Essays on Physical and Financial Well-being in Low- and Middle-income Countries

Von der Wirtschaftswissenschaftlichen Fakultät der
Gottfried Wilhelm Leibniz Universität Hannover
zur Erlangung des akademischen Grades

Doktorin der Wirtschaftswissenschaften
- Doctor rerum politicarum -
(Dr. rer. pol.)
genehmigte Dissertation

von

M.A. Yuanwei Xu,
geboren am 27. September 1995 in Jiangsu, China

2022

Referent: Prof. Dr. Lukas Menkhoff
International Economics, DIW Berlin (Deutsches Insitut für Wirtschaftsforschung)
Wirtschaftswissenschaftliche Fakultät, Humboldt-Universität zu Berlin

Korreferentin: Prof. Dr. Krisztina Kis-Katos
Wirtschaftswissenschaftliche Fakultät, Georg-August-Universität Göttingen

Tag der Promotion: 17.03.2022

Abstract

Many socio-economic factors are threatening the physical, mental, and financial well-being among the population residing in low- and middle-income countries, including financial stress, discrimination, violence, health issues, and inequalities. In the five chapters of my dissertation, I exploit experimental or quasi-experimental natures of socio-economic shocks in China, Peru and Uganda, to measure their impacts on the physical and financial well-being on individuals. The first chapter of this dissertation provides information on the study background, the data and presents an overview of each chapter. The remainder of the dissertation consists of five essays stretching out over the next chapters.

China is well-known for its skewed sex ratios and huge demographic imbalance. In the first chapter, I exploit a cross-cohort spatial comparison to analyze the impact of demographic imbalance on the payment of brideprices and dowries in China. This chapter exploits the variation of sex ratios exposed to children born in the same natal family but born in different years. I use a difference-in-difference estimation and provide the very first empirical evidence that an increase in male surplus leads to higher incidence and value of brideprices, but it has no effect on dowries. Subsequent investigations reveal that brideprices and dowries carry different significance in the Chinese society, which explains the co-existence and co-development of both payments in the 21st century.

Chapter 3 analyzes the negative health impact of rising housing prices in China. There is massive housing price appreciation since the early 2000s, which caused huge financial stress to households because of the strong social norm of owning rather than renting a house or apartment. We highlight that the high competition among males in the marriage market due to the "missing women" phenomenon is an important factor that contributes to such a negative health effect.

Chapter 4 shows that natural disasters would elevate the incidence of intimate partner violence through higher alcohol use by the male partner, higher likelihood of co-residing with the male partner, and an increase in male intra-household economic power. Moreover, we highlight that the access to protective institutions, such as the "women justice centers", are of vital importance to women's vulnerability of intra-household violence. This paper points out one important but often neglected socio-economic consequence of natural disasters, and provides a strong policy implication for the post-disaster relief and reconstruction.

The last two chapters are based on data collected in Uganda. The fourth chapter evaluates the impact of a financial education program on the use of digital finance (mobile money) as well as the business performance and household finance among micro-entrepreneurs in rural Uganda. We design a randomized saturation experiment, and find a positive impact of financial education on the use of mobile money, saving promotion,

and business investments on targeted entrepreneurs. At the same time, we do not find evidence of positive spillover effects on the untreated ones in treated clusters. Instead, the estimated spillover coefficients often show negative (albeit insignificant) signs. In the last chapter, we document that the COVID-19 lockdown increases the financial stress of the micro-entrepreneurs in rural Uganda in the long term.

Keywords: Gender, Health, Family, Financial Well-being, China, Uganda, Peru.

Acknowledgements

Writing this thesis has been an extraordinary experience and would not have been possible without the help and support of many people.

First of all, I want to thank Prof. Lukas Menkhoff. He gave me the opportunity to pursue my Ph.D. and the freedom to follow the research ideas I am passionate about. Thanks for giving me the chance of running a field experiment in Uganda, and for the helpful comments on my other research papers. I am very grateful to Prof. Krisztina Kis-Katos, who has been giving me constructive feedback since very early stages of my papers, as well as the great mental support and care. She is my role model in academia, and she convinced me to try an academic career. I would not have been able to finish this thesis with joy and confidence without her support. In addition, I would also like to thank Prof. Andreas Löschel for offering me a position after this short PhD funding. Without it I might have quit this PhD in winter 2020 and looked for another job to make my ends met. Moreover, this PhD journey would not have started without the aspiration by Prof. Stephan Klasen and Prof. Sebastian Vollmer. Thank you for your encouragement during my exchange studies and for building and maintaining the development kingdom in Göttingen.

I also thank the German Research Foundation for the financial support during my thesis via the Research Training Group on "Globalization and Development" (RTG 1723). Being part of the GlaD program was a great opportunity. It was the first time I felt being a part of one group during my stay in Germany. I did not feel alone thanks to the care and company of the 12 other fellow GlaD students (Laura Barros, Andrea Cinque, Anna Gasten, Richard Haarburger, Tobias Korn, Johannes Matzat, Tatiana Orozco Garcia, Anna Reuter, Lisa Rogge, Claudia Schupp, Henry Stemmler, Lukas Wellner) and Dr. Friederike Lenel. I could not have focused so much on research without the support and the great coordination work by Prof. Krisztina Kis-Katos, Prof. Martin Gassebner, Dr. Melanie Grosse, Dr. Celia Mewes, Antje Juraschek, Carmen Fernandez Leon and Helga Hemmersmeier. I would also like to thank Prof. Stephan Klasen and Prof. Andreas Wagner for taking care of us, and you will always live in our heart. In Berlin, I also benefited from the help and coordination of Anja Kegel. In Münster (before the chair of Löschel switched to Bochum), Ulrike Mink's great support enabled me to fully focus on my research. Your help and company made me enjoy the PhD journey.

I am grateful to my co-authors, Laura Barros, Jana Hamdan, Tim Kaier, Lukas Menkhoff, and Feicheng Wang. Working with you has been an invaluable experience, and I am thankful to have had the opportunity to work with you. Thank you, Jana, in addition,

for being a great field mate during the field work in Uganda. Your optimism was essential for me to go through those crazily stressful field days. I am also grateful for the guest speakers in Göttingen and Berlin. Thank you for offering office hours and meeting junior PhD students. Thanks as well for the warm encouragement and helpful comments. I am also thankful for the lovely people I met in the conferences, seminars, and workshops, who not only gave me feedback on my research papers but also motivated me to move further in the academic career.

Thanks to all my colleagues from Göttingen, Hannover and Berlin. In Göttingen and Hannover, thank you, my GlaD crew, as well as the lovely ones from the previous cohorts and other chairs, for all the lunch and coffee breaks and the fruitful discussions. In Berlin, I am grateful to my mentor, Dr. Helke Seitz, for your support during my most struggling time in the PhD journey. Thank you, Jana, again, for making our office in DIW Berlin a warm and lovely place. I would also like to thank Dr. Melanie Koch and Lukas Boer in addition for the nice relaxing chats, and also for switching to English when I was present. I always felt relieved when you were there. I am the only foreigner in the chair, and I was not able to speak fluent German in the beginning. I was not brave enough to ask people to switch to English and felt being excluded from many discussions. Later, I realized how important it is to be able to speak German to attend conferences, workshops, and seminars in German speaking areas. The language skill serves as an entry ticket to join most of the chats in the coffee breaks, lunch or dinner, although all the economists are fluent in English. The German economic network is highly dominated by native German speakers but nevertheless I, being part of the foreign minority, would feel more comfortable if fellow economists would be more considerate and eager to switch to English when a foreigner is around. Thanks to all of you, the past four years have been one of the greatest experiences of my life.

This thesis would not have existed without the support from my fitness group. Sport was extremely important for me to relieve my stress and keep a balanced lifestyle. I am very grateful for Alex from Göttingen for the amazing cycling courses. Moreover, I would love to express my gratitude to my amazing trainers in Berlin who offered fantastic training sessions, especially during the long lockdown time. These include: Antonia, Asal, Emanuela, Maike, Martina, and Lea. Thanks as well to Jungmi (Jay) and Marina for offering their living rooms as the training space. Without those sessions I could not have overcome both lockdowns, and I could not have worked on my research papers with a smile. I would also like to thank Martina Vitsentsos and Mitch Vanderlinden specifically for encouraging me to be brave and to accomplish so many missions impossible in the

journey of becoming a fitness trainer. It was a great life changing experience, during which I learned to speak loudly with power and strength in public and gained confidence enormously. I could not have made it without your endless support and patience. You are the best mentors one could ever wish for. I also thank my training partners, Stefan, Kathi, Maite, for your company and support. With you there I enjoy the sessions much much more.

I am as well super grateful for my lovely friends from my hometown, most of whom are fighting a life in Shanghai now. Thank you for always being so close to me mentally and for understanding me so well, and you make me feel I am not a lonely weirdo in this world. It is amazing how easy we are able to communicate to each other. In addition, the songs by Jay Chou accompanied my final stages in putting this thesis together. They have been a part of my life since I started attending schools in 2001. Surprisingly, they still sound as amazing as then. Thanks to those songs as well, I am constantly reminded how beautiful my mother tongue, Mandarin, is.

Moreover, and naturally, I thank my parents, Chu Li and Xu Gang (in Chinese, the family name comes before the given name), and my grandparents, aunt and uncle from my mother's side for your unconditional love and support. I am very grateful that you have always been strict with my studies and encouraged me to become a better version of me and reach a higher goal in my life. Thank you as well, that you do not discriminate me because of my gender, and that you always push me to work hard and fight a life like a boy. I also thank my partner, Oliver Braun, for your love, patience and unwavering support in everything I do. You are always there for me, no matter what, which is the best gift one could ever hope for. I am also grateful for Oliver's parents, Lars and Marion. Thank you for making my life here in Germany less struggling, and thank you for offering me another home here. Life would not be so much fun without you.

Last but not least, I would like to thank both the Chinese and German society and the well-functioning institutions. I feel extremely lucky that I was born in China and was exposed to the culture and value. I am grateful for all the people I have encountered in my childhood and adolescence, from whom I learned that I should work hard, but I shall never allow any obstacles to give me a long low time, and that I should always take good care of my own physical, mental, and financial well-being. I am lucky as well that the schooling system in my hometown was completely merit-based. The schools and universities of high quality were mostly public and there has never been a financial barrier to get access to descent education during my younger life in China. I am also thankful

that German public universities do not charge tuition fees, also for international students, and that the cost of living in Germany is not ridiculous. People with my socio-economic background would have a very low chance of pursuing a PhD degree, and the German society made it possible for me. I would hardly have made it this far if I had tried the same path in some other western countries, like the US or the UK, where international students bear huge amounts of tuition fees. However, fighting a life in Shanghai with my socio-economic background would even be harder than in Germany. It takes the wealth accumulation of at least two generations to have a life with some quality, safety and leisure there as a migrant. I do hope that it could become easier for internal migrants to integrate well in first-tier cities in China in the future.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

> "Many people would also agree with Amartya Sen, the economist, philosopher and Nobel Prize Laureate, that poverty leads to an intolerable waste of talent. As he puts it, poverty is not just a lack of money; it is not having the capability to realize one's full potential as a human being."
>
> *Abhijit V. Banerjee*

Many socio-economic factors are threatening the physical and financial well-being among the population residing in low- and middle-income countries (LMIC), preventing them from realizing the "full potential as a human being". These include, for example, financial stress, discrimination, violence, health issues, and inequalities. Gender discrimination is a prevalent issue in the developing world. For example, there is a "missing women" phenomenon in South and East Asia since the early 1980s, which is caused by sex selective pre- and post-natal abortions (Sen, 1992). In addition, across the globe, more than one third of the women suffer from any type of intimate partner violence (WHO, 2013). Financial instability and stress is widely faced by the vast majority of the population not only in LMIC but also in the developed world. Understanding the contributors of the threats to well-beings in LMIC would have important policy implications to improve people's lives.

"Missing Women" Phenomenon. It is estimated that more than 100 million women are missing in the world (Sen, 1992), and most of them are found in China and India. As shown in Figure 1.1, the most skewed sex ratios are observed in Asia, predominantly in China, India, Pakistan and Vietnam. In China, the imbalanced sex ratio could be attributed to four factors. First, China, like other East Asian societies,

has a long history and strong culture of son preference. Unlike sons, daughters would marry out of the family and are not able to continue the lineage by passing down the family names. Second, the introduction of the One Child Policy substantially increased the opportunity cost of having a child of the "unwanted" sex" (Bulte et al., 2011; Ebenstein, 2010; Li et al., 2011). Prenatal and postnatal abortions of daughters were in practice secretly. Third, the ultra-sound technology was widely available in the 1980s (Chen et al., 2013), which significantly reduced the cost of prenatal selection. Last but less well-known, the rural land reform, the Household Responsibility System, which took place from 1978 to 1986, is also found to contribute to the sex selection (Almond et al., 2019). The rural land reform shifted agrarian decision-making from the collective to individual households, and more males were selected due to their higher productivity in agriculture. All these reforms and technology advancement lead to higher sex imbalance and male surplus in China since the 1980s.

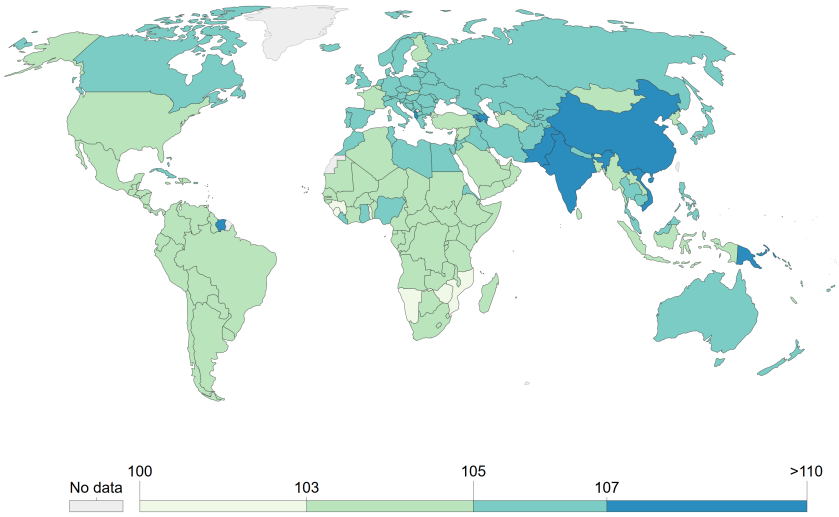


Figure 1.1: Sex Ratio at Birth in the World in 2019

Sources: This map is visualized by Ritchie and Roser (2019) based on data from the United Nations Population Division. The sex ratio is calculated by the number of new-born boys for every 100 new-born girls.

The surplus of men and the lack in women have caused several issues in the Chinese society. The unmarried rate among the young men increases, especially those with low socio-economic status (Huang and Zhou, 2015). The left-over unmarried men associates to higher crime rates (Edlund et al., 2013). Higher competition among men to attract a female partner pushes them to engage in more risky income generating activities (i.e. entrepreneurship) (Wei and Zhang, 2011b). Moreover, parents with

unmarried sons also save more to increase their attractiveness in the marriage market (Wei and Zhang, 2011a).

The influence of the fierce marriage competition among young men has also extended to the housing choices. There is a strongly rooted marriage culture in China in which men are obliged to obtain a house before marriage. Houses in China are important status goods that signal wealth and social status (Wei et al., 2017) and men have stronger incentives to pursue bigger houses to increase their competitiveness in the marriage market. In the meantime, China witnesses the real estate prosperity and rapid housing price growth since the 2000s. These factors taken together have caused a huge financial stress to young men in the marriage markets and their parents.

Micro-entrepreneurs in Rural Uganda. Many households are living and businesses are operating at (or under) the substance levels in low income countries. As seen in Figure 1.2, almost all the sub-Saharan African countries exhibit high poverty headcount ratios. It has been documented that poor people are less patient (Duflo and Banerjee, 2011), and they may focus more on current consumption and save less. As a result of the low capital accumulation, they would have less resource for risk coping or business investments, and thus are trapped in poverty. It is thus of high policy relevance to explore how to empower the vulnerable population and drag them out of this poverty trap.

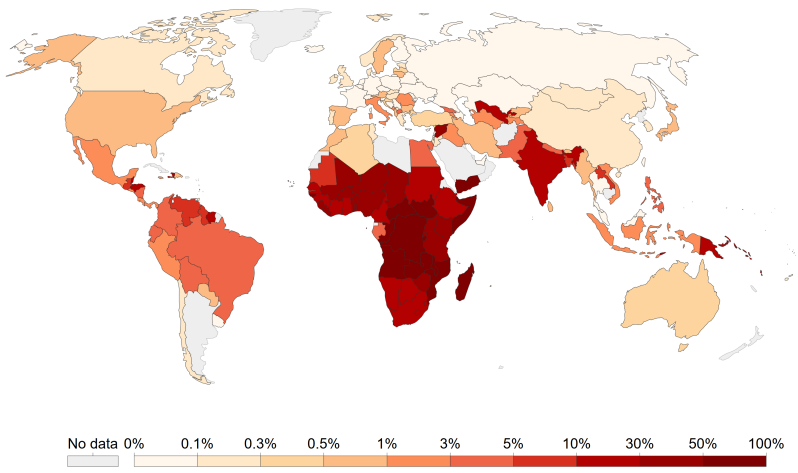


Figure 1.2: The Share of Population in Extreme Poverty in the World in 2013

Sources: This map is visualized by Roser and Ortiz-Ospina (2013) based on data from the World Bank PovcalNet. Note: extreme poverty is defined as living on less than 1.90 international dollars (purchase power parity adjusted) per day. The measure relates to household income or consumption per capita.

In Uganda, about 75 percent of the population reside in rural areas (WDI, 2020). Very few are formally employed and most of them work for themselves or family members (Merotto, 2019). The micro-entrepreneurs in rural Uganda usually run a family business in the form of small shops. They may at the same time engage in agricultural production or other income generating activities. A typical micro-enterprise either does not expand much in scale or would disappear after a few years (McKenzie and Woodruff, 2017).

This dissertation contributes to uncovering different factors that threaten the physical, mental and financial well-being of the population living in LMIC. It consists of five essays that are presented over the next chapters. The first two essays focus on the gender discrimination and financial stress in China, the third chapter sheds light on gender violence in the aftermath of natural disasters in Peru, and the remaining two essays concentrate on the financial well-being among micro-entrepreneurs in rural Uganda. In particular, the first essay studies the impact of sex imbalance on marriage payments and the different significance that brideprices and dowries carry in the Chinese society. The second essay documents the adverse impact of housing price appreciation on health in China. The third essay files the elevated intimate partner violence (IPV) in the aftermath of natural disasters in Peru. Essays four and five concentrate on the financial well-being of the micro-entrepreneurs in rural Uganda. Specifically, the fourth essay evaluates the direct impact and spillover effects of a financial education program, and the fifth one clips the persistent negative impact of COVID-19 lockdown rules on the micro-entrepreneurs.

The findings of this thesis provide the following policy implications. First, in China, there is need of social programs based in schools and workplaces where harmful gender norms should be eliminated to prevent further sex discrimination and selection. Second, in China and other countries with rapid real estate appreciation, complementary policies that alleviate the negative health consequences should be considered in the design of housing policies. Third, in the post disaster relief programs, the facilitation of protective institutions for women need to be considered as well. Fourth, the development finance projects in the global south need to include all eligible individuals and households into the programs, as no positive spillover effect is observed. Finally, COVID-19 causes a long lasting rise in financial stress among individuals in low income economies. By further promoting the adoption of mobile financial services and ensuring consumer protection, specifically with regard to loans and repayment flexibility, governments in low income settings may improve general welfare.

Data Sources

All chapters are based on household survey data in China and Peru, as well as self-collected data on micro-entrepreneurs in Uganda.

Specifically, in China, I use survey data from the China Health and Nutrition Survey (CHNS), the China Health and Retirement Longitudinal Study (CHARLS), the China Family Panel Study (CFPS) and the census data. CHNS is the longest panel survey in China, and the publicly available data includes ten waves from 1989 to 2015. CHARLS is the only nationally representative dataset that systematically collects information on the brideprice and dowry payments in China. In total there are three waves between 2011 and 2015, and I use the survey data collected in 2013, where the child marriage data is available. The CFPS dataset is the largest nationally representative panel dataset. The survey starts in 2010 and contains five waves until 2018. I use the waves between 2010 and 2016 in this dissertation. With regards to the Chinese census data, I project the sex ratios among different age cohorts within cities using the census data in 2000.

In addition, we utilize the Demographic and Health Survey (DHS) data in Peru from 2000 to 2009 for Chapter 4. We combine the DHS data with the geo-spatial data on earthquakes.

Finally, the last two chapters in Uganda are based on primarily collected data in 2019, 2020 and 2021. The data collected in 2019 is based on face-to-face interviews. Then we follow up with the sample first by means of phone interviews in winter 2020. In spring 2021, we follow up with the attriters in the phone surveys using a face-to-face interview. All the survey data used in this dissertation is representative of the national or regional population.

Summary of Chapters

Chapter 2 analyses the effects of sex imbalance on marriage payments. Although there have been arguments that the rising male surplus caused the high brideprices in China, there was no empirical evidence supporting the causal link. In addition, dowries increased along with brideprices. This is contradicting the predictions from the family economic model, that brideprices and dowries should change in opposite directions (Becker, 1991). This paper provides the first empirical evidence documenting the demographic effects on marriage payments, and reveals the different significance of brideprices and dowries in the Chinese society.

Empirically, I use the household-level data containing information on marriage payments by parents collected in the 2013 survey of the CHARLS. In addition, I

calculate the residential sex ratios projected from the 2000 census, and match them with the marriage payment information from the CHARLS data using the child's birth city and year. The identification of a causal link between demographic imbalance and marriage payment relies on the comparison between children who are born in the same natal family but exposed to different sex ratios as they are born in different years. Controlling for the natal family fixed effects allows for taking into account many cross-family heterogeneous factors affecting the marriage formation, including norms, preferences, etc.

Employing a linear (probability) model, I show that higher male-female sex ratios lead to higher prevalence and value of brideprices, however, dowries are not affected by sex imbalance. Moreover, dowries are paid out with an intergenerational implication as an exchange for future help and care in the old age, but brideprices are not. Overall, I contribute to the literature studying the socio-economic consequences of sex imbalance and unintended consequences of family planning policies in China, as well as the broad body of research on marriage formation and marriage markets.

Chapter 3 documents the negative effect of housing price inflation on individual health status in China. Many countries have experienced considerable housing price growth in the past decades. We contribute to the emerging literature on the socio-economic consequences of the rising housing prices, not only in China but also in other countries, offering further insights into the role of Chinese specific background (huge marriage competition among males due to the "missing women" phenomenon) in such an effect. We employ the CHNS panel dataset and link it with the province-year varying housing prices.

We control for a variety of individual, household, and province characteristics, as well as for individual fixed effects, which take out the time-invariate individual characteristics that could affect health directly. In addition, we use a difference-in-difference type instrument variable approach to address the endogenous housing prices. We find robust evidence that individuals from provinces that experienced more rapid growth in housing prices are more likely to have chronic diseases and hypertension relative to others. In addition, the negative health consequence of housing prices is closely linked to the deeply rooted marriage culture that males are obliged to provide a home for newly formed households and the increasing marriage market competition among young males in China due to the imbalanced sex ratios. Lastly, we show that increasing working activities among the employed, higher levels of stress, and changes in lifestyle, especially rises in tobacco consumption and declines in sleep time among the younger cohort could be other potential channels that shape the health consequences

of housing prices.

In Chapter 4, we investigate whether women are more vulnerable to intimate partner violence (IPV) in the aftermath of earthquakes in Peru. Empirically, we combine the household-level data on IPV from the DHS surveys and spatial data on all earthquakes that happened between 2000 and 2009, which are calculated by the United States Geological Surveys. The exposure to earthquake is measured at the disaggregated DHS cluster level. The occurrence of natural disasters is highly exogenous, allowing for a sound causal inference. In addition, we cross check the results using different specifications as well as using administrative data on the registered domestic violence cases to reduce the concern that women exposed to earthquake might have different reporting behaviors from those who did not experience a large-scale natural disaster.

We find robust evidence that exposure to very strong earthquakes increases the probability that women experience violence by more than ten percentage points on average. The effect persists up to 18 months after the earthquake occurrence. Moreover, we document that the effect is less pronounced for women residing in districts with protective institutions or in rural areas. Further evidence suggests that the increase in IPV following earthquakes is induced by a rise in male intra-household economic power, elevated use of alcohol by the male partner, and higher probability of male partner being in the household.

