, Gharad and Melanie Morten. 2019. "The aggregate productivity effects of internal migration: Evidence from Indonesia." *Journal of Political Economy* 127 (5):2229–2268.
- Buera, Francisco J. 2009. "A dynamic model of entrepreneurship with borrowing constraints: theory and evidence." *Annals of Finance* 5 (3):443–464.
- Buera, Francisco J, Joseph P Kaboski, and Yongseok Shin. 2014. "Macro-perspective on asset grants programs: Occupational and wealth mobility." *American Economic Review* 104 (5):159–64.
- . 2015. "Entrepreneurship and financial frictions: A macrodevelopment perspective." *Annual Reviews of Economics* 7 (1):409–436.
- . 2020. "The Macroeconomics of Microfinance." *The Review of Economic Studies* 88 (1):126–161. URL <https://doi.org/10.1093/restud/rdaa047>.
- Buera, Francisco J and Yongseok Shin. 2013. "Financial frictions and the persistence of history: A quantitative exploration." *Journal of Political Economy* 121 (2):221–272.
- Burgess, Robin, Remi Jedwab, Edward Miguel, Ameet Morjaria, and Gerard Padró i Miquel. 2015. "The value of democracy: evidence from road building in Kenya." *American Economic Review* 105 (6):1817–51.
- Burke, Marshall, Lauren Falcao Bergquist, and Edward Miguel. 2019. "Sell Low and Buy High: Arbitrage and Local Price Effects in Kenyan Markets." *The Quarterly Journal of Economics* 134 (2):785–842. URL <https://ideas.repec.org/a/oup/qjecon/v134y2019i2p785-842..html>.
- Card, David, Jesse Rothstein, and Moises Yi. 2023. "Location, location, location." Tech. rep., National Bureau of Economic Research.
- Carter, Michael R. and Travis J. Lybbert. 2012. "Consumption versus asset smoothing: testing the implications of poverty trap theory in Burkina Faso." *Journal of Development Economics* 99 (2):255–264. URL <https://www.sciencedirect.com/science/article/pii/S0304387812000077>.
- Casaburi, Lorenzo and Rocco Macchiavello. 2019. "Demand and supply of infrequent payments as a commitment device: evidence from Kenya." *American Economic Review* 109 (2):523–555.
- Caunedo, Julieta and Namrata Kala. 2021. "Mechanizing Agriculture." Working Paper 29061, National Bureau of Economic Research. URL <http://www.nber.org/papers/w29061>.
- Chernozhukov, Victor, Iván Fernández-Val, and Blaise Melly. 2013. "Inference on counterfactual distributions." *Econometrica* 81 (6):2205–2268.
- Chetty, Raj and Adam Szeidl. 2007. "Consumption commitments and risk preferences." *The Quarterly Journal of Economics* 122 (2):831–877.
- De Mel, Suresh, David McKenzie, and Christopher Woodruff. 2008. "Returns to capital in microenterprises: Evidence from a field experiment." *The Quarterly Journal of Economics* 123 (4):1329–1372.

- . 2014. “Business training and female enterprise start-up, growth, and dynamics: Experimental evidence from Sri Lanka.” *Journal of Development Economics* 106:199–210.
- Dercon, Stefan. 1998. “Wealth, risk and activity choice: cattle in Western Tanzania.” *Journal of Development Economics* 55 (1):1–42.
- Donaldson, Dave. 2018. “Railroads of the Raj: Estimating the impact of transportation infrastructure.” *American Economic Review* 108 (4-5):899–934.
- Donovan, Kevin. 2021. “The equilibrium impact of agricultural risk on intermediate inputs and aggregate productivity.” *The Review of Economic Studies* 88 (5):2275–2307.
- Doornik, Bernardus Ferdinandus Nazar Van, Armando R Gomes, David Schoenherr, and Janis Skrastins. 2021. “Financial Access and Labor Market Outcomes: Evidence from Credit Lotteries.” Tech. rep.
- Eaton, Jonathan and Samuel Kortum. 2002. “Technology, geography, and trade.” *Econometrica* 70 (5):1741–1779.
- Egger, Dennis, Johannes Haushofer, Edward Miguel, Paul Niehaus, and Michael W Walker. 2019. “General equilibrium effects of cash transfers: experimental evidence from Kenya.” Tech. rep., National Bureau of Economic Research.
- Foster, Andrew D and Mark R Rosenzweig. 2022. “Are there too many farms in the world? labor market transaction costs, machine capacities, and optimal farm size.” *Journal of Political Economy* 130 (3):636–680.
- Franklin, Simon, Clement Imbert, Girum Abebe, and Carolina Mejia-Mantilla. 2024. “Urban public works in spatial equilibrium: Experimental evidence from Ethiopia.” *American Economic Review* 114 (5):1382–1414.