Chapter 5 and 6 are based on a field experiment located in rural western Uganda. They are motivated by the fact that many households are living and businesses are operating at the substance levels. It is of high policy relevance to understand whether certain policies or interventions could effectively improve the well-being of this population at affordable costs.

In the rural part of East Africa, a large population has been financially excluded. Bank branches are far away from them, and many do not have qualified documents to apply for a bank account. They strongly rely on informal ways to conduct financial activities, for example, by putting savings in the form of cash at home, and by borrowing from friends and family only. Many studies have pointed out that the widely spreading mobile money, a digital financial service, could solve the financial exclusion problem. However, Hamdan et al. (2021) pointed out that despite the high mobile money account ownership, the real use of mobile money for financial activities is low. We hypothesize that this could be due to two reasons. Firstly, mobile money exhibits a very complicated cost structure, and the costs are especially high when the transaction amount is low. Thus, the micro-entrepreneurs might pay very high fees and may

not use the service. Secondly, the population may have low trust in mobile money due to their generally low trust in formal financial services. It is observed that very few micro-entrepreneurs have deposits in the mobile money account, and many would withdraw small amounts from the account despite the high withdraw transaction fee. In Chapter 5, we design a financial education program including household finance management, business development, and guidance on how to efficiently use mobile money.

We study the impact of a five-hour financial literacy program and the spillover effects on about 2,000 micro-entrepreneurs in rural Uganda using a two-stage randomized experiment. We first randomize the program at the trading center (or cluster) level, and then we randomize the treatment intensity within clusters. The long-term impact is evaluated 13-19 months afterwards, and we find that those targeted in the treated clusters are more likely to have formal savings, higher formal savings, more and larger business investments, more and higher savings in their mobile money account and more often using new payment functions of mobile money. However, we do not find evidence of positive effects on the spillover group but quite often unexpected negative spillover coefficients which require further attention. This suggests that the future programs need to include as many eligible participants as possible, since there are no positive (or even negative) spillover effects observed.

In the last chapter, we provide evidence on the impact of the COVID-19 lockdown on the financial well-being of micro-entrepreneurs in a low income setting. The analysis is based on regionally representative panel data on 1,975 micro-entrepreneurs from rural Uganda. We first show that several business characteristics predict the non-compliance to the national lockdown, including, for example, running a service business. Further, we document a sharp increase in overall household informal savings, loans, as well as the use of mobile money compared to pre-COVID levels. Moreover, we find a substantial drop in business investments and profits. Difference-in-difference estimation results suggest a decrease in overall financial well-being more than six months after the national lockdown. In addition, we find that the longer the individual business shutdown due to COVID-19, the smaller the *ex-post* business profits, investments and the lower the financial safety. These findings point towards a need to design policy instruments aiding small businesses, and service businesses in particular, to help them step out of their ongoing weakened economic situation.

Chapter 2

Paying for the Selected Son: Sex Imbalance and Marriage Payments in China*

Single-authored Paper

* I am grateful to Siwan Anderson, Anna Gasten, Jana Hamdan, Krisztina Kis-Katos, Lukas Menkhoff, Nancy Qian, Lisa Rogge, Helke Seitz, Maria Micaela Sviatschi, Lingwei Wu, and seminar and conference participants at the University of Göttingen, Tsinghua University, the Annual Meeting of the German Economic Association (VfS) 2020 and 2021, the RGS Doctoral Conference, the Annual Meeting of the Chinese Economists Society, the Oxford Development Workshop for valuable discussions and comments. I would also like to thank Oliver Braun and Markus Rafael Paffhausen for their excellent assistance in proofreading the paper. All remaining errors are mine.

Abstract

This paper shows that the rising surplus of males in China has thwarted male marriage formations by causing an increase in brideprice payments. The identification relies on the comparison between siblings from the same natal family who are born in different birth years and thus exposed to different demographic structures. I find robust evidence that a rise in male-female sex ratios significantly increases the incidence and market value of brideprices, but has no influence on dowries. Such a positive effect on brideprices is found predominantly in natal families characterized by low socioeconomic status, smaller number of children, and with more daughters than sons. Dowries are found to carry an intergenerational function for help and care in parents' old age but brideprices are not. Further investigations suggest that dowry values are positively associated with female welfare. This paper provides the first empirical evidence showing that demographic imbalance causes marriage distortions with a rise in brideprices, and suggests that brideprices and dowries carry different significance in the Chinese society.

JEL-Codes: J12, J13, J16, P21
Keywords: Sex Imbalance; Brideprices; Dowries; Marriage Payments; China

2.1 Introduction

The skyrocketed brideprices, which induce huge financial burdens on families with unmarried sons in China, have gathered international attention and concern[1]. Although there have been arguments that the rising male surplus caused the high brideprices, there was no empirical evidence supporting the causal link. In addition, dowries increased along with brideprices. If the sex imbalance contributes to the inflation of brideprices, why is there no deflation of dowries? This paper provides the first empirical evidence documenting the demographic effects on marriage payments, and reveals the different significance of brideprices and dowries in the Chinese society.

There are two types of marriage payments. Brideprices, defined as transfers of goods or money from the family of the groom to the bride's, are highly prevalent in Sub-Saharan Africa. On the other hand, dowries, transfers of the opposite direction, are extensively documented in South Asia. In most of the low- and middle-income countries, only one type of marriage payment is observed, but China is a unique setting where both brideprices and dowries are paid out [2]. Nevertheless, dowries differ substantially from brideprices in China, with the latter being compulsory while the former being voluntary. Moreover, dowries could be financed with a return portion of the brideprices, mostly in rural areas (Engel, 1984).

China experienced a sharp rise in the male-female sex ratios since the 1980s (Sen, 1992), due to several reforms and technology changes. These include the One Child Policy (Bulte et al., 2011; Ebenstein, 2010; Li et al., 2011), the availability of the ultra-sound technology (Chen et al., 2013), and the rural land reform (Almond et al., 2019). Previous studies document that male surplus leads to higher obstacles in male marriage formation (e.g., Edlund et al., 2013; Huang and Zhou, 2015), but the effect of surging sex ratios on marriage payments remains neglected. In this paper, I empirically show that the marriage market would adjust to such a demographic imbalance through an increase in the incidence of paying and the market value of brideprices.

The rise of both marriage payments and male surplus in China is well illustrated in Figure 2.1. Several policy changes and technological advancement elevated the sex imbalance since the end of the 1980s. The vertical line marks the official start of the One Child Policy in 1979. The male-female sex ratio rose from about one among those born in 1980 to 1.13 in 1990. In the meantime, parents with sons born between 1960

[1] For example, see this BBC post: https://www.bbc.com/news/blogs-trending-35727057 (last accessed on October 27, 2020)

[2] See Figure 2.A.1 in the Appendix (Section 2.8). Although both brideprices and dowries are paid out, the average prevalence of brideprices is higher than dowries in all prefectual cities.

and 1980 had a probability of about 60 to 70 percent of paying brideprices for sons' marriage, but those with sons born after 1980 had a probability of over 80 percent to pay brideprices (see Figure 2.1a). The monetary value, as shown in Figure 2.1b, increased drastically among those with sons born after 1980 as well. The average value of brideprices amounted to an average of about 40,000 Chinese Yuan (about 5,714 USD) among those born around 1990, which was equivalently six to seven times the rural annual individual disposable income, and about twice that in urban areas in 2010. On the other hand, although the incidence of dowries did not see a clear increase from 1960 to 1990, the value of dowries witnessed an increase among daughters born after 1980 as well, but at a lower speed compared to brideprices.

In this paper, I study two research questions, (i) whether the male surplus contributed to the inflation of marriage payments in China; and (ii) why brideprices and dowries co-existed and co-evolved. I explore the role of the sex imbalance in marriage payments by exploiting the spatial and temporal variation of demographic shifts and linking that with child-level marriage payment information by birth cities and years. Specifically, I use the household-level data containing information on marriage payments by parents collected in the 2013 survey of the China Health and Retirement Longitudinal Study (CHARLS). It is a nationally representative survey with individuals aged 40 years and above, with the primary goal of studying the physical functioning, financial well-being, and intergenerational relationships of the elderly population in China. This dataset has several advantages. First, it is the only nationally representative dataset that systematically collects detailed information on marriage payments based on self-reports from parents in China. Second, the complete birth history of parents is recorded. This allows to identify and address numerous family and child characteristics, including family size, birth order and birth interval of each child, which might fundamentally affect the allocation of family resources among children in the form of marriage payments. Third, parent-child pair-wise information on monetary transfers and care-taking is well documented in the dataset. This enables to explore underlying incentives of marriage payments in intergenerational decision-making other than the supply and demand changes in the marriage market. For the main explanatory variable regarding demographic imbalance, I calculate the city and birth year level residential sex ratios projected from the 2000 census.

The identification of a causal link between demographic imbalance and marriage payment relies on the comparison between children who are born in the same natal family but exposed to different sex ratios as they are born in different years. There are several potential threats in causal inference. To begin with, there is potential

endogeneity in residential sex ratios which arises from the gender profile of internal migration. Females are more likely to leave economically depressed areas than males (Fan and Huang, 1998; Edlund, 2005), and thus, the sex ratios in areas with better economic conditions might be lower. On the one hand, I control for natal family fixed effects to adjust for initial geographic variations in economic conditions and development. On the other hand, I additionally control for province-year specific per capita GDP and annual disposable income to capture the income effect as a robustness check. Secondly, the sex ratios could be affected by the temporal variations in the intensity of family planning policy and technology advancement. For example, children born in years with higher penalty rates of a second birth might be exposed to higher male-female sex ratios. I further control for the province-year variation of One Child Policy fine rates as calculated by Ebenstein (2010) and the availability of ultra-sound technology in capital cities by Chen et al. (2013) as a robustness check. However, the significance and size of the effect remain unchanged. Lastly, there might be the concern that parents might change sex selection in light of the high marriage payments (Bhalotra et al., 2020). However, the marriage payments are paid out upon children's marriage, which is about 20 years after the birth. It is highly unlikely in the Chinese context that the brideprices paid out at marriage would change the sex ratios at birth.

The empirical results show several interesting findings. To begin with, I find robust evidence that higher male-female sex ratios lead to higher prevalence and value of brideprices, however, dowries are not affected by sex imbalance. This suggests that brideprices respond to demand and supply shocks in the marriage market but dowries do not. Second, the positive demographic effect on brideprices is only observed among males with low socio-economic status, which confirms that the brideprices in China carry the function as a status good to improve male attractiveness or position in the marriage competition. In addition, the family structure matters. The effect is found dominantly on sons from natal families with fewer children and with more daughters than sons. Finally, I further examine the role of marriage payments in intergenerational as well as intra-household relations. Dowries are paid out with an intergenerational implication as an exchange for future help and care in the old age, but brideprices are not. Lastly, higher dowry values are associated with higher female welfare.

This paper is related to several strands of literature. First, this work directly adds to the emerging literature studying the socio-economic consequences of sex imbalance and unintended consequences of family planning policies in China. For instance, high male-female sex ratios associate to higher crime rates (Edlund et al., 2013), increased unmarried rate of males (Huang and Zhou, 2015), a boost in entrepreneurship and

economic development (Wei and Zhang, 2011b), and elevated savings rate (Wei and Zhang, 2011a). It is well documented that the One Child Policy is one important cause of surging sex imbalance in China (Ebenstein, 2010; Bulte et al., 2011; Li et al., 2011). Moreover, high One Child Policy fines lead to birth mis-reporting (Merli and Raftery, 2000; Huang et al., 2016), child abandonment and abduction (Bao et al., 2019). There has been theoretical evidence that the imbalanced sex ratios would cause marriage squeeze (Ebenstein and Sharygin, 2009). This paper is one of the first to empirically investigate the causal effects of sex imbalance on marriage distortions in China, and provide an additional perspective by shedding light on the rise in brideprice payments.

Second, this paper fits within the broad body of research on marriage formation and marriage markets. There are several adjustments in the marriage market induced by shocks of demographic imbalance, including marriage payments (Rao, 1993; Botticini and Siow, 2003), marriage sorting (Abramitzky et al., 2011), and timing of the marriage entry (Anderson, 2007b; Corno et al., 2020). Using gold prices as exogeneous variation for dowry value, higher dowry value is found to lead to higher girl mortality (Bhalotra et al., 2020) but also higher female bargaining power in the new household (Menon, 2020) in India. The findings in this paper provide consistent evidence from China, where there is limited empirical research on marriage payments.

Last, this paper contributes to a set of literature on determinants and interpretation of marriage payments. As summarized in Anderson (2007a), the direction and magnitude of marriage payments depend on certain social and family characteristics. However, a vast majority of the studies are cross-country analyses, and provide anecdotal and qualitative evidence but not causal inference. A standard theoretical model was developed in Becker (1991), where brideprices and dowries were assumed to be the same in nature and were negatively associated with productivity, socio-economic status, and degrees of social equality. However, the co-existence of brideprices and dowries in China suggests that such an assumption might not be valid, and the findings in this paper show that marriage payments have other implications than being merely marriage market clearance prices. A new interpretation of dowries as pre-mortem inheritance was brought up in the theoretical models developed by Zhang and Chan (1999) and Botticini and Siow (2003), where dowries could increase a bride's control over household resources, rather than being a price for grooms. This paper is the first study which provides empirical evidence using individual-level data showing that brideprices would change in response to demographic imbalance. Moreover, it suggests that brideprices feature a status good in marriage market competition, and dowries involve intergenerational decision-making in China. The differences in their interpre-

tation and functions are important in explaining the co-existence in the context of China.

The remainder of this paper is organized as follows. Section 2.2 provides a description of the various datasets used in this paper, before Section 2.3 discusses the empirical approach. Section 2.4 presents empirical results on the demographic effect of marriage payments based on the CHARLS data, including the main results, robustness checks, as well as heterogeneity analyses. Section 2.5 discusses the intergenerational and intra-household roles of marriage payments. Finally, Section 2.6 concludes.

2.2 Data

The analysis combines data from two main sources: (i) demographic imbalance at the city level is projected from the census data in 2000; and (ii) the individual child-level marriage payment information is taken from the 2013 survey of the CHARLS.

For the main explanatory variable regarding demographic imbalance, I construct a panel of the male-female sex ratios among residents in each birth year and prefecture city using the census data collected in 2000. Specifically, I assume a spousal age gap of two years[3] and a five-year age window for spouse matching[4]. In the CHARLS sample, the distribution of age at marriage does not differ substantially between males and females, except that males enter marriage at an older age compared to females (see Figure 2.A.4 in the Appendix (Section 2.8)). The 2000 census contains information on the city of residence at the time of enumeration and the province of birth. Thus, the 2000 census can be used to calculate two sex ratios, which are the residential sex ratio by city and the birth sex ratio by birth province. The ideal demographic structure should be proxied by the residential sex ratios in the marriage year and city using the natality data. However, there is no nationally representative data on vital registration of births, deaths, or marriages (Yang et al., 2005). Thus, I could only rely on projections from the decennial population censuses. The residential sex ratios are preferred over birth sex ratios because it measures the demographic structure taking into account not only local natives but also internal migrants. On the contrary, the birth sex ratios do not consider in- and out-migrants. Using birth sex

[3] According to the 2000 census, males are on average two years older than their female partners among all married couples.

[4] For example, the 1980-born sex ratio for males is the same as the 1982-born sex ratio for females, which is calculated as the number of males born between 1978 and 1982 over the number of females born between 1980 and 1984 within the same city. Please see Figure 2.A.3 in the Appendix (Section 2.8) as an example

ratios might induce measurement errors in cities where migration outflow or inflow is gender specific. It is well noted in Fan and Huang (1998) and Edlund (2005) that females, relative to males, are more likely to leave economically depressed areas. Thus, the residential sex ratio is better able to precisely represent demographic imbalance than the birth sex ratio. Empirically, it tracks the birth sex ratio by birth province quite closely because internal migration is not pervasive in China. According to the CHARLS data, 75 percent of the children reside in the village of birth after marriage. According to another nationally representative micro dataset, the China Family Panel Study (CFPS) (Institute of Social Science Survey, 2015), less than five percent of the population migrate out of the birth province. Alternatively, city level residential sex ratios could be calculated using the 1990 census. Unlike the 2000 census where both birth province and current residential location are recorded, the 1990 census only collected current city and province of residence. However, internal migration was more restricted in 1990 than in 2000, and thus the residential sex ratio calculated using the 1990 census might not precisely reflect the sex ratios faced by younger cohorts who get married after 1990. However, from the CHARLS data, 83 percent of the children whose parents reported their year of marriage were married after 1990. In conclusion, I use the city-level residential sex ratio calculated using the 2000 census as the main measure for local sex imbalance, while the province-level birth sex ratio as well as the residential sex ratio projected from the 1990 census are used as alternative measurements in robustness checks.

The data on marriage payments as well as the socio-economic information come from the CHARLS survey data collected in 2013 (Zhao et al., 2015). This dataset records detailed information on physical and mental health, relationship with children, as well as care-taking status of individuals aged over 40 years in China. CHARLS 2013 covered 28 out of a total of 31 provinces in the mainland and, within each sampling unit it used probability proportional to size sampling method to be nationally representative. Moreover, both individual and household weights are available for most of the respondents in the data. The dataset collects detailed socio-economic information of parents and each of their children, as well as parent-child pairwise information on intergenerational transfers and care giving. Specifically, the CHARLS 2013 survey records whether and how much parents have paid brideprices or dowries when each child got married. Moreover, the geographic location of the respondents are made available at the city level, and this allows me to match with the variation of sex ratios. CHARLS 2013 also tracks the migration history in detail, and I only include parent-child pairs where parents have no migration record after the child's birth to

avoid measurement errors. The marriage payment information was assessed in 2013, however, only very few born after 1990 were married before 2013. In addition, the dataset includes only a few parent-child pairs where children are born before 1960. This is perhaps because the parents with children born prior to 1960 are usually from pre-1940 birth cohorts, and many of them might have already passed away in 2013. Thus, I restrict the analysis to the parent-child pairs where the child was born between 1960 and 1990.

The final sample includes 13,927 parent-child pairs with children were married before 2013 and have at least one sibling in the natal family[5]. Table 2.A.1 in the Appendix (Section 2.8) presents descriptive statistics of the key variables. The average residential sex ratio is 1.05, that is, 105 males per 100 females. 56 percent of the parents have paid marriage payments, with an average value of 9,291 Yuan (approximately 1,327 USD using current exchange rate). [6] On average, 51 percent of the records are parent-son pairs, and the rest are parent-daughter records. 83 percent of the children have at least primary school education, but 43 percent of the parents have not finished primary school. 98 percent of the children are biological children. 30 percent of the parents currently live in urban areas, and they have 3.76 children on average.

The increasing time trend of marriage payments coincides with that of economic development. After several successful reforms and market-oriented liberalization, China has witnessed drastic economic growth since the 1980s. The rise of marriage payments might simply be an income effect if the trend of demographic shifts correlates with that of economic growth. In order to control for this, I collect province-level data on per capita Gross Domestic Product (GDP) and disposable income from the *China Statistical Yearbook* from 1960 to 2010. I match the income data using the marriage year if the children's marriage year is reported by parents[7]. Otherwise, I assume that males get married at 23 years old and females at 21 years old, and I match the province income record with the year when they turn 23 or 21.

Lastly, the sex ratios might be endogeneous due to spatial and temporal variations in policies which might directly affect sex selection at birth, including family planning policies and ultra-sound technologies. I extract the fine rates of the One Child Policy measured in years of family income calculated by Ebenstein (2010) and the roll-out data of ultra-sound technology by Chen et al. (2013). These data are used as additional confounders to check the robustness of the demographic effect on marriage payments.

[5] In the CHARLS sample, only 4.5 percent of the families have only one child.

[6] The value of marriage payments takes a missing value if parents reported not having paid marriage payments when a child got married.

[7] In the sample, 77.23 percent of the children's marriage year was reported by parents.

2.3 Empirical Strategy

I consider the following model to investigate whether the rise in marriage payments could be explained by male surplus.

$$Payment_{ihct} = \alpha + \beta Sex\ Ratio_{ct} + \lambda \mathbf{X}_i + \mu_h + \zeta_t + \epsilon_{ihct} \tag{2.1}$$

On the left hand side, the outcome of interest is either a dummy indicating whether the parents h of child i, who is born in city c and year t, have paid marriage payments when the child was married, or the payment value after adjusting for inflation using Consumer Price Index (CPI). In cases that the outcome variable is the incidence of marriage payments, I use a linear probability model for the estimation. The marriage payments take the form of brideprices for sons and dowries for daughters.

On the right hand side, α is a constant and $Sex\ Ratio_{ct}$ is the average male-female sex ratio in city c and birth year t. The coefficient of interest, β, therefore measures the average impact of demographic imbalance on child-level marriage payments. $\mathbf{X}_i$ is a vector of child-level characteristics within the natal family that may represent their relative importance or preference from parents among all the siblings, which as a result might affect his or her marriage payments from parents. These include, for example, the birth order. As displayed in Figure 2.A.2 in the Appendix (Section 2.8), parents have a lower probability of paying marriage payments to the second-born child with the same sex relative to the first one. If their parents paid either brideprices or dowries, the value is generally lower than the first-born's as well. Specifically, these controls include sex, educational attainment, birth order, birth order within the same sex, length of birth spacing[8], age difference from parents[9], and a biological child dummy. The parents' decision on each child's marriage payments might also depend on parents' own characteristics. I control for the parent or natal family fixed effects, μ_h, to absorb all observable and unobservable heterogeneities regarding marriage formation and practices across different natal families. As such, the identification of the demographic effect originates from within-family variations in marriage payments across siblings who are exposed to different demographic structures as they are born in different years. Note that the family fixed effects also capture all time-invariant characteristics that

[8] For unique child and first child, birth spacing takes the value zero.

[9] I calculate the maximum age difference from both parents if both parents are identifiable, or the age difference from one if only one parent could be identified. In the sample, I was able to identify information of both parents among 75 percent of the children.

might affect marriage payments at the city and province level. At the time horizon, the birth year fixed effects, ζ_t, take into account all macro shocks to all children who are born in the same year. Lastly, ϵ_{ihct} is the error term. Standard errors are clustered at the city level to allow for potential correlations between individuals within cities.

The estimation of demographic effect on marriage payments relies on comparison between siblings from the same natal family. This requires that the estimation is only based on the sample where the natal family has more than one child. In addition, when the outcome variable is specified as brideprices (or dowries), then the estimation is based on the sample where the natal family has more than one son (or daughter). This might lead to a potential problem that the estimation might be based on an extremely small sample, especially when estimating the effect on the average market value of brideprices or dowries. Thus, I further include an interaction term between sex ratios and the male dummy to estimate the differential effect between brideprices and dowries based on all the records where parents have paid marriage payments when the child was married. Specifically, the model takes the following form:

$$Payment_{ihct} = \alpha + \beta_1 Sex\,Ratio_{ct} + \beta_2 Sex\,Ratio_{ct} \times Son_i + \lambda \mathbf{X}_i + \mu_h + \zeta_t + \epsilon_{ihct} \tag{2.2}$$

Then β_1 estimates the total effect of demographic imbalance on dowry value, and $\beta 2$ denotes the additional effect on the value of brideprices relative to dowries.

2.4 The Effect of Sex Imbalance on Marriage Payments

2.4.1 Main Results

We started with the demographic effect on the incidence of marriage payments and the baseline OLS regression results are set out in Table 2.1.

Columns (1), (3) and (5) report the results by controlling for parent fixed effects and birth year fixed effects, and columns (2), (4) and (6) further control for child characteristics. When children get married, parents are more likely to pay marriage payments for those children who are exposed to a higher male-female sex ratio, compared to those exposed to lower sex ratios (column (2)). The coefficient in column

(4) suggests that parents have a higher probability of paying brideprices for those sons who are faced with more drastic male surplus. However, the incidence of dowry payment does not vary with the demographic structure (column (6)). The coefficients presented in Panel B are adjusted for survey weights, and a similar effect is observed, as shown in Panel A. However, about 300 child-parent records in the sample were not assigned survey weights. Thus, I rely on specifications without weight adjustments in the following analyses.

Specifically, as indicated in column (4), when the male-female sex ratio increases by 0.01, that is, when there is one additional male for every 100 females, the incidence of brideprices is increased by 0.26 percentage points on average. The overall sex ratio increased from about one to 1.13 from the cohorts born in 1960 to those born in 1990, implying that the male surplus contributes to an increase in the incidence of brideprices of 3.4 percentage points. In contrast, the overall incidence of brideprices rose by around 10 percentage points (see Figure 2.1a), and the sex imbalance explains about a third of the total increase in brideprices. Moreover, if brideprices and dowries were similar in nature, an increase in sex ratios should have a negative effect on the incidence of dowry payment. However, this is not observed according to the estimate presented in column (6). The coefficient is positive, but insignificant. This further suggests that dowries might differ substantially in nature from brideprices in China, and could not be regarded as marriage market clearance prices as assumed in Becker (1991). I will further discuss the role of marriage payments in detail in Section 2.5.

As a second step, I further investigate whether sex imbalance contributes to the rise in marriage payment value (Table 2.2).

From columns (1) and (2) it is observable that an increase in sex imbalance does not have any effect on the value of dowry payments, but it leads to higher value of brideprices. Specifically, if there is one additional male per 100 females, then the value of brideprices would increase by about 60 Yuan (8.6 USD) on average. However, such a positive effect is not observed if I specify the natural logarithm of the value or its share of disposable income as the outcome variable. Moreover, when comparing children with siblings of the same sex, there is no significant effect of the male surplus on the value of brideprices (or dowries) paid by parents, as shown in Table 2.A.2 in the Appendix (Section 2.8).

Overall, the baseline results indicate that the sex imbalance in China increases the incidence and value of brideprices, but has no effect on dowries. The results here suggest that only brideprices respond to this specific shock in the marriage market.

2.4.2 Robustness Checks

In this section, I check the robustness of the positive effect on the incidence and value of brideprices driven by demographic imbalance. Specifically, I check whether the finding is sensitive to alternative measurements of sex imbalance and to including additional controls for various confounders. Further, I perform falsification tests to check whether a pre-sample time trend exists.

Alternative Measurements The main results in Table 2.1 are estimated based on the residential sex ratios as the measure of sex imbalance. In this section, I consider two alternative measurements of sex ratios, the birth sex ratios and the residential sex ratios projected from a previous census.

The results set out in Table 2.3 confirm the positive impact of male surplus on the incidence of brideprice payment (see columns (2) and (5)). The effect sizes are similar to the results in Table 2.1 and they are both significant at the one percent level.

Economic Development as Confounders

China has witnessed drastic economic growth since the 1980s. Both demographic imbalance and economic boom take the same increasing time trend. Thereby, the question arises whether the increase in brideprices could simply be an income effect. To this end, I further control for the province-level per capita GDP and disposable income varying with years, as recorded in various issues of the *China Statistical Yearbook* from 1960 to 2010. I match the income measures with the children's year of marriage.

The estimation results with income as additional control are set out in Table 2.4. The coefficient estimates in columns (2) and (5) suggest that a 0.01 increase in sex ratio is associated with about a 0.26-0.27 percentage points increase in the incidence of brideprices, and the coefficients are both significant at the one percent level. According to the estimates in columns (3) and (6), sex imbalance does not affect the incidence of dowries. Overall, the results with economic growth as confounders confirm the baseline results as shown in Table 2.1. In addition, it seems that economic growth is negatively associated with marriage payments in China. This is in line with the anecdotal evidence summarized in Anderson (2007a), that marriage payments seem to disappear with industrialization and modernization.

Sample Restriction to Non-migrant Children

In the sample, 27 percent of the children reside outside of the birth village at the time of interview. However, it could not be inferred from the data whether the change of residence took place before or after the marriage. Thus, those 27 percent of the children might get married after the migration and thus might be exposed to

a different demographic structure. In particular, there would be measurement errors with the sex imbalance if they migrate out of the birth city. I restrict the sample to the children who did not migrate out of the current village and check the robustness of the effect. The results are presented Table 2.A.3 in the Appendix (Section 2.8). I find a similar effect size of demographic shifts on the incidence of brideprices and no effect on dowries. The results indicate that the baseline effect is not driven by children who migrated out of the birth village and might be exposed to another demographic structure.

One Child Policy and the Ultra-sound Technology In addition, it is well documented that rising male surplus is driven by the family planning policy (Ebenstein, 2010) and the availability of ultra-sound technology (Chen et al., 2013; Almond et al., 2019). It is likely that areas exposed to different policy intensity and technological advancement might have underlying characteristics that correlate with the development of marriage payments directly. For example, the availability of ultra-sound technology might be a sign of local wealthiness and openness, and thus the local residents might have more resources to pay marriage payments. I further control for the fine rates of the One Child Policy as calculated in Ebenstein (2010) and the availability of ultra-sound technology (Chen et al., 2013) in the birth year and province as additional confounders, respectively. If the temporal variation of city-level sex ratios is driven by such policy exposures, then the effect size or significance should be reduced once the measures of policy exposure are controlled for. However, according to the results set out in Table 2.5, the effect of sex ratios on the incidence of brideprices remains the same in size and significance as the baseline results in Table 2.1. The evidence indicates that the demographic effect on marriage payments are not threatened by potential endogeneity driven by policy exposure and technological advancement.

Marriage Year Fixed Effects The tradition of marriage formation might as well depend on the timing of the marriage. I additionally control for the marriage year fixed effects and the estimation results are set out in Table 2.6. The demographic effect on the incidence of brideprices (in column (2)) is of a larger size and significant at 1 percent level after partialing out the marriage year characteristics. In column (6), where I additionally adjust the economic development, One Child Policy intensity, and marriage year fixed effects, we could still observe the positive demographic effect on brideprices, as in Table 2.1.

Robustness of the Demographic Effect on Marriage Payment Value

In addition, I check the robustness of the demographic effect on the value of marriage payments, and the results are set out in Table 2.A.4 in the Appendix (Sec-

tion 2.8). The positive effect of male surplus on the value of brideprices is robust to controlling for economic development, intensity of One Child Policy, availability of ultra-sound technology, and marriage year fixed effects. However, such an effect is not observed when restricting the sample to those children who did not migrate out of the birth village.

Falsification Test

Finally, I conduct a falsification test to check whether the effect of demographic imbalance on brideprices is driven by pre-sample time trends. This would happen if individuals in cities where sex imbalance became higher were more likely to receive marriage payments from parents even before the shift took place. To this end, I match the marriage payment record of those born between 1960 and 1990 with sex ratios in 1970-2000, and in 1965-1995, respectively, that is, ten and five years following the birth. The demographic imbalance of future cohorts should not be correlated with the marriage payment of current ones unless there were some long-run common trends. The results are reported in Table 2.7. The estimated coefficients are insignificant with very large standard errors in all specifications, suggesting that there is no relationship between the marriage payments of current cohorts and future demographic shifts. This demonstrates that the positive relationship between sex imbalance and the incidence of brideprices identified earlier is not driven by long-run trends.

2.4.3 Heterogeneity

The previous sections show that brideprices respond to the demographic shocks in the marriage market but dowries do not. These findings indicate that brideprices might be paid out to improve the male status in the marriage competition. If that is the case, then the demographic effect should be stronger among males with low attractiveness. Moreover, since the marriage payments are usually paid out by parents, the family structure might have an interplay in the allocation of family resources as well. In this section, I utilize natal family characteristics, including socio-economic status, and family structures, to investigate from which sub-groups the demographic effect of marriage payments is driven. I use a full set of controls as specified in equation (2.1) as well as an income measure proxied by the natural logarithm of per capita GDP to rule out the income effect in all specifications.

Socio-economic Status - Marriage Payments as Status Good

It is well documented in the assortative matching literature that both males and females have an incentive of marrying "up" the socio-economic ladder (Abramitzky

et al., 2011). In the Chinese case, female scarcity is of disproportional disadvantage for males with or born in relatively low levels of socio-economic status. It is also noted in Huang and Zhou (2015) that male surplus leads to a higher unmarried rate especially for those with low socio-economic status. If we consider one of the functions of brideprices to be a status good to increase male attractiveness in the marriage market, we should expect that the demographic effect on brideprices is more pronounced among males with lower socio-economic status at the time of marriage.

Educational attainments and wealth are the usual measures for the individual socio-economic status. Unfortunately, it is difficult to trace down the wealth of males at the time of their marriage using the cross-sectional data collected in 2013. However, educational attainment usually remains unchanged after the marriage formation. Thus, I use the highest educational attainment of the parents and that of the children to study how the effect changes among those from families characterized by or with different levels of educational attainments. Parents (or children) of relatively low education are defined as those with lower than median educational attainments among parents with children (or children) within the same ten-year age cohort. Compulsory schooling laws came into effect in China in the 1980s (Fang et al., 2012), and there is a huge expansion of educational institutions since the late 1990s. Thus, the availability of educational resource is uncomparable among those born in different decades from 1960 to 1990. If low educational attainment is defined as having no or primary schooling, then most of those born between 1960 to 1980 would fall into that category. Thus, the heterogeneity here would simply be due to differences of birth cohorts, instead of the relative educational attainment among all the potential competitors.

The estimation results are set out in Table 2.8. Columns (1)-(4) present the heterogeneous effects according to parents' educational attainments and columns (5)-(8) according to children's. From columns (1)-(2) and (5)-(6) it could be seen clearly that male surplus only causes higher incidence of brideprices among males with relatively low levels of education or born to parents with low education, but not those with a high educational background. The coefficient estimates from columns (3)-(4) and (7)-(8) indicate that the incidence of dowries does not seem to be affected by demographic shifts regardless of the socio-economic background.

Family Structure

The structure of the natal family could have a vital influence on the allocation of family resources among children (Barcellos et al., 2014; Lei et al., 2017). For example, big families might be financially more constrained, and hence those parents are less likely to adjust the marriage payments according to marriage market dynamics.

Moreover, in families with many sons, parents might not respond to sex imbalance either. On the one hand, they might be financially less flexible, since they would need to prepare brideprices for many sons. On the other hand, sons with many brothers are more homogeneous among siblings, and thus their parents might be less likely to allocate more resources to one specific son in light of a marriage market shock.

I first examine whether parents with different family sizes would respond to sex ratios faced by children differently. In columns (1)-(2) and (5)-(6) in Table 2.9, I classify natal families into big and small sizes according to the total number of children. Families with a higher than median headcount are classified as big size, and others are classified as small size. The empirical results show that a rise in male surplus increases the incidence of brideprices only among sons from natal families of small size. This suggests that parents with many children would hardly adjust resources allocated to sons when they face a marriage squeeze.

I then check the potential heterogeneity by whether daughters exhibit a majority among the natal family than sons. The results in columns (3)-(4) show that such a positive demographic effect on the incidence of brideprices only exists in families with more daughters than sons.

Heterogeneous Effect on the Value of Marriage Payments Similarly, the coefficients in Table 2.A.5 in the Appendix (Section 2.8) indicate that parents characterized by low education, a big family size, and having a majority of daughters would increase the value of brideprices when there is a rise in male surplus faced by sons. No significant effect on the value of dowries is observed across all specifications.

Other Heterogeneities Lastly, I also investigate the potential heterogeneous effect depending on other characteristics, including child's place of residence (rural or urban areas), birth cohort (after or before 1975), whether the child was the first-born or last-born in the family, and the sex of the first birth and last birth in the natal family. The results are set out in Table 2.A.6, 2.A.7, and 2.A.8. The only heterogeneity observed is that the demographic effect on the incidence of brideprices is more pronounced in urban than rural areas. However, this difference is not significant at 5 percent level.

2.5 Motivations of Marriage Payments in China

The previous analyses suggest that the rising sex imbalance only causes more prevalent and higher brideprices but has no effect on dowries. Then why do we observe an increase in dowries along with the rise of brideprices? In this section, I attempt to

study the different functions of marriage payments in the Chinese society. Specifically, I analyze the intergenerational and intra-household role of both kinds of marriage payments.

2.5.1 The Intergenerational Role

Firstly, I investigate the intergenerational role of marriage payments using the CHARLS 2013 survey data. I use a similar econometric specification as the main model, which is shown in equation (2.1). Specifically, it is specified as follows:

$$Y_{iht} = \alpha + \beta Payment_i + \theta Payment_i \times Son_i + \lambda \mathbf{X}_i + \mu_h + \zeta_t + \epsilon_{iht} \tag{2.3}$$

The dependent variable on the left-hand side is the outcome of interest, including intergenerational transfer, help and contact between child i and parent h. On the right hand side, the main explanatory variable, $Payment_i$ indicates whether the parents of child i have paid marriage payments or not, or the value of the payment. Son_i is a dummy, taking the value one if the child is a son, otherwise zero. The interaction term, $Payment_i \times Son_i$, is included to identify the differential correlative effect of paying brideprices relative to dowries. Thus, the coefficient estimates of interest, β and θ, measure the association between dowries and the intergenerational outcomes, and the additional association of brideprices on top of dowries respectively. Similar to equation (2.1), $\mathbf{X}_i$ contains a set of variables indicating the child characteristics. μ_z is the natal family fixed effects, ζ_t is the birth year fixed effects. The robust standard error, ϵ_{iht}, is clustered within cities.

Marriage payment is likely to be an implicit intergenerational contract if support to elder parents is only observed from those children whose parents have paid brideprices or dowries. The support could either take the form of monetary or goods transfers, or in the form of physical care-giving and contact.

To begin with, I examine whether children with marriage payments would give more resources to parents when they are old. I specify several measures of intergenerational transfer as outcome variables, including transfer amounts from child to parents, from parents to child, and the net amount from child to parents. To reduce recall bias, CHARLS records both the regular transfers and the transfer amount from the previous year during the interview. Both measurements are considered in the analysis. I take the natural logarithm of the monetary values plus one as the outcome variable in all specifications except the net transfer amounts, and use the monetary

value adjusted for inflation in the last two specifications where net transfer amount is specified as the outcome variable.

The estimation results are set out in Table 2.10. I do not find any evidence that marriage payments are paid out in exchange for old-age monetary help. On the contrary, those children who have received marriage payments from parents would give parents less (columns (1)-(2)) and receive more from them (columns (3)-(4)), compared to siblings without marriage payments by parents. The coefficients of the interaction terms in columns (5) and (6) suggest that sons who received brideprices, relative to daughters who received dowries, would give smaller amounts to parents in the net monetary exchange. This suggests that brideprices are more often paid out as a transaction price in the marriage market, but are not associated with old age monetary support.

On the other hand, providing physical care and having frequent contact is another perspective of old-age care other than giving parents monetary transfers. I examine whether marriage payments are associated with intergenerational care and contact using a linear probability model as specified by equation (2.3). The outcomes include the incidence of parents getting essential help from the child, whether the child provides help regarding housework and money management respectively, whether parents expect to receive help from the child in the future, and whether they see the child or have contact with the child at least once a month.

Table 2.11 presents the role of marriage payment in intergenerational care and contact. The incidence of marriage payments is only positively associated with the incidence of getting essential help from children, but not with other care or contact. In addition, a higher value of dowry is positively associated with the incidence of getting help from children regarding money management and expecting future help. Moreover, the coefficients of the interaction term between marriage payments and the son dummy are mostly negatively significant. The net effect of paying brideprices is insignificant or negative, suggesting that dowries carry intergenerational functions, but brideprices do not.

2.5.2 The Intra-household Role

In this section, I further explore the intra-household role of dowries. Although CHARLS 2013 recorded detailed parent-child pairwise information, one of the biggest drawbacks of the dataset is that there is no information collected regarding characteristics of the child's partner, post-marriage perception or status in the new household. If dowries

could improve the daughter's welfare in the new family, parents might be incentivized to allocate more resource to daughters upon marriage. To check the intra-household motivation of marriage payments, I use an alternative household dataset, the 2014 wave of the China Family Panel Studies (CFPS), which includes dowry values paid by female's parents upon her marriage and allows for identification of partner's socio-economic characteristics. CFPS is conducted by the Institute of Social Science Survey at Peking University biannually since 2010. It is available for four years, and the 2014 wave covers a total of 29 provinces and municipalities. Dowry payment information is available in the waves conducted in 2014 and 2016, but only the 2014 wave includes the identifier of the partner. This allows me to study the role of dowries in intra-household bargaining power. There are two main drawbacks with CFPS. First, although I could construct wife-husband pairwise records with detailed socio-economic information on each side, there is no information on the brideprice payments from the husband side. However, dowries in China might be financed as a return portion of brideprices and it is hard to study the role of sex imbalance on dowries without the information on brideprices. Second, it is difficult to identify the role of parents' socio-economic status on dowries, because the data only tracks down very few parents who still reside with daughters. Thus, I use the 2014 wave of the CFPS to conduct the correlative analysis and investigate the intra-household role of dowries on post-marriage female welfare.

Specifically, the following model is used for the analysis.

$$Outcome_{ijhp} = \alpha + \beta Dowry_i + \lambda \mathbf{X}_i + \delta \mathbf{Y}_j + \theta \mathbf{Z}_h + \mu_p + \epsilon_{ijhp} \tag{2.4}$$

The dependent variable includes several measures for intra-household female welfare, for example, the time spent on housework and work, as well as her perception of marriage of female i married to partner j residing in household h and province p. On the right hand side, the main explanatory variable, $Dowry_i$, measures the dowry payment from her parents. These include the incidence or the value of the dowry. The main coefficient of interest, β, thus measures the average effect of dowries on intra-household female status. Female welfare might be affected by her socio-economic status and her husband's. Thus, I control for a set of characteristics, including age, education level, and her working status ($\mathbf{X}_i$) as well as her husband's ($\mathbf{Y}_j$). At the household level, household per capita net income and the status of house ownership are adjusted ($\mathbf{Z}_h$), as these might affect female labor supply and general perceptions. Province fixed effects, μ_p, absorb the aggregated-level observable and unobservable

characteristics. Thus the estimation of β exploits within-province variation of dowry payments.

The estimation results are set out in Table 2.12. In panel A, where the incidence of dowries is the main explanatory variable of interest, we only find significant evidence that females with dowries are less likely to be satisfied with their partner regarding contribution to housework. However, in panel B, where the dowry value is examined, we see clearly that a higher dowry value improves female welfare in the new household. Higher dowry value is negatively associated with female time allocated on household chores, but positively associated with work time. Similarly, females having a higher dowry value are less likely to be satisfied with their partner's contribution to housework, although they spend less time on housework.

In conclusion, I find evidence that higher dowry is associated with higher female welfare in the new household, which might be one of the incentives for parents to decide to allocate more resource to daughters upon their marriage.

2.6 Conclusion

In this paper, I investigate whether male surplus contributes to the rising marriage payment phenomenon in China. I calculate sex ratios projected from the census data as a proxy for demographic imbalance, and combine this with individual-level data on marriage payment from the 2013 survey of CHARLS. The empirical strategy exploits the within-family and cross-cohort variations.

The results show that male surplus induced a rise in the incidence and value of brideprices, but had no effect on dowries. The effect was found dominantly for males with or from families characterized by low socio-economic status, those with many siblings, and where the natal family has more daughters than sons. Further investigations show that dowries carry strong intergenerational functions for future care or help in the old age, but brideprices do not. In addition, higher dowries are associated with higher female welfare, which might incentivize parents to pay dowries to daughters at time of their marriage. Taken together, these findings suggest that brideprices differ from dowries in China. The former functions as status goods in marriage market formation, while the latter carries strong intergenerational and intra-household implications.

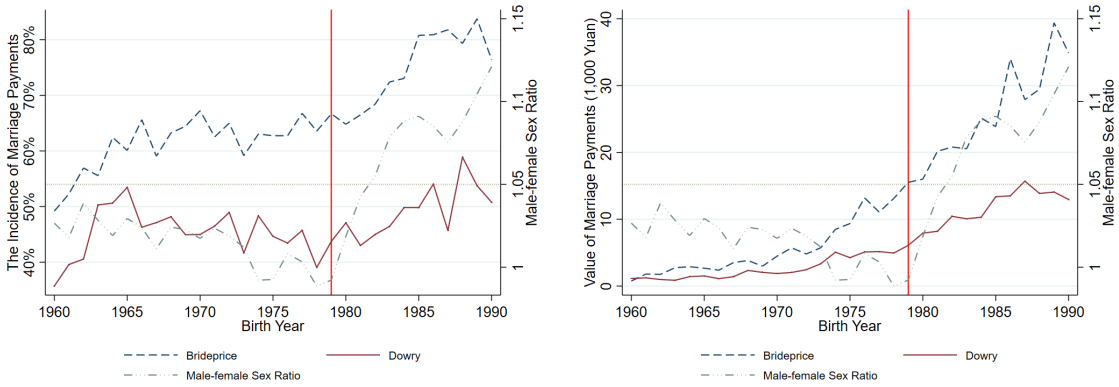
This paper provides the first empirical evidence showing that the male surplus contributed to the rise of brideprices in China. Son preference still persists in the society, and there is need of social programs based in schools and workplaces where

harmful gender norms should be eliminated to prevent further sex discrimination and selection. Moreover, this paper reveals the different significance of brideprices and dowries in the Chinese society, which rationalizes their co-existence from an empirical perspective. The interpretations of the two marriage payments need to be considered in further development of theoretical models.

There are also limitations of the dataset used in this study. It is impossible to identify who receives the marriage payments (child or child's in-laws), and there is no information on whether and how much the child's in-laws have paid. There is need of household data including these missing information to analyze the dynamics and determinants of marriage payments more precisely.

2.7 Chapter 2 - Figures and Tables

Figure 2.1: Sex Imbalance and Marriage Payments: Cohort Born 1960-1990



(a) Incidence (b) Value (CPI adjusted)

Data Source: CHARLS (2013) and census data (2000). Author's own calculation.

Table 2.1: Sex Imbalance and the Incidence of Marriage Payments

	Pooled		Brideprice		Dowry	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Without Survey Weights						
Residential Sex Ratio	0.098*	0.142***	0.261***	0.258***	0.010	0.009
	(0.055)	(0.051)	(0.093)	(0.091)	(0.083)	(0.081)
Observations	13,927	13,927	5,053	5,053	4,798	4,798
R^2	0.647	0.677	0.803	0.806	0.814	0.816
Panel B: With Survey Weights						
Residential Sex Ratio	0.099*	0.138***	0.262***	0.267***	0.034	0.026
	(0.058)	(0.052)	(0.091)	(0.088)	(0.087)	(0.086)
Observations	13,660	13,660	4,947	4,947	4,717	4,717
R^2	0.655	0.682	0.802	0.806	0.817	0.819
Child Controls		✓		✓		✓
Parent FE	✓	✓	✓	✓	✓	✓
Birth Year FE	✓	✓	✓	✓	✓	✓

Notes: The table reports the OLS estimation results of Equation (2.1) using a linear probability model. The dependent variable is a dummy, taking the value one if parent paid marriage payment when child was married, otherwise zero. Panel A does not take into account survey weights, and panel B adjusts for survey weights. Child controls include sex, age, education level, order of birth, order of birth among the children with the same sex, birth interval between the child and the next older child (zero if the child is the oldest child), age of parent when the child was born, a dummy indicating whether the child is biological child of the parents. Standard errors are clustered at the city level in all specifications.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors in parentheses.

Table 2.2: Sex Imbalance and the Value of Marriage Payments

	Value (1,000 Yuan)		ln Value		% Disp Income	
	(1)	(2)	(3)	(4)	(5)	(6)
Residential Sex Ratio	3.499	2.549	0.209	0.124	0.063	0.413
	(2.957)	(2.813)	(0.232)	(0.230)	(0.774)	(0.785)
Residential Sex Ratio × Son	5.581**	6.015**	-0.331	-0.293	0.382	0.295
	(2.744)	(2.738)	(0.250)	(0.250)	(0.735)	(0.722)
Son	1.158	0.188	1.122***	1.036***	1.052	1.333*
	(2.789)	(2.751)	(0.266)	(0.265)	(0.783)	(0.771)
ln GDP p.c.		3.862***		0.343***		-1.380***
		(0.616)		(0.037)		(0.240)
Observations	6,194	6,194	6,194	6,194	6,160	6,160
R^2	0.744	0.750	0.785	0.790	0.601	0.612
Child Controls	✓	✓	✓	✓	✓	✓
Parent FE	✓	✓	✓	✓	✓	✓
Birth Year FE	✓	✓	✓	✓	✓	✓

Notes: The table reports the OLS estimation results of Equation (2.2) without survey weight adjustment. The dependent variable is the absolute monetary value of marriage payment a parent has paid for child's marriage adjusted by province-year varying consumer price index in all specifications. Child controls include sex, age, education level, order of birth, order of birth among the children with the same sex, birth interval between the child and the next older child (zero if the child is the oldest child), age of parent when the child was born, a dummy indicating whether the child is biological child of the parents. Standard errors are clustered at the city level in all specifications.
*** $p < 0.01$, ** $p < 0.05$, *$p < 0.1$. Standard errors in parentheses.