- Galor, Oded and Joseph Zeira. 1993. “Income distribution and macroeconomics.” *The Review of Economic Studies* 60 (1):35–52.
- Gertler, Paul, Sean Higgins, Aisling Scott, and Enrique Seira. 2023. “Using Lotteries to Attract Deposits.” Tech. rep., National Bureau of Economic Research.
- Ghatak, Maitreesh and Nien-Huei Jiang. 2002. “A simple model of inequality, occupational choice, and development.” *Journal of Development Economics* 69 (1):205–226.
- Gochberg, William. 2021. “The social costs of titling land: Evidence from Uganda.” *World Development* 142:105376. URL <https://www.sciencedirect.com/science/article/pii/S0305750X20305040>.
- Gollin, Douglas, David Lagakos, and Michael E Waugh. 2014. “The agricultural productivity gap.” *The Quarterly Journal of Economics* 129 (2):939–993.
- Gollin, Douglas and Christopher Udry. 2021. “Heterogeneity, Measurement Error, and Misallocation: Evidence from African Agriculture.” *Journal of Political Economy* 129 (1):1–80. URL <https://doi.org/10.1086/711369>.
- Goodman-Bacon, Andrew. 2021. “The long-run effects of childhood insurance coverage: Medicaid implementation, adult health, and labor market outcomes.” *American Economic Review* 111 (8):2550–2593.
- Grant, Matthew and Meredith Startz. 2021. “Cutting out the middleman: The structure of chains of intermediation.” Working paper.
- Greaney, Brian P, Joseph P Kaboski, and Eva Van Leemput. 2016. “Can self-help groups really be “self-help”?” *The Review of Economic Studies* 83 (4):1614–1644.
- Haushofer, Johannes and Jeremy Shapiro. 2016. “The short-term impact of unconditional cash transfers to the poor: experimental evidence from Kenya.” *The Quarterly Journal of Economics* 131 (4):1973–2042.

- . 2018. “The long-term impact of unconditional cash transfers: experimental evidence from Kenya.” Tech. rep., Busara Center for Behavioral Economics, Nairobi, Kenya.
- Herskowitz, Sylvan. 2021. “Gambling, saving, and lumpy liquidity needs.” *American Economic Journal: Applied Economics* 13 (1):72–104.
- Holt, Charles A and Susan K Laury. 2002. “Risk aversion and incentive effects.” *American Economic Review* 92 (5):1644–1655.
- Jayachandran, Seema. 2006. “Selling labor low: Wage responses to productivity shocks in developing countries.” *Journal of Political Economy* 114 (3):538–575.
- Jedwab, Remi and Adam Storeygard. 2020. “The average and heterogeneous effects of transportation investments: Evidence from Sub-Saharan Africa 1960-2010.” Tech. rep., National Bureau of Economic Research.
- Jensen, Robert. 2007. “The Digital Divide: Information (Technology), Market Performance, and Welfare in the South Indian Fisheries Sector.” *Quarterly Journal of Economics* 122 (3):879–924.
- Jensen, Robert and Nolan Miller. 2018. “Market Integration, Demand, and the Growth of Firms: Evidence from a Natural Experiment in India.” *American Economic Review* 108 (12):3583–3625.
- Kaboski, Joseph P and Robert M Townsend. 2011. “A structural evaluation of a large-scale quasi-experimental microfinance initiative.” *Econometrica* 79 (5):1357–1406.
- Kaplan, Greg and Giovanni L. Violante. 2014. “A Model of the Consumption Response to Fiscal Stimulus Payments.” *Econometrica* 82 (4):1199–1239. URL <https://ideas.repec.org/a/wly/emetrp/v82y2014i4p1199-1239.html>.
- Kaplan, Greg, Giovanni L. Violante, and Justin Weidner. 2014. “The Wealthy Hand-to-Mouth.” *Brookings Papers on Economic Activity* 45 (1 (Spring)):77–153. URL <https://ideas.repec.org/a/bin/bpeajo/v45y2014i2014-01p77-153.html>.
- Kwang, Ng Yew. 1965. “Why do people buy lottery tickets? Choices involving risk and the indivisibility of expenditure.” *Journal of Political Economy* 73 (5):530–535.
- LaFave, Daniel and Duncan Thomas. 2016. “Farms, families, and markets: New evidence on completeness of markets in agricultural settings.” *Econometrica* 84 (5):1917–1960.
- Lagakos, David, Mushfiq Mobarak, and Michael Waugh. 2018. “The Welfare Effects of Encouraging Rural-Urban Migration.” Working Papers 2018–002, Human Capital and Economic Opportunity Working Group.