Table 2.3: Sex Imbalance and the Incidence of Marriage Payments - Alternative Measurements

	Pooled	Brideprice	Dowry	Pooled	Brideprice	Dowry
	(1)	(2)	(3)	(4)	(5)	(6)
Residential Sex Ratio (1990)	0.125***	0.208***	0.054			
	(0.036)	(0.070)	(0.063)			
Birth Sex Ratio				0.145***	0.247***	0.016
				(0.047)	(0.086)	(0.074)
Observations	13,927	5,053	4,798	13,927	5,053	4,798
R^2	0.678	0.807	0.816	0.678	0.806	0.816
Child Controls	✓	✓	✓	✓	✓	✓
Parent FE	✓	✓	✓	✓	✓	✓
Birth Year FE	✓	✓	✓	✓	✓	✓

Notes: The table reports the OLS estimation results of Equation (2.1) using a linear probability model without survey weight adjustment. Child controls include sex, age, education level, order of birth, order of birth among the children with the same sex, birth interval between the child and the next older child (zero if the child is the oldest child), age of parent when the child was born, a dummy indicating whether the child is biological child of the parents. Standard errors are clustered at the city level in all specifications.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors in parentheses.

Table 2.4: Sex Imbalance and the Incidence of Marriage Payments - Economic Development as Confounders

	Pooled	Brideprice	Dowry	Pooled	Brideprice	Dowry
	(1)	(2)	(3)	(4)	(5)	(6)
Residential Sex Ratio	0.147***	0.272***	0.020	0.142***	0.261***	0.020
	(0.051)	(0.092)	(0.080)	(0.052)	(0.091)	(0.081)
ln GDP p.c.	-0.025**	-0.027*	-0.038**			
	(0.009)	(0.014)	(0.015)			
ln Disposable Income p.c.				-0.033***	-0.028*	-0.049**
				(0.012)	(0.017)	(0.020)
Observations	13,921	5,049	4,798	13,856	5,014	4,773
R^2	0.678	0.807	0.816	0.678	0.807	0.816
Child Controls	✓	✓	✓	✓	✓	✓
Parent FE	✓	✓	✓	✓	✓	✓
Birth Year FE	✓	✓	✓	✓	✓	✓

Notes: The table reports the OLS estimation results of Equation (2.1) using a linear probability model without survey weight adjustment. The dependent variable is a dummy, taking the value one if parent paid marriage payment when child was married, otherwise zero. Child controls include sex, age, education level, order of birth, order of birth among the children with the same sex, birth interval between the child and the next older child (zero if the child is the oldest child), age of parent when the child was born, a dummy indicating whether the child is biological child of the parents. Standard errors are clustered at the city level in all specifications.
*** $p < 0.01$, ** $p < 0.05$, *$p < 0.1$. Standard errors in parentheses.

Table 2.5: Sex Imbalance and the Incidence of Marriage Payments - Family Planning Policies and Availability of Ultra-sound Technology as Confounders

	One Child Policy			Ultra-sound Technology		
	Pooled	Brideprice	Dowry	Pooled	Brideprice	Dowry
	(1)	(2)	(3)	(4)	(5)	(6)
Residential Sex Ratio	0.142***	0.256***	0.012	0.153***	0.276***	-0.049
	(0.051)	(0.091)	(0.081)	(0.054)	(0.097)	(0.090)
Fine in years of income	0.002	-0.029	0.033			
	(0.026)	(0.054)	(0.039)			
Ultra-sound tech available				-0.024	0.011	-0.054*
				(0.017)	(0.026)	(0.028)
Child Controls	✓	✓	✓	✓	✓	✓
Parent FE	✓	✓	✓	✓	✓	✓
Birth Year FE	✓	✓	✓	✓	✓	✓
Observations	13,927	5,053	4,798	11,867	4,336	4,042
R^2	0.677	0.806	0.816	0.673	0.806	0.811

Notes: The table reports the OLS estimation results of Equation (2.1) using a linear probability model. The dependent variable is a dummy, taking the value one if parent paid marriage payment when child was married, otherwise zero. Child controls include sex, age, education level, order of birth, order of birth among the children with the same sex, birth interval between the child and the next older child (zero if the child is the oldest child), age of parent when the child was born, a dummy indicating whether the child is biological child of the parents. Standard errors are clustered at the city level in all specifications.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors in parentheses.

Table 2.6: Sex Imbalance and the Incidence of Marriage Payments - Additional Marriage Year Fixed Effects

	Pooled	Brideprice	Dowry	Pooled	Brideprice	Dowry
	(1)	(2)	(3)	(4)	(5)	(6)
Residential Sex Ratio	0.145**	0.318***	-0.033	0.131**	0.312***	-0.048
	(0.066)	(0.112)	(0.108)	(0.065)	(0.110)	(0.109)
ln GDP p.c.				0.076**	0.032	0.064
				(0.035)	(0.056)	(0.058)
Fine in years of income				-0.001	0.025	0.021
				(0.026)	(0.063)	(0.051)
Observations	10,059	3,602	3,133	10,057	3,600	3,133
R^2	0.682	0.812	0.812	0.682	0.812	0.812
Child Controls	✓	✓	✓	✓	✓	✓
Parent FE	✓	✓	✓	✓	✓	✓
Birth Year FE	✓	✓	✓	✓	✓	✓
Marriage Year FE	✓	✓	✓	✓	✓	✓

Notes: The table reports the estimation results of Equation (2.1) using logistic models and the estimated marginal effects without survey weight adjustment. The dependent variable is a dummy, taking the value one if parent paid marriage payment when child was married, otherwise zero. Child controls include sex, age, education level, order of birth, order of birth among the children with the same sex, birth interval between the child and the next older child (zero if the child is the oldest child), age of parent when the child was born, a dummy indicating whether the child is biological child of the parents. Standard errors are clustered at the city level in all specifications.
*** $p < 0.01$, ** $p < 0.05$, *$p < 0.1$. Standard errors in parentheses.

Table 2.7: Falsification Test

Dep. var:	Incidence (=1)			Value (1,000 Yuan)		
Marriage Payments (1960-1990)	Pooled	Brideprice	Dowry	Pooled	Brideprice	Dowry
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Linked with sex ratios with a 10-year lead						
Residential Sex Ratio (1970-2000)	-0.002	0.081	0.098	-2.544	-13.298	3.416
	(0.091)	(0.104)	(0.129)	(5.283)	(9.166)	(6.596)
Observations	6,681	2,111	2,062	2,768	1,052	706
R^2	0.716	0.845	0.826	0.786	0.862	0.778
Panel B: Linked with sex ratios with a 5-year lead						
Residential Sex Ratio (1965-1995)	-0.101	-0.140	0.064	-1.579	4.488	5.338
	(0.072)	(0.100)	(0.119)	(5.423)	(5.903)	(3.370)
Observations	10,491	3,569	3,520	4,523	1,776	1,237
R^2	0.698	0.815	0.817	0.736	0.834	0.762
Child Controls	✓	✓	✓	✓	✓	✓
Parent FE	✓	✓	✓	✓	✓	✓
Birth Year FE	✓	✓	✓	✓	✓	✓

Notes: The table reports the OLS estimation results of Equation (2.1) without survey weight adjustment. The dependent variable in columns (1)-(3) is a dummy, taking the value one if parent paid marriage payment when child was married, otherwise zero. The dependent variable in columns (4)-(6) is the value of marriage payments, measured in 1,000 Chinese Yuan. Child controls include sex, age, education level, order of birth, order of birth among the children with the same sex, birth interval between the child and the next older child (zero if the child is the oldest child), age of parent when the child was born, a dummy indicating whether the child is biological child of the parents. Standard errors are clustered at the city level in all specifications.
*** $p < 0.01$, ** $p < 0.05$, $^{*}p < 0.1$. Standard errors in parentheses.

Table 2.8: Sex Imbalance and the Incidence of Marriage Payments - The Role of Socio-economic Status

	Parent Education				Child Education			
	Brideprice		Dowry		Brideprice		Dowry	
	Low (1)	High (2)	Low (3)	High (4)	Low (5)	High (6)	Low (7)	High (8)
Residential Sex Ratio	0.352***	-0.044	0.055	-0.150	0.338***	0.174	0.030	-0.055
	(0.119)	(0.171)	(0.096)	(0.153)	(0.118)	(0.195)	(0.084)	(0.193)
ln GDP p.c.	-0.031*	-0.031	-0.027	-0.054**	-0.021	-0.060*	-0.046**	-0.044
	(0.016)	(0.022)	(0.018)	(0.026)	(0.016)	(0.036)	(0.019)	(0.040)
Observations	3,128	1,652	2,981	1,615	3,001	976	3,336	688
R^2	0.818	0.804	0.824	0.822	0.829	0.822	0.826	0.836
Child Controls	✓	✓	✓	✓	✓	✓	✓	✓
Parent FE	✓	✓	✓	✓	✓	✓	✓	✓
Birth Year FE	✓	✓	✓	✓	✓	✓	✓	✓

Notes: The table reports the OLS estimation results of Equation (2.1) using a linear probability model without survey weight adjustment. The dependent variable is a dummy, taking the value one if parent paid marriage payment when child was married, otherwise zero. Parents with high education means that parents' education level is above the median education level attained among all parents with children born in the same ten-year cohort (1960-1969, 1970-1979, and 1980-1990). Similarly, child with high education means that his or her education level is above the median education level attained among all the children born in the same ten-year cohort. Child controls include sex, age, education level, order of birth, order of birth among the children with the same sex, birth interval between the child and the next older child (zero if the child is the oldest child), age of parent when the child was born, a dummy indicating whether the child is biological child of the parents. Standard errors are clustered at the city level in all specifications.
*** $p < 0.01$, ** $p < 0.05$, *$p < 0.1$. Standard errors in parentheses.

Table 2.9: Sex Imbalance and the Incidence of Marriage Payments - Family Structure

	Brideprice				Dowry			
	Family Size		Majority Daughters		Family Size		Majority Daughters	
	Small (1)	Big (2)	Yes (3)	No (4)	Small (5)	Big (6)	Yes (7)	No (8)
Residential Sex Ratio	0.259**	0.235	0.542***	0.138	0.017	0.030	-0.001	0.392
	(0.126)	(0.148)	(0.160)	(0.112)	(0.136)	(0.089)	(0.079)	(0.533)
ln GDP p.c.	-0.027*	-0.022	-0.025	-0.025	-0.059***	-0.003	-0.040***	0.080
	(0.016)	(0.024)	(0.025)	(0.015)	(0.016)	(0.021)	(0.015)	(0.064)
Observations	3,507	1,539	1,468	3,581	2,974	1,824	4,585	204
R^2	0.816	0.790	0.834	0.798	0.816	0.822	0.816	0.859
Child Controls	✓	✓	✓	✓	✓	✓	✓	✓
Parent FE	✓	✓	✓	✓	✓	✓	✓	✓
Birth Year FE	✓	✓	✓	✓	✓	✓	✓	✓

Notes: The table reports the OLS estimation results of Equation (2.1) using a linear probability model without survey weight adjustment. The dependent variable is a dummy, taking the value one if parent paid marriage payment when child was married, otherwise zero. Sample is split into big or small family according to the number of children in the natal family until 2013. Sample is also split according to whether more than half of the children are female in the natal family. Child controls include sex, age, education level, order of birth, order of birth among the children with the same sex, birth interval between the child and the next older child (zero if the child is the oldest child), age of parent when the child was born, a dummy indicating whether the child is biological child of the parents. Standard errors are clustered at the city level in all specifications.
*** $p < 0.01$, ** $p < 0.05$, *$p < 0.1$. Standard errors in parentheses.

Table 2.10: Marriage Payments and Inter-generational Monetary Transfers

Dependent Variable:	ln (Monetary Amounts + 1)				Monetary Amounts	
	Received From Child		Given To Child		Net From Child	
	Last Year (1)	Regularly (2)	Last Year (3)	Regularly (4)	Last Year (5)	Regularly (6)
Panel A: The Incidence of Marriage Payment						
Paid Marriage Payment	-0.235**	-0.131*	0.127*	0.026	340.427	445.960
	(0.110)	(0.077)	(0.073)	(0.036)	(504.267)	(446.917)
Paid Marriage Payment × Son	-0.086	-0.029	0.007	-0.009	-1587.389**	-1123.136*
	(0.124)	(0.087)	(0.096)	(0.047)	(755.294)	(605.052)
Son	-0.400***	-0.032	0.309***	0.081**	61.098	221.443
	(0.100)	(0.066)	(0.070)	(0.033)	(402.319)	(225.222)
Observations	10,934	10,917	11,099	11,082	10,901	10,880
R^2	0.681	0.776	0.649	0.608	0.401	0.219
Panel B: The Value of Marriage Payment						
Payment Value (1,000 Yuan)	0.007	0.004	0.003	-0.007	456.778	403.891
	(0.008)	(0.006)	(0.008)	(0.006)	(332.958)	(336.833)
Value × Son	-0.002	-0.004	0.011	0.005	-648.848**	-349.504
	(0.007)	(0.005)	(0.008)	(0.005)	(303.820)	(291.896)
Son	-0.552***	-0.010	0.177*	0.067	2025.255*	-34.833
	(0.118)	(0.099)	(0.105)	(0.056)	(1062.040)	(579.061)
Observations	4,774	4,769	4,822	4,815	4,757	4,747
R^2	0.702	0.813	0.686	0.648	0.423	0.221
Child Controls	✓	✓	✓	✓	✓	✓
Parent FE	✓	✓	✓	✓	✓	✓
Birth Year FE	✓	✓	✓	✓	✓	✓

Notes: The table reports the OLS estimation results of Equation (2.3) without survey weight adjustment. The dependent variable in columns (1)-(4) is the logarithm of the monetary amount of inter-generational transfer plus one between parents and children recorded in 2013, and the dependent variable in column (5)-(6) is the monetary amount of net transfer from child to parent. Child controls include sex, age, education level, order of birth, order of birth among the children with the same sex, birth interval between the child and the next older child (zero if the child is the oldest child), age of parent when the child was born, a dummy indicating whether the child is biological child of the parents, and seven fixed effects regarding child residence status (in the same household but economically dependent, in the same household but economically independent, same or adjacent courtyard, another household but in the same neighbourhood, another neighbourhood in the same district, other districts, abroad). Standard errors are clustered at the city level in all specifications.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors in parentheses.

Table 2.11: Marriage Payments and Inter-generational Help and Contact

Dependent Variable: The Incidence of Inter-generational Help and Contact						
	Essential Help	Household Chores	Manage Money	Future Help	See Monthly	Contact Monthly
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: The Incidence of Marriage Payment						
Paid Marriage Payment	0.006**	-0.001	0.003	-0.020	0.006	0.007
	(0.003)	(0.006)	(0.003)	(0.013)	(0.017)	(0.016)
Paid Marriage Payment × Son	-0.009**	-0.014	-0.011**	0.017	-0.023	-0.031
	(0.004)	(0.009)	(0.005)	(0.017)	(0.023)	(0.024)
Son	0.006*	0.031***	0.011***	0.089***	0.014	-0.001
	(0.003)	(0.007)	(0.004)	(0.014)	(0.016)	(0.021)
Observations	13,589	13,616	13,616	13,616	10,731	6,207
R^2	0.602	0.560	0.404	0.823	0.654	0.711
Panel B: The Value of Marriage Payment						
Payment Value (100,000 Yuan)	0.007	0.044	0.018*	0.165**	-0.042	0.038
	(0.010)	(0.032)	(0.010)	(0.083)	(0.114)	(0.082)
Value × Son	-0.006	-0.080**	-0.025**	-0.137*	0.012	-0.018
	(0.009)	(0.032)	(0.013)	(0.079)	(0.108)	(0.093)
Son	-0.001	0.023***	0.003	0.117***	-0.017	-0.059***
	(0.003)	(0.007)	(0.003)	(0.016)	(0.018)	(0.020)
Observations	6,092	6,094	6,094	6,094	4,672	2,494
R^2	0.660	0.630	0.426	0.840	0.683	0.739
Child Controls	✓	✓	✓	✓	✓	✓
Parent FE	✓	✓	✓	✓	✓	✓
Birth Year FE	✓	✓	✓	✓	✓	✓

Notes: The table reports the OLS estimation results of Equation (2.3) using a linear probability model without survey weight adjustment. The dependent variable is a dummy, indicating whether parents have received help from children regarding essential needs, household chores, or money management, or whether they expect to get help, or whether they have frequent contact with children in 2013. Child controls include sex, age, education level, order of birth, order of birth among the children with the same sex, birth interval between the child and the next older child (zero if the child is the oldest child), age of parent when the child was born, a dummy indicating whether the child is biological child of the parents, and seven fixed effects regarding child residence status (in the same household but economically dependent, in the same household but economically independent, same or adjacent courtyard, another household but in the same neighbourhood, another neighbourhood in the same district, other districts, abroad).

*** $p < 0.01$, ** $p < 0.05$, *$p < 0.1$. Standard errors in parentheses.

Table 2.12: Dowry and Female Bargaining Power - Evidence from CFPS 2014

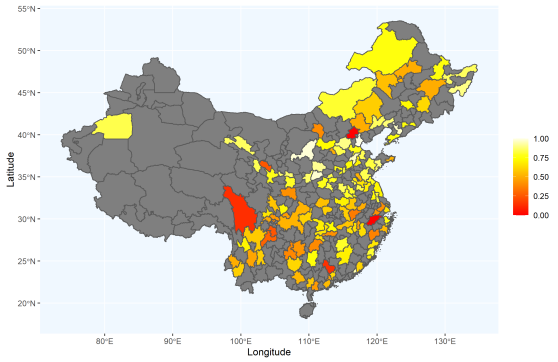
	Time Use			Marriage Satisfaction (=1)		
	Housework Weekdays	Housework Weekend	Work	Overall	Economic Contribution	Housework Contribution
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A						
Incidence of Dowry	0.042	0.042	-0.028	-0.014	-0.013	-0.035***
	(0.045)	(0.045)	(0.097)	(0.012)	(0.012)	(0.012)
Observations	9,481	9,481	6,170	9,464	9,461	9,464
R^2	0.082	0.082	0.085	0.028	0.030	0.038
Panel B						
ln Dowry Value	-0.035**	-0.035**	0.075**	-0.002	0.002	-0.014***
	(0.018)	(0.018)	(0.038)	(0.005)	(0.005)	(0.005)
Observations	6,596	6,596	4,519	6,580	6,579	6,582
R^2	0.099	0.099	0.087	0.024	0.026	0.036
Female Controls	✓	✓	✓	✓	✓	✓
Partner Controls	✓	✓	✓	✓	✓	✓
Household Controls	✓	✓	✓	✓	✓	✓
Province FE	✓	✓	✓	✓	✓	✓

Notes: The table reports the OLS estimation results. The dependent variable is a continuous measure of hours spent on housework from column (1)-(2) and hours spent on work in column (3). The depend variable from column (4) to (6) is a dummy, taking the value one if the female respondent reports to be very satisfied with marriage overall, regarding partner's economic contributions, or household contribution, otherwise zero. Female controls and partner controls include age, education level, and working status; household control include household per capita net income and the status of house ownership.

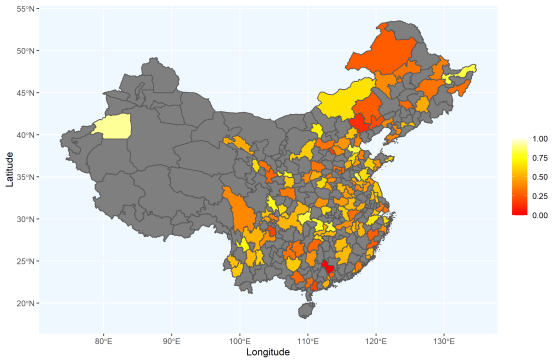
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors in parentheses.

2.8 Chapter 2 - Appendix

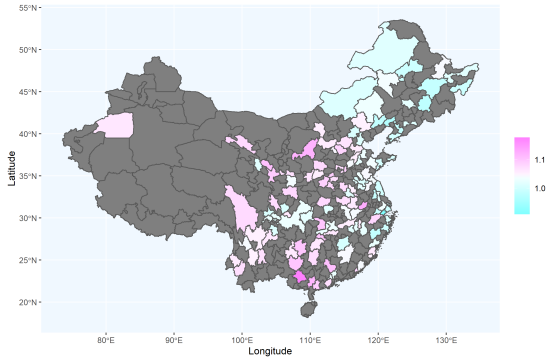
Figure 2.A.1: The Average Prevalence of Brideprices and Dowries, and the Average Sex Ratios Across Prefectual Cities in Mainland China



(a) Incidence of Brideprices



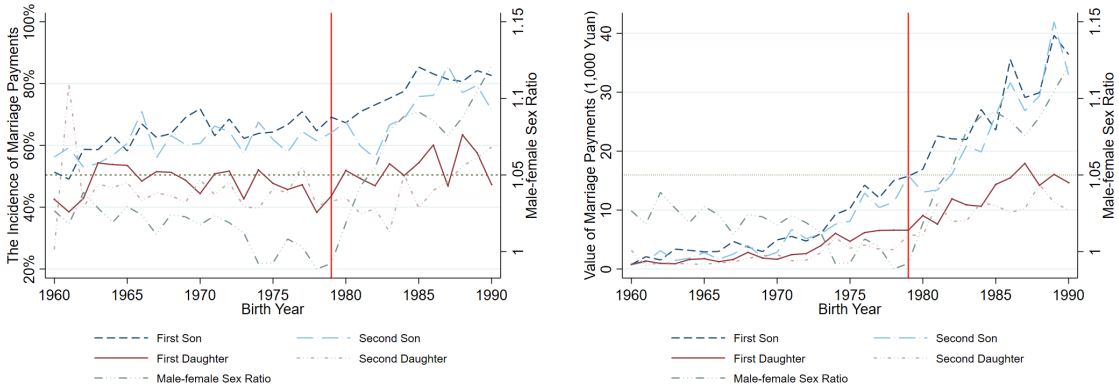
(b) Incidence of Dowries



(c) Average Sex Ratios

Data Source: CHARLS (2013) and census data (2000). Author's own calculation.

Figure 2.A.2: Sex Imbalance and Marriage Payments By Birth Order: Cohort Born 1960-1990



(a) Incidence (b) Value (CPI adjusted)

Data Source: CHARLS (2013) and census data (2000). Author's own calculation.

Figure 2.A.3: An Example of the Construction of Sex Ratios

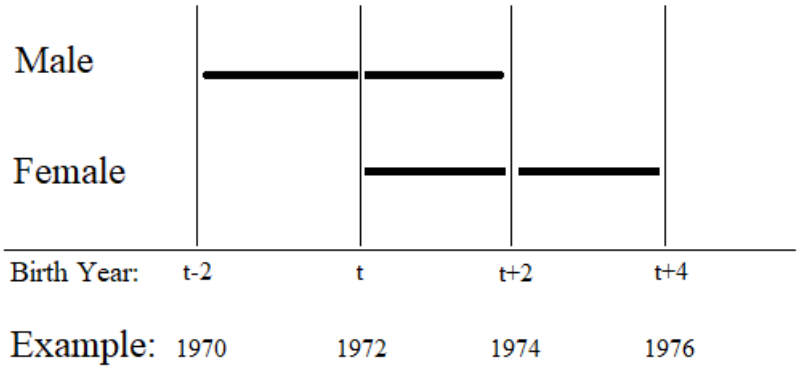
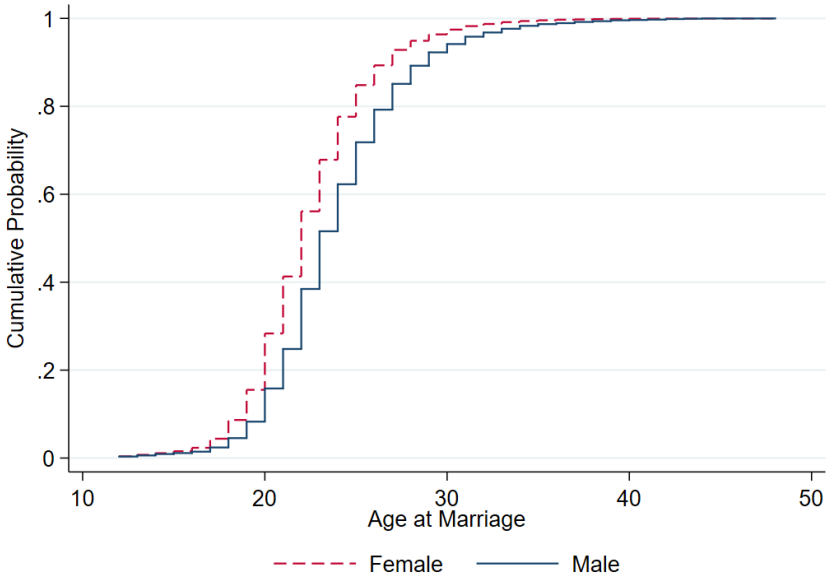


Figure 2.A.4: Cummulative Probability of Age at Marriage among Males and Females



Data Source: CHARLS (2013). Author's own calculation.

Table 2.A.1: Summary Statistics

	mean	sd	min	max
Sex Imbalance				
Residential Sex Ratio	1.05	0.14	0.66	1.73
Residential Sex Ratio (1990)	1.08	0.15	0.68	1.95
Birth Sex Ratio	1.05	0.15	0.69	1.74
Marriage Payments				
Marriage Payments (=1)	0.56	0.50	0.00	1.00
Value of Marriage Payments (CPI adjusted)	9,290.82	17,022.40	0.00	283,055.22
ln Marriage Payment Value + 1	8.02	1.75	0.00	12.55
Share as Per Capita Annual Disposable Income	2.32	4.16	0.00	149.68
Child Characteristics				
Male (=1)	0.51	0.50	0.00	1.00
Birth Year	1,975.17	7.27	1,960.00	1,990.00
Age	37.83	7.27	23.00	53.00
Educational Attainment				
Less than primary school (=1)	0.17	0.36	0.00	1.00
Primary to middle school (=1)	0.63	0.48	0.00	1.00
High school or vocational education (=1)	0.16	0.37	0.00	1.00
Bachelor or above (=1)	0.04	0.19	0.00	1.00
Biological Child (=1)	0.98	0.13	0.00	1.00
Birth Order	2.36	1.32	1.00	10.00
Birth Order Within the Same Sex	1.65	0.88	1.00	9.00
Birth Interval (years)	2.18	2.30	0.00	18.00
Age Different with Parent (years)	26.62	5.40	16.00	68.00
Parent Characteristics				
Age	63.46	8.33	41.00	91.00
Educational Attainment				
Less than primary school (=1)	0.43	0.63	0.00	1.00
Primary school (=1)	0.27	0.44	0.00	1.00
Some high school (=1)	0.23	0.42	0.00	1.00
Vocational education (=1)	0.07	0.26	0.00	1.00
Bachelor or above (=1)	0.00	0.06	0.00	1.00
Household Wealth Quintile				
1 (poorest)	0.23	0.42	0.00	1.00
2	0.24	0.42	0.00	1.00
3	0.21	0.40	0.00	1.00
4	0.17	0.38	0.00	1.00
5 (richest)	0.15	0.36	0.00	1.00
Urban (=1)	0.30	0.46	0.00	1.00
Number of Children	3.76	1.46	2.00	10.00
Income Measures				
ln GDP p.c.	8.53	1.16	5.16	11.44
ln Annual Disposable Income p.c.	8.03	0.97	4.89	10.30
Observations	13,927			

Table 2.A.2: Sex Imbalance and the Value of Marriage Payments

	Absolute Monetary Amounts					
	Pooled	Brideprice	Dowry	Pooled	Brideprice	Dowry
	(1)	(2)	(3)	(4)	(5)	(6)
Residential Sex Ratio	6082.269**	-4438.697	-343.596	5336.309**	-7259.203	-530.454
	(2806.918)	(5562.292)	(2102.235)	(2645.614)	(5418.002)	(2070.570)
ln GDP Per Capita				3846.445***	4741.393***	657.484*
				(613.573)	(731.060)	(387.205)
Observations	6,194	2,580	1,737	6,194	2,580	1,737
R^2	0.744	0.818	0.813	0.749	0.825	0.813
Child Controls	✓	✓	✓	✓	✓	✓
Parent FE	✓	✓	✓	✓	✓	✓
Birth Year FE	✓	✓	✓	✓	✓	✓

Notes: The table reports the OLS estimation results of Equation (2.1) without survey weight adjustment. The dependent variable is the absolute monetary value of marriage payment a parent has paid for child's marriage adjusted by province-year varying consumer price index in all specifications. Child controls include sex, age, education level, order of birth, order of birth among the children with the same sex, birth interval between the child and the next older child (zero if the child is the oldest child), age of parent when the child was born, a dummy indicating whether the child is biological child of the parents. Standard errors are clustered at the city level in all specifications.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors in parentheses.

Table 2.A.3: Sex Imbalance and the Incidence of Marriage Payments - Non-migrant Children Sample

	Pooled	Brideprice	Dowry
	(1)	(2)	(3)
Residential Sex Ratio	0.180***	0.237***	0.074
	(0.055)	(0.079)	(0.095)
Child Controls	✓	✓	✓
Parent FE	✓	✓	✓
Birth Year FE	✓	✓	✓
Observations	9,318	3,354	3,050
R^2	0.710	0.840	0.849

Notes: The table reports the OLS estimation results of Equation (2.1) using a linear probability model. The dependent variable is a dummy, taking the value one if parent paid marriage payment when child was married, otherwise zero. Child controls include sex, age, education level, order of birth, order of birth among the children with the same sex, birth interval between the child and the next older child (zero if the child is the oldest child), age of parent when the child was born, a dummy indicating whether the child is biological child of the parents. Standard errors are clustered at the city level in all specifications.

*** $p < 0.01$, ** $p < 0.05$, *$p < 0.1$. Standard errors in parentheses.

Table 2.A.4: Sex Imbalance and the Value of Marriage Payments - Robustness Checks

	(1)	(2)	(3)	(4)	(5)
Residential Sex Ratio	1541.279	2657.947	1849.213	-528.743	-463.736
	(2835.210)	(2785.666)	(2106.115)	(3469.859)	(3440.404)
Residential Sex Ratio × Son	2487.780	5961.585**	6327.571**	9512.624***	9492.740***
	(3039.401)	(2748.268)	(2462.003)	(3242.724)	(3255.793)
Son	3155.456	235.443	-892.915	-2724.945	-2707.666
	(3156.335)	(2760.791)	(2507.661)	(3226.428)	(3237.869)
ln GDP p.c.	3872.577***	3849.193***	2784.676***	4860.647***	4833.989***
	(591.685)	(616.884)	(391.937)	(1386.007)	(1405.365)
Fine in years of income		1401.743			634.108
		(1821.359)			(2163.774)
Ultra-sound Tech Available			-177.212		
			(744.532)		
Observations	4,225	6,194	5,363	4,836	4,836
R^2	0.777	0.750	0.770	0.761	0.761
Child Controls	✓	✓	✓	✓	✓
Parent FE	✓	✓	✓	✓	✓
Birth Year FE	✓	✓	✓	✓	✓
Marriage Year FE				✓	✓

Notes: The table reports the OLS estimation results of Equation (2.1) without survey weight adjustment. The dependent variable is the absolute monetary value of marriage payment a parent has paid for child's marriage adjusted by province-year varying consumer price index in all specifications. Compared to Table 3.A.3, column (1) is estimated based on children who have not migrated out of birth villages, columns (2)-(5) are estimated based on the full sample, with additional controls. Child controls include sex, age, education level, order of birth, order of birth among the children with the same sex, birth interval between the child and the next older child (zero if the child is the oldest child), age of parent when the child was born, a dummy indicating whether the child is biological child of the parents. Standard errors are clustered at the city level in all specifications.

*** $p < 0.01$, ** $p < 0.05$, *$p < 0.1$. Standard errors in parentheses.

Table 2.A.5: Sex Imbalance and the Value of Marriage Payments - Heterogeneity

	Heterogeneity Variable			
	Family Low Edu	Child Low Edu	Small Family	More Daughters Than Sons
	(1)	(2)	(3)	(4)
Residential Sex Ratio × Son × Var	1.817*	0.241	4.699***	3.255***
	(0.981)	(1.004)	(0.810)	(0.798)
Residential Sex Ratio × Var	3.826	-2.729	1.074	2.801
	(2.718)	(2.608)	(1.992)	(2.895)
Residential Sex Ratio × Son	2.477	5.902	4.102	4.963*
	(3.057)	(3.758)	(2.634)	(2.822)
Residential Sex Ratio	-2.323	6.065	0.868	0.219
	(3.972)	(4.686)	(2.536)	(3.388)
Son	1.164	0.000	-1.535	-1.267
	(2.670)	(3.010)	(2.908)	(2.817)
ln GDP p.c.	3.886***	3.871***	3.850***	3.812***
	(0.611)	(0.618)	(0.613)	(0.612)
Observations	6,194	6,194	6,194	6,194
R^2	0.750	0.750	0.752	0.751
Child Controls	✓	✓	✓	✓
Parent FE	✓	✓	✓	✓
Birth Year FE	✓	✓	✓	✓

Notes: The table reports the OLS estimation results of Equation (2.1) without survey weight adjustment. The dependent variable is the absolute monetary value of marriage payment (unit 1,000 yuan) a parent has paid for child's marriage adjusted by province-year varying consumer price index in all specifications. Child controls include sex, age, education level, order of birth, order of birth among the children with the same sex, birth interval between the child and the next older child (zero if the child is the oldest child), age of parent when the child was born, a dummy indicating whether the child is biological child of the parents. Standard errors are clustered at the city level in all specifications.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors in parentheses.

Table 2.A.6: Sex Imbalance and the Incidence of Marriage Payments - Urban-Rural and Cohort Heterogeneity

	Urban/Rural				Born After 1975			
	Brideprice		Dowry		Brideprice		Dowry	
	Urban	Rural	Urban	Rural	Yes	No	Yes	No
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Residential Sex Ratio	0.442**	0.209*	-0.208	0.103	0.320**	0.283*	-0.040	-0.019
	(0.184)	(0.111)	(0.208)	(0.086)	(0.139)	(0.145)	(0.163)	(0.092)
ln GDP p.c.	-0.059**	-0.007	0.017	-0.058***	-0.062**	-0.012	-0.037*	-0.028
	(0.023)	(0.016)	(0.028)	(0.017)	(0.024)	(0.018)	(0.019)	(0.026)
Observations	1,470	3,579	1,404	3,394	1,743	2,328	1,933	2,031
R^2	0.814	0.811	0.815	0.821	0.818	0.818	0.836	0.822
Child Controls	✓	✓	✓	✓	✓	✓	✓	✓
Parent FE	✓	✓	✓	✓	✓	✓	✓	✓
Birth Year FE	✓	✓	✓	✓	✓	✓	✓	✓

Notes: The table reports the OLS estimation results of Equation (2.1) using a linear probability model without survey weight adjustment. The dependent variable is a dummy, taking the value one if parent paid marriage payment when child was married, otherwise zero. Child controls include sex, age, education level, order of birth, order of birth among the children with the same sex, birth interval between the child and the next older child (zero if the child is the oldest child), age of parent when the child was born, a dummy indicating whether the child is biological child of the parents. Standard errors are clustered at the city level in all specifications.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors in parentheses.

Table 2.A.7: Sex Imbalance and the Incidence of Marriage Payments - Heterogeneity by Birth Order

	First-born		Last-born	
	Brideprice	Dowry	Brideprice	Dowry
	(1)	(2)	(3)	(4)
Residential Sex Ratio	0.276***	0.010	0.358***	0.035
	(0.104)	(0.088)	(0.122)	(0.107)
Residential Sex Ratio × First-born (=1)	-0.010	0.012		
	(0.099)	(0.093)		
First-born (=1)	0.017	0.013		
	(0.109)	(0.101)		
Residential Sex Ratio × Last-born (=1)			-0.062	-0.002
			(0.114)	(0.114)
Last-born (=1)			0.050	0.003
			(0.124)	(0.119)
ln GDP p.c.	-0.027*	-0.037**	-0.036**	-0.053***
	(0.014)	(0.015)	(0.015)	(0.017)
Observations	5,049	4,798	4,296	3,922
R^2	0.807	0.816	0.825	0.832
Child Controls	✓	✓	✓	✓
Parent FE	✓	✓	✓	✓
Birth Year FE	✓	✓	✓	✓

Notes: The table reports the OLS estimation results of Equation (2.1) using a linear probability model without survey weight adjustment. The dependent variable is a dummy, taking the value one if parent paid marriage payment when child was married, otherwise zero. Child controls include sex, age, education level, order of birth, order of birth among the children with the same sex, birth interval between the child and the next older child (zero if the child is the oldest child), age of parent when the child was born, a dummy indicating whether the child is biological child of the parents. Standard errors are clustered at the city level in all specifications.

*** $p < 0.01$, ** $p < 0.05$, *$p < 0.1$. Standard errors in parentheses.

Table 2.A.8: Sex Imbalance and the Incidence of Marriage Payments - The Sex of First-borns and Last-borns

	Brideprice				Dowry			
	First Birth		Last Birth		First Birth		Last Birth	
	Son	D.	Son	D.	Son	D.	Son	D.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Residential Sex Ratio	0.243**	0.387**	0.264**	0.317*	0.152	-0.015	0.205*	-0.059
	(0.110)	(0.174)	(0.101)	(0.178)	(0.208)	(0.089)	(0.123)	(0.105)
ln GDP p.c.	-0.022	-0.047*	-0.031*	-0.016	-0.064*	-0.033**	-0.046	-0.039**
	(0.015)	(0.026)	(0.016)	(0.031)	(0.037)	(0.016)	(0.032)	(0.017)
Observations	3,732	1,317	3,867	1,182	1,069	3,729	1,596	3,202
R^2	0.799	0.835	0.801	0.839	0.850	0.809	0.822	0.819
Child Controls	✓	✓	✓	✓	✓	✓	✓	✓
Parent FE	✓	✓	✓	✓	✓	✓	✓	✓
Birth Year FE	✓	✓	✓	✓	✓	✓	✓	✓

Notes: The table reports the OLS estimation results of Equation (2.1) using a linear probability model without survey weight adjustment. The dependent variable is a dummy, taking the value one if parent paid marriage payment when child was married, otherwise zero. D. stands for daughter. Child controls include sex, age, education level, order of birth, order of birth among the children with the same sex, birth interval between the child and the next older child (zero if the child is the oldest child), age of parent when the child was born, a dummy indicating whether the child is biological child of the parents. Standard errors are clustered at the city level in all specifications. *** $p < 0.01$, ** $p < 0.05$, *$p < 0.1$. Standard errors in parentheses.

Chapter 3

The Health Consequence of Rising Housing Prices in China*

with:
Feicheng Wang

* This article uses data from China Health and Nutrition Survey (CHNS). We thank the National Institute of Nutrition and Food Safety, China Center for Disease Control and Prevention, Carolina Population Center, the University of North Carolina at Chapel Hill, the NIH (R01-HD30880, DK056350, and R01-HD38700) and the Fogarty International Center, NIH for financial support for the CHNS data collection and analysis files from 1989 to 2006, and the China-Japan Friendship Hospital as well as Ministry of Health for additional support for the CHNS data collection from 2009 onwards. Xu acknowledges funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – project RTG 1723. We are grateful to Krisztina Kis-Katos, Stephan Klasen, Lukas Menkhoff, Helke Seitz, Lingwei Wu, Qiang Wu, and seminar and conference participants at the University of Göttingen, Tsinghua University, Nanjing University, University of Clérmont Auvergne, the Eastern Economic Association Meeting 2020 in Boston, the Development Economic Network Berlin (DENeB) discussion rounds in Berlin, the European Economic Association Virtual Congress 2020, the Annual Meeting of the German Economic Association (VfS) 2020, the Spring Meeting of Young Economists 2021, DIAL Conference, and DGGÖ PhD Workshop, for valuable discussions and comments. We would also like to thank Oliver Braun and Matthew Rudh for their excellent assistance in proofreading the paper. All remaining errors are ours. Declarations of interest: none.

Abstract

This paper examines the health consequence of rising housing prices in China by exploiting spatial and temporal variation in housing price appreciation linked to individual-level health data from 2000 to 2011. Using an instrumental variable approach, we find that increases in housing prices significantly raise the probability of having chronic diseases. The deep-rooted marriage culture that males are obligated to buy a home before getting married and the growing marriage market competition are the main channels that exacerbate the negative health effects, particularly for parents with marriage-age sons. We also show that increased work intensity, high levels of stress, and changes in lifestyle are other possible channels. Our results highlight the unintended health consequences of the real estate market prosperity.

JEL Codes: I10, I12, R21, R31, G51
Keywords: Housing Prices; Chronic Diseases; Health; Marriage Competition; China

3.1 Introduction

Many countries have experienced considerable housing price growth in the past decades. China provides a good example of the real estate market prosperity, with housing prices growing at an annual rate of eight percent since the 2000s,[1] a higher rate than both average household disposable income and GDP (Glaeser et al., 2017). China's first-tier cities have witnessed an even faster growth, with an average annual rate of 13 percent from 2003 to 2013 (Fang et al., 2016). Rapid housing appreciation has affected individuals and households significantly, leading to tightened household consumption (Waxman et al., 2020), lower female labor force participation (Fu et al., 2016), postponed marriage (Wrenn et al., 2019), and more inter-generational co-residence (Li and Wu, 2019), among other consequences. This paper focuses on a less studied dimension, i.e. the health consequences of the rapid growth in housing prices, aiming to contribute to a deeper understanding of the effects of high housing prices in both China and other countries.

The Chinese housing market was centrally regulated until the late 1990s, when China enacted sweeping reforms privatizing the housing sector. Since then, the real estate market has experienced a remarkable boom. Despite rapid housing appreciation, house ownership rates have been growing steadily, reaching over 90 percent in 2004, as shown in Figure 3.A.1 in Appendix A (Section 3.8.1). Such high home ownership rates originate from Chinese social norms, in which owning a home is preferred to renting. Despite the high ownership rates, this preference for home ownership has still imposed heavy financial burdens on individuals and households with the surging housing prices. This is possible for house owners if they intend to purchase a more spacious home. This is not uncommon because houses in China are important status goods that signal wealth and social status (Wei et al., 2017) and individuals have strong incentives to pursue bigger houses[2]. The other group who is especially vulnerable to the rapid housing price growth consists of young unmarried men and their parents who wish to

[1] Author's own calculation based on housing price data (1998-2013) obtained from various issues of China Statistical Yearbook. Also see Figure 3.1.

[2] In the Chinese context, those households owning a house could hardly benefit from the housing price appreciation. On contrary, they could still be affected by the rising prices due to three reasons. Firstly, the apartments are usually purchased for living, and only very few households would own extra residence for investment or to rent out. When the real estate prices increase, the rise in the book value of the wealth could hardly be translated into cash inflow. Secondly, the upgrade of housing is rapid in China. Even if the households already owns a house or an apartment, they would still have the motivation of purchasing a bigger and newer residence. Thirdly, those parents, especially those with marriage-age sons, would have the pressure of preparing a dwelling space for the children.

purchase a house to increase their competitiveness in the marriage market (Wei et al., 2017). This stems from a strongly-rooted marriage culture in China in which men are obliged to obtain a house before marriage, and this makes the strain of needing to buy a home particularly sever for young men on the marriage market (Wrenn et al., 2019).

Increases in housing expenditure and the financial stress induced by it may directly affect individual's mental health, which could subsequently deteriorate physical health. One may work more intensely to alleviate their financial burdens when housing prices are high, leading to both mental and physical health problems. As a response to this increased work intensity, individuals may get less sleep or reduce their leisure time, resulting in a higher probability of chronic diseases. In this paper, we explore the health consequences of rising housing prices in China, in particular its impact on chronic diseases, by exploiting the temporal and spatial variation in housing market boom since the late 1990s.

We consider the prevalence of chronic diseases to be the main health outcome of interest in this paper. This is mainly due to the concern that individuals' physical health is less likely to react immediately and intensely to changes in housing prices, and more likely to respond gradually and to be reflected by chronic diseases. In the last few decades, China has witnessed a significant increase in the prevalence of chronic diseases. As shown in Figure 3.1, about 16 percent of individuals had chronic diseases in the early 2000s; by 2011 this number had almost doubled. The share of those with hypertension shows a similar trend. The increased prevalence of chronic diseases could be induced by a wide range of factors, including an unhealthy diet, lack of physical activities, tobacco use, excessive alcohol intake, constant mental stress, and environmental pollution, among other things. In this paper, we propose that the rapid rise in housing prices could be an additional driving force.

We link individual health records with housing prices at the province level to explore whether individuals in areas experiencing greater housing price growth are more likely to acquire chronic diseases. Individual-level data is collected from the China Health Nutrition Survey (CHNS) between 2000 and 2011, when China experienced considerable housing price appreciation. The main advantage of this database lies in the fact that it reports rich information on an individual's health conditions and health-related activities, such as diagnoses of over 16 types of chronic diseases as well as biomarkers of blood pressure, lifestyle and health-related behaviors, including consumption of tobacco and alcohol, sleep time, sports activities, etc. These allow us to precisely and comprehensively measure individuals' health conditions and explore

the potential mechanisms behind changes in health outcomes. The panel nature of the data allows for further controls of unobserved time-invariant confounders by including individual fixed effects, which could substantially reduce potential omitted variable biases. To identify a causal link between housing prices and health, we employ an instrumental variable approach. We construct the instrumental variable for housing prices by combining exogenous variations in both initial land supply across provinces and national interest rates over the years. The former captures cross-province housing supply capacities and the latter tackles the national trend in housing demand. Our instrumental variable implies that a province initially more constrained by land supply would witness a higher growth in housing prices when confronting a positive housing demand shock nationwide.

Our empirical results show several interesting findings. We find robust evidence that individuals from provinces that experienced more rapid growth in housing prices were more likely to have chronic diseases and hypertension relative to others. Such a negative health impact holds when we use alternative measures of housing prices and health outcomes, an alternative instrumental variable, and when we control for additional possible confounding factors.

The negative health consequence of housing prices is closely linked to the deeply rooted marriage culture that males are obliged to provide a home for newly formed households and the increasing marriage market competition among young males in China due to the imbalanced sex ratio. This is reflected in our findings that the negative health effects are observed only for parents with at least one marriage-age son during the housing boom, but are not significant for those without sons, those having only old sons who are most likely to be married, or sons who are still too young to worry about marriage market competition. This indicates an inter-generational effect of housing prices on health arising from marriage market competition among marriage-age males. Our results also indicate that increasing working activities among the employed, higher levels of stress, and changes in lifestyle, especially rises in tobacco consumption and declines in sleep time among the younger cohort could be other potential channels that shape the health consequences of housing prices.

Our paper contributes to several strands of literature. First, it adds to an emerging set of studies that document the socio-economic impact of rising housing prices. At the individual and household level, rising housing prices are shown to, for example, induce higher saving rates (Chamon and Prasad, 2010), lead to declines in consumption (Waxman et al., 2020), increase wages (Liang et al., 2016), delay marriages (Wrenn et al., 2019), increase inter-generational co-residence (Li and Wu, 2019), and discour-

age female labor force participation and raise fertility among house owners (Dettling and Kearney, 2014; Fu et al., 2016). Our paper is among the first to examine the health consequences of rising housing prices in an emerging economy context and enriches this strand of literature by providing additional evidence on the health impact of the housing market prosperity.

Second, this paper complements the broad literature that documents the health effects of socio-economic changes, such as globalization and pollution. Fan et al. (2020) find negative health effects of input tariff reductions following China's WTO accession, through increased working hours. Export expansion is shown to increase infant mortality due to deterioration of environmental quality (Bombardini and Li, 2020). He et al. (2020) show that air pollution due to straw burning significantly increases mortality. The findings in this paper underline the role of an important albeit less researched socio-economic factor - housing prices - in causing health problems. Our findings also speak directly to the literature studying the health effects of financial stress (e.g. Sweet et al., 2013; McInerney et al., 2013; Prentice et al., 2017; Guariglia et al., 2021).