- Larcom, Shaun, Ferdinand Rauch, and Tim Willems. 2017. “The benefits of forced experimentation: Striking evidence from the London underground network.” *The Quarterly Journal of Economics* 132 (4):2019–2055.
- Lee, David S. 2009. “Training, wages, and sample selection: Estimating sharp bounds on treatment effects.” *The Review of Economic Studies* 76 (3):1071–1102.
- Lu, Will Jianyu. 2020. “Transport, Infrastructure and Growth: Evidence from Chinese Firms.” *Working Paper*.
- Lybbert, Travis J. and Christopher B. Barrett. 2011. “Risk-Taking Behavior in the Presence of Nonconvex Asset Dynamics.” *Economic Inquiry* 49 (4):982–988. URL <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1465-7295.2009.00198.x>.
- Lybbert, Travis J., Christopher B. Barrett, Solomon Desta, and D. Layne Coppock. 2004. “Stochastic Wealth Dynamics and Risk Management among a Poor Population.” *The Economic Journal* 114 (498):750–777. URL <http://www.jstor.org/stable/3590237>.

- McKenzie, David J and Christopher Woodruff. 2006. "Do Entry Costs Provide an Empirical Basis for Poverty Traps? Evidence from Mexican Microenterprises." *Economic Development and Cultural Change* 55 (1):3–42. URL <https://ideas.repec.org/a/ucp/ecdecc/y2006v55i1p3-42.html>.
- Meissner, Thomas, Xavier Gassmann, Corinne Faure, and Joachim Schleich. 2023. "Individual characteristics associated with risk and time preferences: A multi country representative survey." *Journal of Risk and Uncertainty* 66 (1):77–107.
- Midrigan, Virgiliu and Daniel Yi Xu. 2014. "Finance and misallocation: Evidence from plant-level data." *American Economic Review* 104 (2):422–58.
- Monte, Ferdinando, Stephen J Redding, and Esteban Rossi-Hansberg. 2018. "Commuting, migration, and local employment elasticities." *American Economic Review* 108 (12):3855–3890.
- Murphy, Kevin M., Andrei Shleifer, and Robert W. Vishny. 1989. "Industrialization and the Big Push." *Journal of Political Economy* 97 (5):1003–1026. URL <http://www.jstor.org/stable/1831884>.
- Pfaff, A. 1999. "What Drives Deforestation in the Brazilian Amazon? Evidence from Satellite and Socioeconomic Data." *Journal of Environmental Economics and Management* 37 (1).
- Piketty, Thomas. 1997. "The dynamics of the wealth distribution and the interest rate with credit rationing." *The Review of Economic Studies* 64 (2):173–189.
- Roback, Jennifer. 1982. "Wages, rents, and the quality of life." *Journal of political Economy* 90 (6):1257–1278.
- Rosen, Sherwin. 1979. "Wage-based indexes of urban quality of life." *Current Issues in Urban Economics* :74–104.
- Rosenzweig, Mark R and Hans P Binswanger. 1993. "Wealth, weather risk and the composition and profitability of agricultural investments." *The economic journal* 103 (416):56–78.
- Roy, Andrew Donald. 1951. "Some thoughts on the distribution of earnings." *Oxford economic papers* 3 (2):135–146.
- Santos, Paulo and Christopher B Barrett. 2011. "Persistent poverty and informal credit." *Journal of Development Economics* 96 (2):337–347.
- Sotelo, Sebastian. 2020. "Domestic Trade Frictions and Agriculture." *Journal of Political Economy* 128:2690–2738.
- Tarozzi, Alessandro, Jaikishan Desai, and Kristin Johnson. 2015. "The impacts of microcredit: Evidence from Ethiopia." *American Economic Journal: Applied Economics* 7 (1):54–89.
- Tibshirani, Robert. 1996. "Regression shrinkage and selection via the lasso." *Journal of the Royal Statistical Society: Series B (Methodological)* 58 (1):267–288.
- Townsend, Robert M. 1994. "Risk and insurance in village India." *Econometrica* :539–591.
- Tsivanidis, Nick. 2023. "Evaluating the impact of urban transit infrastructure: Evidence from Bogota's Transmilenio." .
- Uganda Bureau of Statistics. 2021. "The National Labour Force Survey 2021 – Main Report." URL https://www.ubos.org/wp-content/uploads/publications/11_2022NLFS_2021_main_report.pdf.
- Van Leemput, Eva. 2016. "A Passage to India: Quantifying Internal and External Barriers to Trade." International Finance Discussion Papers 1185.
- Vitali, Anna. 2024. "Consumer search and firm location: theory and evidence from the garment sector in uganda." *Technical Report* .
- Wineman, Ayala and Thomas S Jayne. 2018. "Land prices heading skyward? An analysis of farmland values across Tanzania." *Applied Economic Perspectives and Policy* 40 (2):187–214.