This paper is also related to the growing literature studying the interplay between housing prices and the marriage market, and particularly marriage competition, in China. Due to the imbalanced sex ratio (Wei and Zhang, 2011a), competition for a bride between unmarried young men in the marriage market has risen substantively. To improve competitiveness, households with an unmarried son tend to buy a more expensive home to signal wealth, which has pushed up housing prices (Wei et al., 2017). The high housing prices in turn further intensify the marriage market competition, imposing much pressure on the young adults and their family and subsequently delaying marriages (Wrenn et al., 2019). Our paper shows that marriage market competition is not only a driver of housing value appreciation, but also interacts with housing prices and leads to negative health consequences.

The rest of the paper proceeds as follows. Section 3.2 describes the main data sources used in the paper and discusses measures of housing prices and health outcomes. Section 3.3 presents our empirical approach and identification strategies. Section 3.4 reports empirical results, including the main results and a battery of robustness checks. Section 3.5 examines several potential mechanisms, and Section 3.6 concludes.

3.2 Data Sources and Measures

Our primary dataset is the China Health and Nutrition Survey (CHNS). It is an individual-level longitudinal survey conducted by the Population Center at the University of North Carolina at Chapel Hill since 1989.[3] It covers nine provinces for earlier waves, and extends to 15 in 2011. The provinces are selected to ensure that the sample is representative for China's eastern, central, and western regions. For consistency concerns, we consider only the nine provinces surveyed in all waves in our analysis. The sample within each province is drawn following a multi-stage, random cluster sampling approach. The raw dataset includes around 3,456 to 7,319 households and 11,860 to 20,914 individuals across years. An important feature of the dataset is that it reports rich information on individuals' health status, sickness history, and health-related behaviors, such as diet, tobacco consumption, alcohol intake and physical activities, etc. This allows for a precise measure of individual's health condition, as well as potential behaviors associated with health status. It also records detailed information on individual's demographic characteristics, labor market participation, and household characteristics, which enables us to control for an extensive number of confounders that may affect individuals' health. Compared to other individual-level databases in China that are often repeated cross-sections, the longitudinal nature of the CHNS makes it possible to track individuals across time and hence control for unobserved time-invariant characteristics by including individual fixed effects in the estimation.

The CHNS conducted ten waves until 2015, which well covered the period of China's housing market boom as well as a long period before the boom.[4] For the purpose of this paper, we mainly use the five waves ranging between 2000 and 2011. The choice of the sample is based on two concerns. First, the Chinese housing market was not liberalized until the end of the 1990s and the data on housing prices is only available since 1998. As we link the CHNS data with housing prices, our sample could only start from 2000. The first four waves prior to 2000 are, however, used for testing the existence of a pre-trend in the empirical analysis. Second, all waves except 2015 contain detailed information on three common types of chronic diseases - hypertension, diabetes, and cardiovascular diseases (including heart diseases and stroke) - whereas the 2015 wave reports only information on cancer. We therefore exclude 2015 from our main analysis in order to have consistent measures of the outcome variable. The

[3] More details about the CHNS could be found from https://www.cpc.unc.edu/projects/china.
[4] The ten waves are for the years of 1989, 1991, 1993, 1997, 2000, 2004, 2006, 2009, 2011, and 2015.

2015 wave, however, additionally surveyed individuals' mental health status. We use this cross-sectional data to investigate the impact of housing prices on mental health, a potential channel through which housing prices affect chronic diseases.

We link the CHNS data with records of housing prices at the province level using individuals' province location.[5] Our province-level data on housing prices is from various issues of the *China Real Estate Statistics Yearbook*, in which housing prices are measured in Chinese Yuan per square meter. The *China Real Estate Statistics Yearbook* reports the average selling price for each type of commercialized buildings by use, including all residential houses, villas and deluxe apartments, office and business buildings, as well as the average price of all types of commercial buildings. We use the natural log of average residential housing prices as our main measure of housing prices as they affect the entire population, while other types may only affect certain groups of individuals. We consider the price of business and office buildings and the average price level of all types of buildings as alternative measures to check the robustness of our results.

We restrict our sample to working-age population aged 15 to 60. The final sample includes 32,111 observations of 9,515 individuals from nine provinces. Table 3.A.1 in Appendix A (Section 3.8.1) presents descriptive statistics of our key variables. The average housing price is 2,691.23 Yuan (approximately 384 US dollar) per square meter, and the average residential housing price is slightly lower at 2,505.53 Yuan (approximately 358 US dollar) per square meter. A typical Chinese house with the size of 90 square meters could cost 32,000 US dollars. This is over 16 times the annual income of the average person in our sample. In contrast, the median price of a house in a major US metropolitan area is about 162,000 US dollars, which is less than 3.5 times the average annual income (Dettling and Kearney, 2014). These numbers imply that housing expenditures impose a much higher financial burden on Chinese than American households.

The main measure of our health outcome is the prevalence of chronic diseases. This arises from the concern that changes in housing prices may not induce health problems immediately, but rather individuals may accumulate health problems gradually, which eventually may develop into chronic diseases. The identification of individuals with chronic diseases is based on either self-reported health information or on their biomarker records. An individual is identified to have a chronic disease if either hypertension, diabetes, or cardiovascular diseases is reported or detected through biomarker

[5] We get access to the province location only for each individual.

or medicine intake.[6] One concern is that diabetes and cancer are relatively recently detected chronic diseases, the prevalence of which may rise with technological changes. However, hypertension can be diagnosed at a very low cost, and has been prevalent in China since the late 1980s. Thus, we consider the prevalence of hypertension as an alternative, specific measure of the health outcome of interest. This could rule out the increase in chronic diseases due to heterogeneous technology changes in health care across provinces.[7] Table 3.A.1 shows that the overall prevalence of chronic diseases is about 18.7 percent and the prevalence of hypertension is about 14.6 percent. Compared to the U.S., where nearly half of the population suffers from at least one type of chronic disease (Raghupathi and Raghupathi, 2018), China has a much lower prevalence.

3.3 Empirical Methodology

3.3.1 Econometric Specification

The aim of this paper is to uncover the health consequences of rising housing prices. To this end, we link individuals' health outcomes from the CHNS with province-level housing prices to investigate whether individuals from provinces with a larger growth in housing prices are more likely to have health problems than others. Specifically, we consider the following linear probability model:

$$P(CD_{ihprt} = 1) = \alpha + \beta HP_{pt} + \lambda \mathbf{X}_{it} + \delta \mathbf{Y}_{ht} + \gamma \mathbf{Z}_{pt} + \mu_i + \upsilon_{rt} + \varepsilon_{ihprt} \tag{3.1}$$

where CD_{ihprt} is a dummy variable indicating whether an individual i from household h, province p that belongs to region r, has chronic diseases in year t. HP_{pt} denotes measures of housing prices for province p in year t. The coefficient of interest β measures the average effects of rising housing prices on the probability of having chronic diseases conditional on other control variables.

$\mathbf{X}_{it}$ is a vector of individual characteristics that may affect one's health conditions, including age, age squared, education level, marital status, etc. $\mathbf{Y}_{ht}$ denotes a set of household characteristics, including household income, household size, and household registration type (rural or urban). Notably, the inclusion of household income controls

[6] For example, an individual has hypertension if he or she is diagnosed with hypertension by a physician, or take medicines to reduce blood pressure, or the average systolic blood pressure is over 140 mmHg or the average diastolic blood pressure reading is over 90 mmHg.

[7] We control for local economic development and medical resources at the province level in our regressions to further capture the effects of technology advances.

for positive wealth effects of real estate appreciation, especially for house owners and real estate investors (Atalay et al., 2017), which may further affect households' consumption and eventually the health condition of family members (Aladangady, 2017; Waxman et al., 2020). This allows us to concentrate on the health effects of housing prices due to high financial pressure for home buyers or potential home buyers. We also control for a set of time-variant health-related factors at the province level, denoted by $\mathbf{Z}_{pt}$. These include the number of medical personnel per 1,000 local residents as a measure of availability of healthcare resources, the natural log of waste water emissions (in 10,000 tons) and sulfur dioxide emissions (in tons) to proxy for environmental quality, local GDP per capita to account for the level of economic development as well as unemployment rates capturing the labor market dynamics.

We include individual fixed effects, denoted by μ_i, to control for both observed and unobserved time-invariant individual characteristics. As such, the identification of the health effects derives from variations in health conditions across time at the individual level. Notice that if an individual did not change household or province location, the individual fixed effects also capture household-level and province-level time-invariant characteristics that may be related to health. These could capture systematic differences in health conditions, for example, between households with and without genetic diseases, or between individuals from the relatively more developed eastern coastal provinces and those from the less developed inland regions. However, to account for the possibility that some individuals move out of a household (e.g. due to marriage or work changes), we check the robustness of our results by additionally controlling for household fixed effects.

In equation (3.1), we also include region-year fixed effects, denoted by υ_{rt}, to control for differential trends in average health outcomes across regions (Fan et al., 2018; Bombardini and Li, 2020). Here a region denotes eastern, central or western China. This is especially important in the Chinese context due to its large geographic area and the imbalanced economic development across regions. During our sample period, the Chinese government launched various programs to stimulate and support the development of the western and central regions, such as the Great Western Development Strategy starting from 2000 and the Rise of Central China Plan from 2004. The inclusion of region-year fixed effects effectively capture regional policy changes that may affect both the real estate market and health. Finally, ε_{ihprt} is the error term. Following Cameron et al. (2011) and Deschenes et al. (2020), we use a two-way clustering approach at the individual and the province-year level. This controls for serial correlations in the error term for each individual and possible correlations between

individuals within provinces for each year.[8] We also check clustered standard errors at the individual level or at the province-year level and our results are rather similar.

3.3.2 Identification Strategy

The identification of a causal relationship between housing prices and health relies on the assumption that housing prices are uncorrelated with the error term. Such an assumption would be violated if there are unobserved factors that are simultaneously correlated with both housing prices and individuals' health conditions. One example of such factors is internal migration. Migrant inflows may increase the demand for houses, especially in regions where migrants without a local household registration are not allowed to purchase houses, thus inducing higher housing prices.[9] Meanwhile, migrant workers may have relatively poor health conditions, particularly due to the fact that they often work in occupations with poorer working conditions and earn relatively lower incomes (Meng, 2012). As such, rising inflows of migrants would lead to an overestimate of the health effects of housing prices.[10] However, without accurate records of migrants, and due to the fact that housing policies are often designed at the city level, we are not able to control for the impact of migration. Unobserved factors related to local economic development could also bias our estimates. Local housing prices may be inflated with regional economic growth, which further affects health due to, for instance, pollution. We control for local GDP per capita and pollutant emissions to capture differences in economic development across provinces. Unobserved heterogeneity in local infrastructure development and production activities could still be a concern.

To address the potential endogeneity of housing prices, we employ an instrumental variable approach. An ideal instrument should be correlated with housing prices but uncorrelated with the error term. Our identification strategy relies on exogenous

[8] Recent advances by Cameron and Miller (2015) and Abadie et al. (2017) underline that in a fixed effects model setting like our case, correlation of errors within clusters due to common shocks can be absorbed by cluster-specific fixed effects. Applying to our setting, individual (and also province) fixed effects in the estimation can address possible correlation of the error term for each individual across years as well as common shocks to individuals within provinces. Comparing clustered standard errors with heteroskedasticity-robust standard errors does not reveal significant differences.

[9] In many Chinese cities, housing policies prohibit migrants without a local household registration from purchasing houses as a tool to adjust housing prices.

[10] Immigration may also reduce housing prices, which would then cause an underestimate of the health effects of housing prices. Sá (2015) finds that immigration reduces housing demand as natives, especially high-income natives, move to other regions in the UK. This, however, rarely happens in China because of the high migration costs attributed to the hukou system.

shocks that explain variations in housing prices through affecting housing demand and supply. Specifically, our instrumental variable is constructed as the pre-sample land supply at the province level multiplied by national interest rates, with the former measuring housing supply constraints across provinces and the latter capturing national trends in housing demand.

Housing supply elasticity with respect to geography-based measures of land supply is a commonly used instrumental variable in predicting housing prices (e.g., Saiz, 2010; Mian and Sufi, 2011; Chaney et al., 2012; Stroebel and Vavra, 2019). The intuition is that regions with relatively less elastic housing supply would observe a higher growth in housing prices with a positive housing demand shock (Hilber and Vermeulen, 2016). In the spirit of this strand of literature, studies on Chinese housing prices often use land supply as the instrument since housing supply elasticity cannot be directly estimated due to data availability constraints.[11] Land use is strictly regulated in China by the central government and provincial government (Liang et al., 2016), allowing land supply to be plausibly exogenous for cities and individuals. However, this does not apply in our context, as our housing price measure is at the province level. If provincial government's decision on land supply is correlated with other policies that may affect individual's health, land supply does not meet the exclusion restriction of a valid instrumental variable. There is evidence that incomes from land sales are an important source of local government revenue (Waxman et al., 2020), which in turn determine government's expenditure on public utilities, including health related items like medical and health care, as well as environmental protection.

To alleviate the above concerns, we use pre-sample land supply in 1997 across provinces, rather than time-variant land supply, to capture the possibility that provinces with higher initial land supply constraints would witness a larger increase in housing prices than others given a housing demand shock. Figure 3.A.2a in Appendix A (Section 3.8.1) shows the scatter plots of changes in log housing prices between 2000 and 2011 and the initial land supply for our sample provinces. It shows a clear negative correlation such that provinces with a higher initial land supply, like Liaoning and Heilongjiang, observed a smaller increase in housing prices. In order to predict variations in housing prices over time, we interact pre-sample land supply with national long-term interest rates. A similar instrument is adopted by Chaney et al. (2012) and Rong et al. (2016). Long-term interest rates not only add time variations but also cap-

[11] For example, Waxman et al. (2020) and Li and Wu (2019) use land sales from the local government to real estate companies as a proxy for land supply to predict city-level housing prices and further identify the impacts on consumption and inter-generational co-residence, respectively. Liang et al. (2016) instruments housing prices with land supply quotas to examine the impact on wages.

ture a national trend in housing demand, as the mortgage rates are approximately the same as the long-term interest rates in China (Rong et al., 2016). Lower interest rates indicate easier access to mortgage, which stimulates housing demand. As changes in interest rates are designed by the central government and could hardly be affected by a single province, they are exogenous to the error term. Specifically, our instrument takes the following form:

$$IV_{pt} = ls_{p,1997} \times i_t \tag{3.2}$$

where $ls_{p,1997}$ denotes land supply of province p in 1997. We measure land supply using planned urban residential area divided by urban population. Data on planned urban residential area and urban population in 1997 is collected from *China Environment Statistical Yearbook* (1998). i_t is the average level of national long-term interest rate in year t.[12] Our instrument indicates that regions with limited land supply are disproportionately more affected by a rise in national housing demand due to a lower interest rate, and we therefore expect that our instrumental variable is negatively correlated with housing prices. Figure 3.A.2b in Appendix A (Section 3.8.1) displays the scatter plots of changes in log housing prices against changes in our instrumental variable as defined in equation (3.2) and shows a negative correlation between the two, as expected.

A remaining concern regarding our instrumental variable is that it may violate the exclusion restriction if initial land supply and/or national interest rates affect individual's health conditions through channels other than housing prices. Provinces with lower initial land supply may have more potential to develop, leading to differential health outcomes. In our regression, we include a set of control variables at the province level to capture possible health effects of economic growth, labor market conditions, medical resources, and pollution. This ensures that channels other than housing prices through which initial land supply affects health are controlled for. In other words, initial land supply is plausibly exogenous conditional on those control variables. Interacting national interest rates with the initial land supply at the province level further ensures that the differential effects of national interest rates across provinces are proportional to initial land supply constraints, thereby ensuring that our instrumental variable could only affect individual's health through housing prices.

[12] In cases where interest rates were adjusted several times within a year, we calculate a weighted average using the share of active days for each rate as weight.

In addition, we consider an alternative instrumental variable following the literature documenting spatial correlation of housing prices. Spatial diffusion of housing prices is found pervasive in many regions such that housing prices are affected by a price shock emanating from surrounding areas, e.g. in the U.S. (Guerrieri et al., 2013), the Netherlands (Teye and Ahelegbey, 2017), the UK (Holly et al., 2011), Taiwan (Chen et al., 2011), and China Mainland (Gong et al., 2016). We specifically calculate the average housing prices of all bordering provinces as an alternative instrument for the housing prices of each sample province, and we expect a positive correlation between the two.

3.4 Empirical Results

3.4.1 Main Results

Table 3.1 reports the 2SLS estimation results of equation (3.1) using initial land supply multiplied by interest rates as the instrumental variable for housing prices and the incidence of any chronic diseases as the dependent variable. We start with a baseline estimation controlling for only individual fixed effects and region-year fixed effects in column (1), and add individual, household, and province-level characteristics in columns (2) to (4), respectively. We report two-way cluster-robust standard errors at the individual and the province-year level in all specifications.

The first-stage results show that our instrumental variable is negatively correlated with housing prices and is highly significant in all specifications, indicating that provinces with a tighter initial land supply constraint would observe a larger increase in housing prices when facing a positive demand shock due to lower interest rates. The first-stage F-statistics allows us to reject the null hypothesis of weak instruments.

The second-stage results show that all estimates for the coefficient on average residential housing prices are positive and statistically significant, indicating that individuals in provinces with a larger growth rate of housing prices are more likely to have chronic diseases than those in provinces with lower growth rates. The size of the coefficients does not change much across specifications when controlling for additional variables. Our preferred specification with a full set of control variables in column (4) suggests that a 10 percent rise in housing prices leads to a 3.29 percentage point increase in the prevalence of chronic diseases, holding other variables constant. This impact is economically sizable, as it represents about 17.59 percent of the average

prevalence of chronic diseases in China during our sample period (see Table 3.A.1 in Appendix A (Section 3.8.1)).

To check whether our results in Table 3.1 are sensitive to model selection, we re-estimate equation (3.1) using logit regressions where we employ the same instrumental variable as specified in equation (3.2). Logit regressions with individual fixed effects only take into account individuals that experienced changes in health status during the sample period, thereby resulting in a reduced sample size. The results are reported in Table 3.A.2 in Appendix A (Section 3.8.1). In line with the results in Table 3.1, the estimated coefficients of residential housing prices are positive and significant in all specifications, indicating a positive effect of the rising housing prices on the incidence of chronic diseases. Our preferred results in column (4) suggest that a 10 percent increase in housing prices is linked to a 0.32 rise in log-odds of chronic diseases holding other variables constant. This translates to a 3.27 percentage point increase in the prevalence of chronic diseases for an individual who has a 10 percent probability of having chronic diseases.

3.4.2 Robustness Checks

In this section, we perform a series of robustness checks on our main results. Specifically, we check whether the negative health effects of housing prices are sensitive to specifications controlling for additional confounders and to alternative measures of housing prices and health outcomes as well as an alternative instrumental variable. We also perform a falsification test to check whether the health consequences of rising housing prices are driven by long-term trends.

3.4.2.1 Balanced Panel and Controlling for Additional Variables

In Table 3.2 we report the estimation results of equation (3.1) based on the balanced panel and controlling for additional possible confounders. In all specifications, we use the prevalence of chronic diseases as the outcome variable and average residential housing prices (in natural log) as the housing price measure, including a full array of control variables and fixed effects as in column (4) of Table 3.1. Column (1) constrains the estimation sample to individuals recorded in all waves. This reduces the sample size considerably. The coefficient estimate, however, continues to show a positive and significant impact of housing prices on the prevalence of chronic diseases.

We add household fixed effects in column (2). This arises from the concern that some individuals may change households over time, for instance, due to marriage or

divorce.[13] Individuals could face different household environments that might affect their health. Consider the case of an individual moving out of the parents' household due to marriage and creating a new one. He or she will deal with new relationships and play a different role in the new household. Unobserved household characteristics related to such a change may result in different health outcomes. The result in column (2) shows that accounting for the effects of changes in households does not alter our main results much.

In our main regression, we include unemployment rates to control for the effects of labor market conditions on health due to employment opportunities. In column (3) of Table 3.2, we use an alternative measure, namely, average wages, as a proxy for labor market conditions. Compared to unemployment rates that reflect mainly the availability of job opportunities and job insecurity, which could adversely affect both physical and mental health (Green, 2011), average wages, however, as a measure of job quality, can mitigate the negative health effects arising from limited jobs (Green, 2020). As shown in column (3), the negative coefficient on average wages, though insignificant, implies the potential mitigation effects. More importantly, the coefficient on housing prices remains positive and statistically positive. Controlling for average wages, the estimated effects of housing prices on health increase modestly compared to the baseline results in Table 3.1.

During the period under investigation, one notable change in the Chinese economy was the rapid pace of openness to the world market following China's accession to the WTO in 2001. Trade openness could affect health through various channels, such as changes in incomes, changes in job opportunities due to import competition or export expansion and the associated adjustment in work intensity, and changes in environmental quality (e.g. Fan et al., 2020; Bombardini and Li, 2020). To control for the impact of trade openness, we include the shares of export and import over local GDP in columns (4) and (5) separately. Interestingly, we do not find significant effects of either trade openness measures on health, possibly due to various channels offsetting each other. Controlling for trade openness, the impact of housing prices on chronic diseases is not affected much.

An additional concern is the negative consumption effects of housing prices in China as documented by Waxman et al. (2020). With higher housing prices, households reduce their non-housing expenditures, including expenses on medical services,

[13] 1.84 percent of individuals in the sample changed their household between 2000 and 2011. However, this change could only be observed among those who reside within the sample provinces throughout the survey years. If individuals moved to a non-sampled province, a problem of sample attrition would arise.

and this may cause negative health effects. While the CHNS data does not collect information on each household's expenditures on housing and non-housing items, we address the role of consumption by controlling for the average medical expenses per capita at the province level. Data on medical expenses is obtained from *China Statistical Yearbook*. The results in column (6) reveal that individuals from provinces with higher expenses on medical related items tend to be less likely to have chronic diseases, though the impact is statistically insignificant. Controlling for medical expenses, the effects of housing prices on the prevalence of chronic diseases remain very close to the baseline estimate. This demonstrates that changes in non-housing expenses play a limited role in affecting health.

Overall, our results in Table 3.2 show consistent evidence that rising housing prices in China increases the occurrence of chronic diseases.

3.4.2.2 Alternative Measures of Housing Prices and Health Outcomes

In our main specification, we measure housing prices using the average residential housing prices. One concern with this measure is that the average residential housing price level is an average over all types of residences, including villas and deluxe apartments. Rising prices of luxurious houses may not strongly affect health, since only a specific group of high-income individuals, who may not be as sensitive to housing price increases as average-income individuals, would be affected. While we do not have price data on residential buildings excluding luxurious houses, the *China Real Estate Statistics Yearbook* separately reports the average price level of villas and deluxe apartments. This allows us to partial out the effects of the more expensive houses by controlling for their average prices in our main regressions. Column (1) of Table 3.3 reports the 2SLS results. It shows that higher prices of villas and deluxe apartments are correlated with a lower probability of having chronic diseases and such a correlation is statistically insignificant. Partially out the potential impacts of high-grade houses, the estimated coefficient on average residential housing prices remains significantly positive but slightly larger in size compared to the result in column (4) of Table 3.1. An individual in a province with a 10 percent increase in housing prices has a 4.32 percentage point higher probability of being affected by chronic diseases.

In column (2) of Table 3.3, we consider the price of houses for business use and office buildings. The health effects of price growth of business and office buildings are less clear-cut, as some entrepreneurs and firms own buildings while others have to rent, with the former more likely to have positive health effects whereas the latter

more likely to affect health negatively. The first-stage result shows that our instrumental variable, initial land supply multiplied by national interest rates, can also well predict changes in the housing prices of business and office buildings. The second-stage coefficient estimate stays positive and significant, but smaller in size than the effect of residential housing prices. A 10 percent price increase is associated with a 1.95 percentage point rise in the incidence of chronic diseases.

China Real Estate Statistics Yearbook also reports average prices of all types of houses, including residential houses, office buildings and houses for business activities, as well as other types. Using this measure, column (3) again shows a positive association between housing prices and the incidence of chronic diseases.

One concern with the health outcome measure is that chronic diseases are self-reported. The rising prevalence of chronic diseases could be due to technological improvements allowing for cheaper, easier detection of chronic diseases, or due to an increasing awareness of health leading individuals to check their health more frequently. The inclusion of individual fixed effects and year fixed effects could partially account for time-invariant heterogeneity in health values among individuals and technological differences across provinces, as well as common trends in the development of medical equipment across years. The unobserved changes in individual's health values or province-specific technology improvement across time could still bias our estimates. To address these concerns, we consider an alternative, specific type of chronic disease – hypertension, as our measure of health outcome. The awareness of hypertension in China can be traced back to the 1980s, and detection of hypertension is easily performed at a low cost. The increasing prevalence of hypertension could therefore hardly be driven by technological improvement or rising awareness.

Columns (4) and (5) of Table 3.3 report the estimation results using the prevalence of hypertension as the outcome variable, conditional on a full set of control variables. The two columns use the natural log of average residential housing prices and the average price of all types of houses as the housing price measure, respectively. The estimated coefficients on both measures are positive and significant, indicating that individuals in provinces with a higher level of housing prices are more likely to have hypertension. Compared to the effects on chronic diseases in general, the effects on hypertension are rather similar in magnitude. This alleviates the concern of measurement errors in the outcome variable.

We further investigate the effects on other measures of health, including the incidence of being sick in the past four weeks and three indicators measuring overweight and obese. The results are set out in Table 3.A.3 in Appendix A (Section 3.8.1). The

estimated coefficients on residential housing prices are only marginally significant in column (3) in which we use an indicator of high waist circumference as the dependent variable; an increase in housing prices is associated with a higher likelihood of having high waist circumference. It is worthy mentioning that waist circumference is based on a direct measure, and is unlikely to be affected by technological advancement in health care or rising awareness. Considering that high waist circumference is found to be highly correlated with the incidence of chronic diseases (Li et al., 2007), this result provides additional supportive evidence that our main findings are not biased by measurement errors.

3.4.2.3 Alternative Instrumental Variable

Table 3.4 presents the estimation results using the average residential housing prices of adjacent provinces as the instrument. The identification of a causal relationship hinges on the assumption that housing prices are spatially correlated but the housing prices in other provinces do not directly affect individual's health. The first-stage results show a highly significant correlation between home and neighbouring housing prices. The first-stage F-statistics rules out the possibility of a weak instrumental variable problem.

The coefficient estimates on housing prices in the second stage remain positive and significant for both specifications using chronic diseases and hypertension as alternative measures of health outcomes in columns (1) and (2), respectively. Relative to the baseline results using initial land supply interacted with national interest rates as the instrumental variable, the estimated effects of housing prices on health appear smaller. A 10 percentage rise in housing prices leads to a 1.43 percentage point increase in the prevalence of chronic diseases and a 1.27 percentage point increase in the prevalence of hypertension. Considering that unobserved spatial correlations between provinces could still be a threat to the causal interpretation of the results, we rely on our main instrumental variable as specified in equation (3.2).

As an additional check of the quality of our primary instrumental variable, we perform a leave-one-out exercise by excluding one province from the sample each time. This examines whether our results are driven by a particular province, especially the outlier province (Heilongjiang) as shown in Figure 3.A.2. The results are presented in Table 3.A.4 in Appendix A (Section 3.8.1). Our results stay almost unchanged leaving out Heilongjiang and the instrumental variable can still well explain housing prices, as shown in column (1). However, excluding Guangxi or Guizhou reports a

weak instrumental variable issue in columns (8) and (9). Importantly, we continue to find negative and significant health effects of rising housing prices in all specifications. These results rule out the possibility of our main findings being driven by a particular province.

3.4.2.4 Falsification Test

We conduct a falsification exercise to check whether the negative health consequences of rising housing prices are driven by long-run time trends. This would happen if individuals in provinces that experience a relatively higher growth in housing prices are more likely to suffer from chronic diseases even before the boom of the real estate market. To this end, we match the health records of individuals from the first four waves of the CHNS collected between 1989 and 1997 with housing prices in 1999-2007 and in 2004-2012, respectively, namely, 10 and 15 years following the survey. Future housing prices should not be correlated with the prevalence of chronic diseases unless there are some long-run common trends. We repeat the estimation of equation (3.1) and the results are reported in Table 3.5, where panel A uses our main instrumental variable and panel B uses the average housing prices in neighboring provinces. The estimated coefficients of housing price measures are insignificant in all specifications. However, this could be due to the weak instrument problem in panel A and in column (2) of panel B. In column (1) of panel B where the weak instrument problem is not present, we still find an insignificant impact of housing prices on health. In sum, the results in Table 3.5 show no supportive evidence of the existence of a pre-trend.

3.5 Examining Potential Mechanisms and Additional Results

Rising housing prices may cause an increase in the prevalence of chronic diseases through various mechanisms. We show in this section that the heavily rooted marriage culture of men providing a home for newly formed households and the marriage market competition among unmarried males play a significant role in yielding negative health consequences. We also show evidence that increased work intensity, rising mental stress as well as changes in lifestyle could be other potential channels.

3.5.1 Marriage Culture and Marriage Market Competition

The imbalanced sex ratio in China caused by the long lasting son preference culture, the One Child Policy, the spread of gender identification technology, as well as the rural land reform (Ebenstein, 2010; Bulte et al., 2011; Chen et al., 2013; Almond et al., 2019; Tan et al., 2021) has substantially raised the competition between young men for a bride. To improve the standing in the marriage market competition, unmarried young men and their parents strive to raise household wealth by saving more, consuming less, and taking risky but better paid jobs (Wei and Zhang, 2011a; Waxman et al., 2020; Tan et al., 2021), and purchase more spacious and expensive homes to signal their wealth (Wei et al., 2017). The latter stems from the heavy marriage culture in China that males are obligated to provide a living space for newly formed households. Owning a house is still a pre-requisite to marrying a potential wife in many regions (Hu and Wang, 2019; Wrenn et al., 2019). High housing prices, along with increased competition in the marriage market, however, have placed much pressure on unmarried young males as a result. This makes them more likely to have health problems than their female counterparts.

The potential negative health effects of the increasing housing prices on young males could also be passed to parents, as Chinese parents often feel obliged to provide financial support for house purchasing to improve their son's success in the marriage market. This originates from a strongly-rooted culture of ensuring the continuity of the family lineage and clan. In addition, many Chinese parents, especially in the rural areas, live with their son's family and rely on them for elderly care (Tan et al., 2021). The rising housing prices therefore could particularly affect families with unmarried sons. The recent study by Tan et al. (2021) finds that parents whose son faces a higher marriage market competition are more likely to take risky jobs and suffer from work-related injuries and even deaths.

To investigate the role of marriage culture and marriage market competition in affecting health, we first examine whether housing prices affect marriage-age males and females differently. Columns (1) and (2) of Table 3.6 report the results separately by sex among the marriage-age individuals. It seems that the rising housing prices have a stronger impact on the likelihood of getting chronic disease for males than females, but the difference is statistically insignificant. The effects on marriage-age females could be related to marital sorting in the marriage market (Sun and Zhang, 2020). With males having more pricey houses, females have the pressure to improve their wealth so as to improve their chance of finding a husband with a similar or higher

wealth level. An additional potential reason is that females with a higher wealth level are more likely to have a stronger bargaining power within the family after marriage (Wei and Zhang, 2011a). As documented by Wrenn et al. (2019), China's housing price growth has delayed the marriage of males and females in a similar manner.

In columns (3) and (4), we turn to the sample of parents to better understand the potential passing-through effects of marriage market competition on them. To this end, we classify parents into those who had at least one marriage-age son aged 15 to 40 when interviewed and those who did not, and compare whether housing prices affect them differently. Parents without a marriage-age son consist of those without children, those with only daughters, those with young sons who have not yet entered the marriage market, and those having married sons. The results for the two types of parents show that only parents with at least one adult son experienced the negative health effects induced by housing prices. The estimated coefficient in column (3) suggests that a 10 percent increase in residential housing prices leads to a 4.51 percentage point increase in the incidence of chronic diseases for parents having at least one marriage-age son.[14]

Considering that males are often the breadwinner of the household and they are more likely to take risky jobs (Tan et al., 2021), the negative health effects of housing prices may be stronger for fathers. In Table 3.A.6 in Appendix A (Section 3.8.1), we replicate regressions in Table 3.A.5 separately for fathers and mothers. Interestingly, both fathers and mothers who have at least one adult son experience a higher probability of having chronic diseases with housing price growth, with the coefficient for mothers even larger, although the difference is statistically insignificant. These results indicate that mothers' health could be equally affected by housing prices, possibly through channels other than work. We explore additional channels in the next section. By contrast, neither fathers nor mothers without a marriage-age son are affected.

The CHNS also allows us to distinguish between households that had a marriage and those not in the last 12 months. Although households without a marriage can also be affected by rising housing prices if they have a marriage in the near future, we

[14] One concern with these results is that they may capture the age effects if parents with an adult son are older than others, and if the health effects of housing prices are stronger for the older cohort. In our sample, the average age for parents with at least one adult son is 46 and for other parents 40. To check this possibility, we first examine the health effects by cohort. As shown in Table 3.A.5 in Appendix A (Section 3.8.1), the negative health effects are only present for the two cohorts aged 31-45 and 46-60, and stronger for the older cohort. To further rule out the age effects, we run a regression that includes all parents and interacts housing prices with a dummy variable indicating the two types of parents, controlling for ages. This ensures that we compare parents at the same age. We find consistent result that housing prices affect only parents with at least one adult son.

nevertheless expect a more salient health effect on households with a marriage. This is exactly what we find in columns (5) and (6) that report the results separately for the two types of households. Those from both types of households experience a higher probability of having chronic diseases with housing price growth, whereas households with a marriage are more strongly affected.

In sum, the results in Table 3.6 suggest that the marriage culture and marriage market competition among marriage-age males are an important channel through which housing prices cause negative health consequences in China.

In Appendix C (Section 3.8.3), we further investigate possible heterogeneity of the results by household income level and household registration status. We find that those from low-income households are more vulnerable to the rising housing prices. Moreover, the negative health impact is not only present among urban households but also among rural residents and rural-to-urban migrants. The impact on the latter two groups could be related to the marriage culture if the sons of rural and migrant households intend to purchase a home in urban areas.

3.5.2 Other Potential Channels

3.5.2.1 Work Intensity

As a response to the rising living costs caused by higher housing prices, one may increase work intensity to improve income and alleviate financial stress. This could be reflected by longer hours with an expectation of higher payment, or by finding a second job to broaden income sources. It is documented in the literature that long working time may cause both mental and physical health problems through reducing leisure time, raising inactivity, affecting sleeping quality (e.g., Cygan-Rehm and Wunder, 2018; Wong et al., 2019). It is therefore likely that the high occurrence of chronic diseases following housing price growth is due to increased work intensity.

We examine the role of work intensity by estimating equation (3.1), replacing the outcome variable with various measures of work intensity: weekly working days, daily working hours, and total weekly working hours (all in natural logs). For this purpose, we constrain our sample to those who reported to be working in the last week. The 2SLS regression results are reported in Table 3.7.

The results reveal a positive association between housing prices and work intensity in all specifications, suggesting that individuals from provinces that experienced a higher housing price growth worked more than those from other provinces. A 10 percent rise in housing prices increases weekly working days by 6.45 percent, daily

working hours by 6.81 percent, and total weekly working hours by 13.92 percent. For an individual working an average of 5.46 days a week, 7.18 hours a day, and 39.49 hours a week, the estimated effects translate to working 0.35 more days a week, 0.49 more hours a day, and 5.50 more hours a week. These results imply that increased work intensity following housing price appreciation could be a channel that causes health problems.

3.5.2.2 Mental Stress

Mental health and physical health are closely related (Yaribeygi et al., 2017; Ohrnberger et al., 2017; Niles and O'Donovan, 2019). Individuals with poor mental health are more likely to develop physical health problems including chronic diseases. Housing prices could affect mental health directly by imposing heavy financial stress to individuals who have the need to purchase houses. This is particularly the case for the disadvantaged groups that have relatively poor financial conditions but a high demand for housing, like migrants, for instance (Li and Liu, 2018). Housing prices could also cause mental health problems indirectly. Our results in Table 3.7 show evidence that individuals work more hours when housing prices are higher. Long working hours in turn may aggravate mental stress (Wong et al., 2019). To explore the role of mental stress in shaping the effects of housing prices on physical health, we regress housing prices on measures of mental health conditions.

Unfortunately, the CHNS did not collect data on individual mental health conditions until 2015, when a new module on mental stress was added to the survey. We therefore use self-reported memory conditions as an indirect measure of mental health for the sample of year 2000-2006, and a direct measure of mental health for the 2015 cross section.[15] Specifically, the 2015 survey asked questions on the frequency of feeling stressful, frustrated, and anxious in the past month, with answers ranging from one to five, with one indicating "never" and five indicating "very often". We define individuals who often or very often experienced any of these distressing feelings as mentally stressed. The regression results based on the two samples are reported in Table 3.8.

Column (1) reports the 2SLS results based on the 2000-2006 sample. The negative coefficient indicates that a rise in housing prices is associated with memory lapses. This provides suggestive evidence that housing prices may worsen mental health. Column

[15] Neither the data on memory nor on mental health is available for 2009 and 2011. Due to missing values in self-reported memory conditions in the sample years between 2000 and 2006, we end up with a substantially smaller sample.

(2) directly checks this relationship based on the 2015 survey. The OLS estimation results based on the 2015 cross section show that individuals from provinces with high housing prices are more likely to feel stressful, providing direct supportive evidence of the effects on mental health, though one needs to interpret the result as only a correlation.

In Appendix B (Section 3.8.2), we use the China Family Panel Survey or CFPS database (Institute of Social Science Survey, 2015) to cross-check this channel. Compared to the CHNS, the CFPS database includes detailed information on mental health measures and covers more provinces. The results in Table 3.B.1 show that individuals from provinces experiencing a higher housing price growth tend to report stronger depression. Taken together, these results provide suggestive evidence that mental health is a potential channel through which rising housing prices raise the incidence of having chronic diseases.

3.5.2.3 Changes in Lifestyle

As a response to increased work intensity and mental stress, individuals may also change their lifestyle. One may increase the consumption of tobacco and alcohol to relieve stress (Ayyagari and Sindelar, 2010). Longer working hours may also reduce time for relaxing activities, such as sleeping and exercise. All these changes in lifestyle could increase the probability of chronic diseases. In Table 3.9, we examine whether housing prices change individuals' consumption of tobacco and alcohol, as well as their average time devoted to sports activities and sleep.

Panels A and B display results for individuals born before 1965 and those born after 1965, respectively. This is motivated by the fact that the younger cohort may adjust their lifestyle and habits relatively easily (Tseng and Lin, 2008). We find that housing price appreciation raises the consumption of tobacco among both cohorts, whereas consumption of alcohol does not respond to housing prices for either cohort. Consumption of tobacco can significantly increase the risk of having respiratory diseases, revealing that increased tobacco consumption could be a reason of higher prevalence of chronic diseases with housing price appreciation. We do not find evidence that individuals change their sports activities following housing price growth. Finally, the results in column (4) show that the younger cohort tends to reduce sleeping time with higher housing prices. Such an impact is not observed for the older cohort. Changes in sleeping time among the younger generation are consistent with the results in Table 3.7 where housing prices increase work intensity.

Results based on the CFPS as shown in Table 3.B.1 in Appendix B (Section 3.8.2) confirm that individuals change their lifestyle as a response to housing price appreciation by raising alcohol consumption. By and large, the results in Table 3.9 suggest that changes in lifestyle, especially increased consumption of tobacco and reduced sleeping time among the younger cohort, could be a mechanism explaining the negative health effects of housing prices.

3.6 Conclusions

Housing prices have been increasing substantially in China since the late 1990s, when the market-oriented reforms to the real estate market were implemented. Recent studies find that the rapid housing price appreciation has profound effects on socio-economic outcomes, such as household consumption, labor market participation, fertility, and marriage. This paper widens the spectrum of this strand of literature by examining the health consequences of rising house prices in the context of China and investigating various potential mechanisms.

Using individual-level data from the CHNS between 2000 and 2011, which is further linked to province-level data on housing prices, we employ an instrumental variable approach to identify a causal relationship between housing prices and the prevalence of chronic diseases. We find robust evidence that the rise in housing prices in China significantly increases the prevalence of chronic diseases. Our baseline results imply that a 10 percent increase in residential housing prices induces a 3.29 percentage point rise in the probability of having chronic diseases.

We find that the negative health effects are closely related to the marriage culture, which lead to males feeling pressured to purchase a household before marriage, as well as the strong competition in the marriage market among young males caused by a shortage of brides. The negative health effect due to the stress of purchasing houses before marriage could be transmitted to parents who have a marriage-age son. We also find evidence that individuals respond to the increase in housing prices with increased work intensity, elevated mental stress, higher tobacco use, and a reduction in sleep time. These findings not only comprehensively illustrate the health-related changes at the individual level, but also highlight the role of culture and social norms in shaping the health effects of rising housing prices.

The rapid real estate appreciation is not unique to China, but is pervasive in many other countries across the globe. Our findings based on the Chinese context shed light on the potential downside of the real estate market growth in other countries.

While the prosperity of the real estate industry could contribute to economic growth, complementary policies that alleviate the negative health consequences should be considered in the design of housing policies.

3.7 Chapter 3 - Figures and Tables

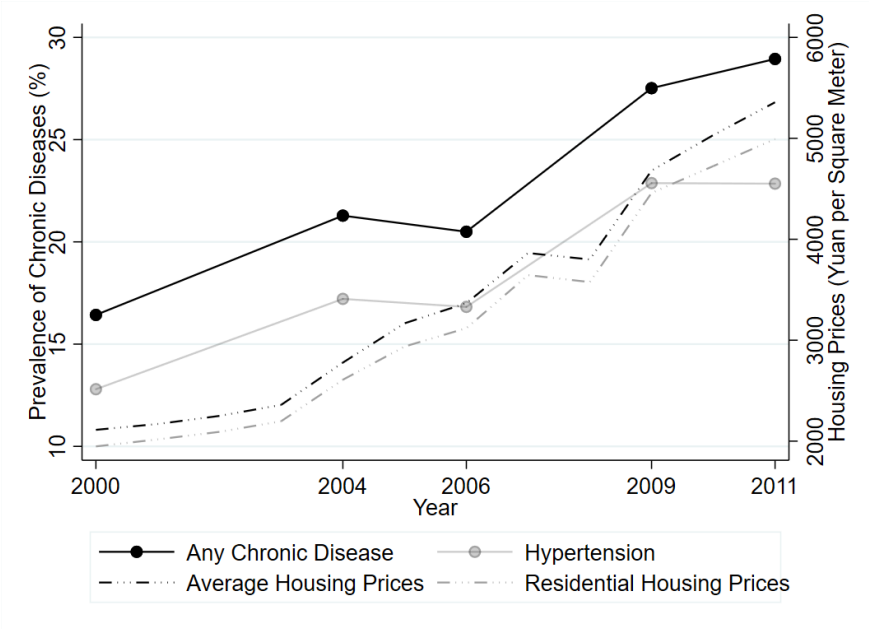


Figure 3.1: Average Housing Prices and the Prevalence of Chronic Diseases in China, 2000-2011

Data sources: Housing price data are from various issues of *China Real Estate Statistics Yearbook*. The prevalence of chronic diseases and hypertension is calculated based on the CHNS dataset.

Table 3.1: Average Residential Housing Prices and the Incidence of Chronic Diseases

	(1)	(2)	(3)	(4)
ln Residential Housing Prices	0.445***	0.415***	0.417***	0.329***
	(0.128)	(0.127)	(0.128)	(0.094)
First-stage Results				
Initial Land Supply×Interest Rates	-0.732***	-0.728***	-0.727***	-0.799***
	(0.196)	(0.195)	(0.195)	(0.186)
First-stage SW F-Statistics	13.96	13.99	13.95	18.36
Individual Controls		✓	✓	✓
Household Controls			✓	✓
Province Controls				✓
Individual Fixed Effects	✓	✓	✓	✓
Region-year Fixed Effects	✓	✓	✓	✓
Observations	32,111	32,111	32,111	32,111
R^2	0.560	0.563	0.563	0.566

Notes: Table reports the 2SLS estimation results of equation (3.1). Individual controls include age, age squared, marital status, and education level; household controls include ln household gross income, ln household size, and a rural or urban dummy; province controls include healthcare staff per 1,000 local residents, ln waste water emissions in 10,000 tons, ln sulfur dioxide emissions in tons, GDP per capita, and unemployment rates. Two-way cluster-robust standard errors at the individual and the province-year level are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, *$p < 0.1$.

Table 3.2: Average Residential Housing Prices and the Incidence of Chronic Diseases: Robustness Checks

	Balanced Panel (1)	HH FE (2)	Avg. Wage (3)	Export Share (4)	Import Share (5)	Medical Expense (6)
ln Residential Housing Prices	0.274**	0.335***	0.499***	0.332***	0.345***	0.324***
	(0.126)	(0.102)	(0.172)	(0.094)	(0.095)	(0.097)
ln Average Wage			-0.162			
			(0.197)			
Export Share				0.000		
				(0.000)		
Import Share					0.000	
					(0.000)	
ln Medical Expense						-0.019
						(0.054)
Control Variables	✓	✓	✓	✓	✓	✓
Individual FE	✓	✓	✓	✓	✓	✓
Region-year FE	✓	✓	✓	✓	✓	✓
First-stage SW F-Stat	17.18	15.95	11.06	21.83	20.60	16.98
Observations	11,250	32,036	32,111	32,111	32,111	32,111
R^2	0.508	0.568	0.564	0.566	0.566	0.566

Notes: This table reports the 2SLS estimation results of various robustness checks. Column (1) presents the results based on a balanced panel with individual appearing in all sample years; column (2) adds household fixed effects to the baseline specification; column (3) controls for average wages; columns (4) and (5) separately control for the shares of exports and imports in local GDP; and column (6) includes province-level per-capita average household expenditure on medical services. All regressions include a full set of control variables as in column (4) of Table 3.1. Two-way cluster-robust standard errors at the individual and the province-year level are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 3.3: Housing Prices and the Incidence of Chronic Diseases: Alternative Measures of Housing Prices and Health Outcomes

	Chronic Diseases			Hypertension	
Housing price measures	Residential	Business	All Types	Residential	All Types
	(1)	(2)	(3)	(4)	(5)
ln Housing Prices	0.432**	0.195***	0.280***	0.323***	0.275***
	(0.171)	(0.060)	(0.076)	(0.092)	(0.073)
ln High-class Residence Prices	-0.085				
	(0.060)				
First-stage Results					
Initial Land Supply×Interest Rates	-0.563***	-1.347***	-0.938***	-0.799***	-0.938***
	(0.154)	(0.244)	(0.181)	(0.186)	(0.181)
First-stage SW F-Statistics	13.37	30.44	26.79	18.37	26.79
Control Variables	✓	✓	✓	✓	✓
Individual Fixed Effects	✓	✓	✓	✓	✓
Region-year Fixed Effects	✓	✓	✓	✓	✓
Observations	32,111	32,111	32,111	32,103	32,103
R^2	0.565	0.567	0.567	0.584	0.585

Notes: The table reports the 2SLS estimation results of equation (3.1) using alternative measures of housing prices in columns (1) to (3) and using the incidence of hypertension as the dependent variable in columns (4) and (5). Average housing prices in column (3) indicate the average price level of all types of houses, including residential houses, villas and high-grade apartments, office buildings and houses for other business activities. All regressions include a full set of control variables as in column (4) of Table 3.1. Two-way cluster-robust standard errors at the individual and the province-year level are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, *$p < 0.1$.

Table 3.4: Average Residential Housing Prices and the Incidence of Chronic Diseases: Using Average Housing Prices in Neighboring Provinces as the Instrumental Variable

	Chronic Diseases (1)	Hypertension (2)
ln Residential Housing Prices	0.143**	0.127**
	(0.068)	(0.058)
First-stage Results		
Average Prices in Neighboring Provinces	0.172***	0.172***
	(0.025)	(0.025)
First-stage SW F-Statistics	45.52	45.52
Control Variables	✓	✓
Individual Fixed Effects	✓	✓
Region-year Fixed Effects	✓	✓
Observations	32,111	32,103
R^2	0.567	0.585

Notes: This table reports the 2SLS estimation results of equation (3.1) using average housing prices in all neighboring provinces as an alternative instrumental variable. All regressions include a full set of control variables as in column (4) of Table 3.1. Two-way cluster-robust standard errors at the individual and the province-year level are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, *$p < 0.1$.

Table 3.5: Residential Housing Prices and the Incidence of Chronic Diseases: Falsification Test

Dependent Variable: Chronic Disease in 1989-1997	Housing Prices in 1999-2007 (1)	Housing Prices in 2004-2012 (2)
Panel A: Main IV		
ln Residential Housing Prices	-0.224	-0.165
	(0.157)	(0.184)
First-stage SW F-Statistics	3.95	2.62
Observations	24,242	24,242
R^2	0.549	0.551
Panel B: Alternative IV		
ln Residential Housing Prices	0.308	0.609
	(0.190)	(0.418)
First-stage SW F-Statistics	10.69	3.87
Observations	24,285	24,817
R^2	0.550	0.536
Control Variables	✓	✓
Individual Fixed Effects	✓	✓
Region-year Fixed Effects	✓	✓

Notes: This table reports the 2SLS estimation results of equation (3.1) using the prevalence of chronic diseases in 1989-1997 as the dependent variable. Column (1) and (2) use residential housing prices in 1999-2007 and those in 2004-2012 as the housing price measure, respectively. Panel A uses initial land supply times interest rates and panel B uses average housing prices in neighbouring provinces as the instrumental variable, respectively. All specifications include a full set of control variables as in column (4) in Table 3.1. Two-way cluster-robust standard errors at the individual and the province-year level are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, $^{*}p < 0.1$.

Table 3.6: Average Residential Housing Prices and the Incidence of Chronic Diseases: The Role of Marriage Market Competition

	Sex (aged 20-40)		Adult Son (Parent Sample)		With a Marriage (Full Sample)	
	Male	Female	Yes	No	Yes	No
	(1)	(2)	(3)	(4)	(5)	(6)
ln Residential Housing Prices	0.406***	0.229**	0.451***	0.103	0.479**	0.257***
	(0.148)	(0.113)	(0.121)	(0.181)	(0.229)	(0.089)
$\beta_a = \beta_b$ (p-value)	0.224		0.025		0.018	
Control Variables	✓	✓	✓	✓	✓	✓
Individual Fixed Effects	✓	✓	✓	✓	✓	✓
Region-year Fixed Effects	✓	✓	✓	✓	✓	✓
First-stage SW F-Statistics	19.68	19.15	17.38	20.47	13.87	17.94
Observations	7,623	8,569	18,240	7,843	5,014	25,451
R^2	0.495	0.507	0.557	0.573	0.567	0.557

Notes: The table reports the 2SLS estimation results of equation (3.1) based on various sub-samples. Parents with a young adult son is defined as those having at least one male child aged 20 to 40 when interviewed. Households with a marriage are those having a marriage happening within the last 12 months. All regressions include a full set of control variables as in column (4) of Table 3.1. Two-way cluster-robust standard errors at the individual and the province-year level are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, *$p < 0.1$.

Table 3.7: Average Residential Housing Prices and Work Intensity

	ln Weekly Working Days (1)	ln Daily Working Hours (2)	ln Weekly Working Hours (3)
ln Residential Housing Prices	0.645***	0.681***	1.392**
	(0.211)	(0.214)	(0.543)
Control Variables	✓	✓	✓
Individual Fixed Effects	✓	✓	✓
Region-year Fixed Effects	✓	✓	✓
First-stage SW F-Statistics	20.34	20.74	20.70
Observations	17,820	17,575	16,149
R^2	0.475	0.565	0.539

Notes: The table reports the 2SLS estimation results of equation (3.1) using measures of work intensity as the dependent variable. All regressions include a full set of control variables as in column (4) of Table 3.1. Two-way cluster-robust standard errors at the individual and the province-year level are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, *$p < 0.1$.

Table 3.8: Residential Housing Prices, Memory and Mental Stress

	Sample: 2000-2006 (2SLS)	Sample: 2015 (OLS)
	Self-reported Memory	Any Mental Stress (=1)
	(1)	(2)
ln Residential Housing Prices	-0.415**	0.077**
	(0.168)	(0.029)
Control Variables	✓	✓
Individual Fixed Effects	✓	
Region-year Fixed Effects	✓	
First-stage SW F-Statistics	37.44	
Observations	1,535	7,515
R^2	0.608	0.014

Notes: This table reports the 2SLS and OLS estimation results of equation (3.1) using measures of memory as the dependent variable for the sample from 2000 to 2006 (column 1), and using measures of mental stress as the dependent variable for the sample of the 2015 cross section (column 2), respectively. All regressions include a full set of control variables as in column (4) of Table 3.1. Two-way cluster-robust standard errors at the individual and the province-year level are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 3.9: Housing Prices and Changes in Lifestyle

	Tobacco (=1)	Alcohol (=1)	Sport (=1)	Sleep Hours
	(1)	(2)	(3)	(4)
Panel A: Pre-1965 Cohort				
ln Average Residential Housing Prices	0.200*	-0.059	0.044	0.700
	(0.102)	(0.131)	(0.086)	(0.712)
First-stage SW F-Statistics	19.57	17.04	17.54	8.21
Observations	6,359	18,416	17,976	13,991
R^2	0.559	0.711	0.487	0.456
Panel B: Post-1965 Cohort				
ln Average Residential Housing Prices	0.307**	-0.105	0.085	-2.389**
	(0.146)	(0.115)	(0.085)	(1.175)
First-stage SW F-Statistics	22.56	20.56	21.13	11.23
Observations	3,845	13,489	10,717	13,229
R^2	0.556	0.659	0.511	0.465
Control Variables	✓	✓	✓	✓
Individual Fixed Effects	✓	✓	✓	✓
Region-year Fixed Effects	✓	✓	✓	✓

Notes: This table reports the 2SLS estimation results of equation (3.1) using various measures of daily activities as the dependent variable. Columns (1), (2) and (3) measure whether an individual consumed tobacco, alcohol, and did sports, respectively. Column (4) uses daily hours of sleep as the dependent variable. All regressions include a full set of control variables as in column (4) of Table 3.1. Two-way cluster-robust standard errors at the individual and the province-year level are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, *$p < 0.1$.

3.8 Chapter 3 - Appendix

3.8.1 Additional Figures and Tables

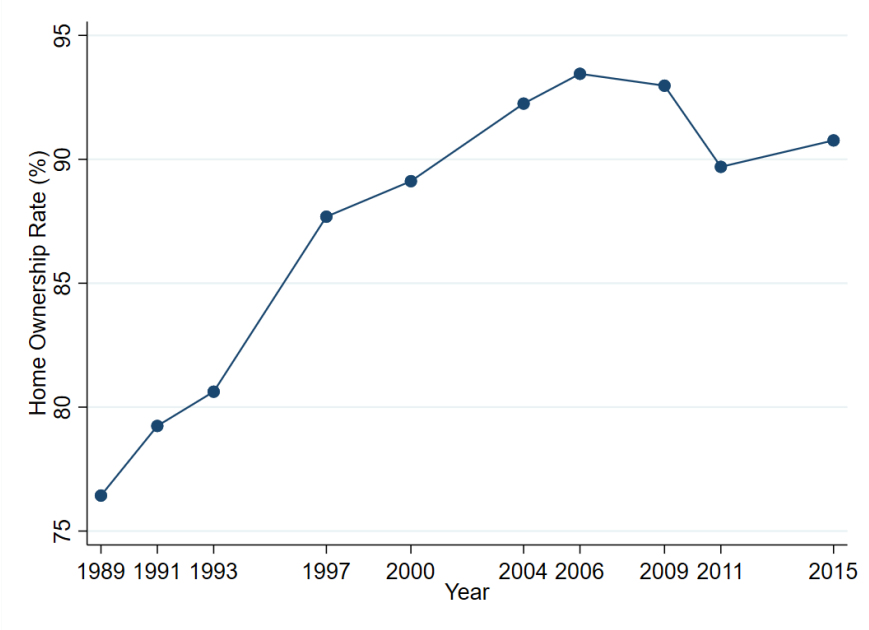
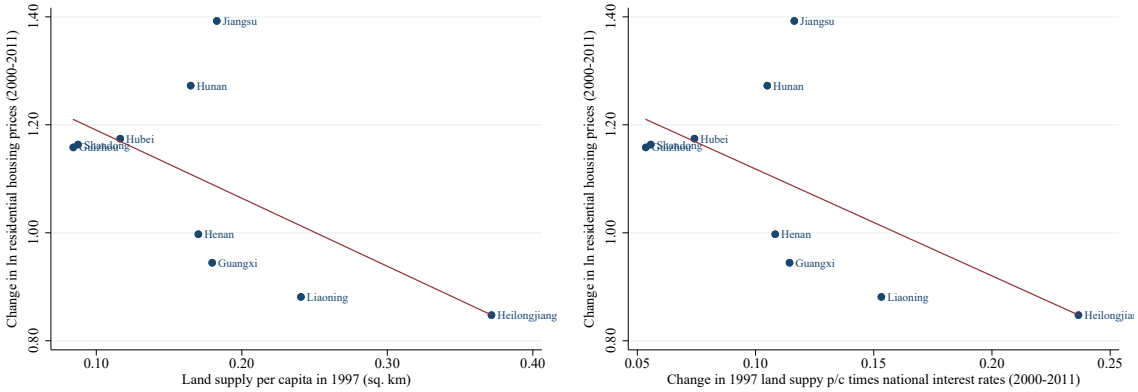


Figure 3.A.1: Home Ownership Rate in China, 1989-2015

Data source: Own calculation based on the CHNS dataset. Home ownership rate is calculated as the share of households owning as opposed to renting the current home in total number of households.



(a) Initial Land Supply (b) Change in IV

Figure 3.A.2: Changes in Housing Prices, Initial Land Supply, and the Instrumental Variable

Notes: Figure 3.A.2a shows the scatter plots of changes in ln housing prices between 2000 and 2011 and initial land supply by province. Initial land supply is measured as planned urban residential area in 1997 divided by urban population. Figure 3.A.2b shows the scatter plots of changes in ln housing prices and the instrumental variable by province. The instrumental variable is defined as equation (3.2).

Table 3.A.1: Summary Statistics of Key Variables

	Mean	S.D.	Min.	Max.
Housing Price Measures				
Average residential housing prices	2,505.530	1,119.830	987.000	6,145.200
ln average residential housing prices	7.730	0.440	6.895	8.723
Prices of high-class residences	4,917.272	2,326.008	1,923.895	11,643.360
ln prices of high-class residences	8.390	0.474	7.562	9.362
Prices of business and office buildings	4,703.045	1,870.552	1,725.620	9,504.770
ln prices of business and office buildings	8.376	0.404	7.453	9.160
Average housing prices of all types	2,691.233	1,155.867	1,079.000	6,554.410
ln average housing prices of all types	7.809	0.421	6.984	8.788
Health Outcomes				
Any chronic disease (=1)	0.187	0.390	0.000	1.000
Hypertension (=1)	0.146	0.353	0.000	1.000
Individual-level Characteristics				
Female (=1)	0.518	0.500	0.000	1.000
Age	42.709	10.666	15.000	60.000
Married (=1)	0.880	0.325	0.000	1.000
Education level				
Below primary	0.145	0.352	0.000	1.000
Primary school	0.203	0.403	0.000	1.000
Lower middle school	0.375	0.484	0.000	1.000
Upper middle school	0.152	0.359	0.000	1.000
Technical or vocational degree	0.069	0.254	0.000	1.000
University of college	0.054	0.226	0.000	1.000
Master or above	0.001	0.027	0.000	1.000
Household-level Characteristics				
ln household income	9.905	1.035	1.609	14.058
ln household size	1.284	0.376	0.000	2.565
Urban (=1)	0.386	0.487	0.000	1.000
Household registration type				
Urban (=1)	0.287	0.452	0.000	1.000
Rural (=1)	0.488	0.500	0.000	1.000
Rural-Urban (=1)	0.187	0.390	0.000	1.000
Urban-Rural (=1)	0.038	0.192	0.000	1.000
Own purchase of current house (=1)	0.679	0.467	0.000	1.000
With at lease one adult son (=1)	0.699	0.459	0.000	1.000
Province-level Characteristics				
Health care staff per capita	49.870	13.355	23.098	72.808
ln waste water emission (10,000 tons)	12.192	0.721	9.933	13.302
ln sulfur dioxide emission (tons)	13.730	0.447	12.309	14.510
GDP per capita	19.049	12.993	2.759	62.290
Unemployment rate (%)	3.851	0.620	2.600	6.500

Notes: Data on housing price measures is from various issues of the *China Real Estate Statistics Yearbook*. Health outcomes, individual- and household-level characteristics are abstracted from the CHNS database. Province-level characteristics are from various issues of *China Statistical Yearbook*. The number of observations is 32,111.

Table 3.A.2: Average Residential Housing Prices and the Incidence of Chronic Diseases: Logit Bootstrapping 2SLS Regressions

	(1)	(2)	(3)	(4)
ln Residential Housing Prices	4.183***	4.032***	4.075***	3.169**
	(1.344)	(1.140)	(1.225)	(1.324)
Individual Controls		✓	✓	✓
Household Controls			✓	✓
Province Controls				✓
Individual Fixed Effects	✓	✓	✓	✓
Region-year Fixed Effects	✓	✓	✓	✓
Observations	10,634	10,634	10,634	10,634
Pseudo R^2	0.113	0.114	0.116	0.118

Notes: This table reports the 2SLS logit model estimation results of equation (3.1) using national interest rates interacted with provincial land supply in 1997 as the instrumental variable (equation (3.2)). Individual controls include age, age squared, marital status, and education level; household controls include ln household gross income, ln household size, and a rural or urban dummy; province controls include healthcare staff per 1,000 local residents, ln waste water emissions in 10,000 tons, ln sulfur dioxide emissions in tons, GDP per capita, and unemployment rates. Robust standard errors are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 3.A.3: Housing Prices and Additional Health Outcomes

	Sick in the Past 4 Weeks (=1) (1)	Overweight (=1) (2)	High Waist Circumference (=1) (3)	High Waist-hip Ratio (=1) (4)
ln Residential Housing Prices	0.052	0.028	0.115**	0.091
	(0.128)	(0.074)	(0.051)	(0.119)
Control Variables	✓	✓	✓	✓
Individual Fixed Effects	✓	✓	✓	✓
Region-year Fixed Effects	✓	✓	✓	✓
First-stage SW F-Statistics	18.39	19.51	19.52	19.75
Observations	32,011	30,040	29,903	29,529
R^2	0.409	0.747	0.630	0.538

Notes: This table reports the 2SLS estimation results of equation (3.1). Sick in the past four weeks is based on self reports. One is overweight if the body mass index exceeds 25. Waist circumference is identified as being high if it is over 102cm for males and over 88cm for females. Waist-hip ratio is identified as being high if waist circumference exceeds 0.90 of hip circumference for males, and 0.85 for females. All regressions include a full set of control variables as in column (4) of Table 3.1. Two-way cluster-robust standard errors at the individual and the province-year level are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, *$p < 0.1$.

Table 3.A.4: Average Residential Housing Prices and the Incidence of Chronic Diseases: Robustness Checks by Excluding One Province Each Time

Excluded Province:	LN	HLJ	JS	SD	HA	HB	HN	GX	GZ
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
ln Residential Housing Prices	0.225***	0.315**	0.360***	0.284***	0.318***	0.329***	0.356***	0.636*	0.638*
	(0.070)	(0.128)	(0.125)	(0.082)	(0.077)	(0.113)	(0.116)	(0.337)	(0.338)
First-stage Results									
Initial Land Supply ×Interest Rates	-0.905***	-1.439***	-0.658***	-0.874***	-0.849***	-0.920***	-0.753***	-0.343*	-0.343*
	(0.174)	(0.432)	(0.156)	(0.204)	(0.174)	(0.251)	(0.190)	(0.192)	(0.192)
First-stage SW F-Statistics	27.18	11.12	17.90	18.30	23.72	13.44	15.78	3.19	3.19
Control Variables	✓	✓	✓	✓	✓	✓	✓	✓	✓
Individual Fixed Effects	✓	✓	✓	✓	✓	✓	✓	✓	✓
Region-year Fixed Effects	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	28,599	28,294	28,525	28,727	28,577	28,730	28,530	28,115	28,791
R^2	0.560	0.571	0.560	0.568	0.568	0.565	0.566	0.568	0.569

Notes: The table reports the 2SLS estimation results of equation (3.1) by excluding observations in one province in each column. LN stands for Liaoning, HJ Heilongjiang, JS Jiangsu, SD Shandong, HA Henan, HB Hubei, HN Hunan, GX Guangxi, GZ Guizhou. All regressions include a full set of control variables as in column (4) of Table 3.1. Two-way cluster-robust standard errors at the individual and the province-year level are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, *$p < 0.1$.

Table 3.A.5: Average Residential Housing Prices and the Incidence of Chronic Diseases: Heterogeneity by Age Group

Age group	15-30	31-45	46-60
	(1)	(2)	(3)
ln Residential Housing Prices	-0.154	0.421***	0.468***
	(0.197)	(0.132)	(0.160)
Control Variables	✓	✓	✓
Individual Fixed Effects	✓	✓	✓
Region-year Fixed Effects	✓	✓	✓
First-stage SW F-Statistics	16.73	20.29	18.46
Observations	4,036	11,448	13,623
R^2	0.519	0.551	0.598

Notes: The table reports the 2SLS estimation results of equation (3.1) by age cohort. All regressions include a full set of control variables as in column (4) of Table 3.1. Two-way cluster-robust standard errors at the individual and the province-year level are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 3.A.6: Average Residential Housing Prices and the Incidence of Chronic Diseases on Parents: Fathers and Mothers

	Adult Son		No Adult Son	
	Father	Mother	Father	Mother
	(1)	(2)	(3)	(4)
ln Residential Housing Prices	0.401***	0.516***	0.351	-0.129
	(0.142)	(0.161)	(0.267)	(0.200)
$\beta_a = \beta_b$ (p-value)	0.856		0.684	
Control Variables	✓	✓	✓	✓
Individual Fixed Effects	✓	✓	✓	✓
Region-year Fixed Effects	✓	✓	✓	✓
First-stage SW F-Statistics	17.79	17.08	21.39	20.04
Observations	8,161	9,779	3,637	3,998
R^2	0.538	0.577	0.563	0.589

Notes: The table reports the 2SLS estimation results of equation (3.1) based on various sub-samples. Parents with a young adult son are defined as those having at least one male child aged 15 to 40 when interviewed. All regressions include a full set of control variables as in column (4) of Table 3.1. Two-way cluster-robust standard errors at the individual and the province-year level are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

3.8.2 Housing Prices, Mental Health, and Changes in Lifestyle: Evidence from the CFPS

In this section, we cross-check the role of mental stress, and lifestyle changes in affecting the health consequences of housing prices using an alternative dataset, the China Family Panel Survey (CFPS) (Institute of Social Science Survey, 2015). It is a longitudinal survey of Chinese families and individuals starting from 2010, and four waves until 2016 are available. In each wave, more than 10,000 households were interviewed. It reports rich information on individual and household characteristics, with the focus on the economic and subjective well-being. Compared to the CHNS, the CFPS covers a larger number of provinces, which allows for larger variations in housing appreciation across regions.[16] Additionally, the CFPS includes information on individual's mental health, time use, and consumption of tobacco and alcohol throughout the four waves.

CFPS measures the adult psychological well-being using scales of the Center for the Epidemiological Studies of Depression (CES-D) (Radloff, 1977). The CES-D is a validated screening tool for detecting depressive symptoms, and is suitable to be used across different racial, gender and age groups (Weissman et al., 1977). The standard CES-D covers a total of eight to 20 items on mental health, and individuals are asked about the frequency of experiencing each item during the past week. The scale of each item ranges from one to five, indicating a frequency from never, rarely, sometimes, often, to very often, with a higher scale indicating a worse mental health. The CFPS covers different numbers of the CES-D items across years, and we calculate the average score to harmonize the variation in item numbers.[17]

We match the CFPS dataset with provincial housing prices and estimate equation (3.1) using measures of mental health, consumption of tobacco and alcohol, and sleep time as the dependent variable. Similar to the analysis using CHNS, we restrict the sample to individuals aged between 15 and 60 years old. As the CFPS covers a shorter panel than the CHNS, we control for household instead of individual fixed effects to allow for more variations. To address the endogeneity of housing prices, we use the same instrumental variable as specified in equation (3.2), i.e. initial land supply interacted with national interest rates. The 2SLS estimation results are set out in Table 3.B.1.

[16] CFPS covers 25 provinces in 2010, 28 in 2012, 29 in 2014, and 30 in 2016.
[17] CFPS has six items on CES-D in 2010, 20 in 2012, 13 in 2014, and nine in 2016.

The first-stage results in Table 3.B.1 indicate that our instrument is negatively correlated with housing prices, with the F-statistics of weak instrument test ranging from 0.96 to 21.89, indicating that in some cases the instrumental variable is weak. Column (1) shows that rising housing prices lead to higher depression scores as measured by CES-D. Columns (2) to (5) show that housing prices affect neither consumption of tobacco nor the frequency of alcohol drinks significantly, but raise the volume of alcohol drinks. Notice that the sample size of those reporting a non-missing number of alcohol drinks is greatly reduced. We should therefore interpret the results cautiously. Lastly, the result in column (6) suggests that housing prices have a negative but insignificant impact on sleeping hours. Overall, the results here are consistent with our earlier findings based on the CHNS in Tables 3.8 and 3.9 that housing price appreciation leads to deteriorated mental health and some changes in lifestyle.

Table 3.B.1: Housing Prices, Depression, Tobacco and Alcohol Use, and Sleep Time: Evidence from the CFPS

	Depression	Tobacco		Alcohol		Sleep
	Score	Yes (=1)	ln Num	Frequent (=1)	ln Num	Hours
	(1)	(2)	(3)	(4)	(5)	(6)
ln Residential Housing Prices	1.234**	-0.046	0.423	0.177	1.630*	-2.690
	(0.599)	(0.068)	(0.365)	(0.135)	(0.858)	(2.905)
First-stage Results						
Initial Land Supply	-0.437***	-0.174*	-0.124	-0.174*	-0.677***	-0.083
× Interest Rates	(0.132)	(0.092)	(0.094)	(0.092)	(0.145)	(0.084)
First-stage SW F-Statistics	10.96	3.57	1.72	3.57	21.89	0.98
Individual Controls	✓	✓	✓	✓	✓	✓
Household Controls	✓	✓	✓	✓	✓	✓
Household Fixed Effects	✓	✓	✓	✓	✓	✓
Region-year Fixed Effects	✓	✓	✓	✓	✓	✓
Observations	69,196	90,418	24,017	90,416	1,813	66,047
R^2	0.434	0.593	0.634	0.392	0.655	0.300

Notes: This table reports the 2SLS estimation results using the depression score, the incidence of tobacco use, ln number of cigarettes per day, the incidence of frequent alcohol use (more than 3 times per week), ln number of alcoholic drinks per week, and sleep hours as dependent variables. Depression score is an average score of different measures of mental depression, each measured with 1-5 scale. The higher the score the worse the mental health. Individual controls include sex, age, age squared, marital status, and level of education; household controls include household income categories, ln household size, and an urban or rural household type dummy. Two-way cluster-robust standard errors at the individual and the province-year level are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

3.8.3 Additional Heterogeneous Results

3.8.3.1 The Role of Household Income

Considering that low-income households often endure heavier financial burdens from housing purchases (Fang et al., 2016), we examine whether households with different income levels are affected differently. In columns (1) and (2) of Table 3.C.1, we classify households into low- and high-income ones based on their average income level across years within each province. Households with a higher than the 75th percentile income level are defined as high-income ones, and others are low-income ones. The empirical results show that housing prices increase the prevalence of chronic diseases only for low-income households, while the high-income ones are immune to the rapid housing price appreciation. This suggests that individuals who are relatively financially constrained are more vulnerable to the negative effects of rising living costs induced by housing price growth.

Table 3.C.1: Average Residential Housing Prices and the Incidence of Chronic Diseases: Heterogeneity by Household Income and Registration Type

	Income Level		Registration Type			
	Low	High	Urban	Rural	Rural-Urban	Urban-Rural
	(1)	(2)	(3)	(4)	(5)	(6)
ln Residential	0.509***	0.221	0.366**	0.344***	0.457**	0.246
Housing Prices	(0.142)	(0.152)	(0.168)	(0.111)	(0.199)	(0.445)
$\beta_a = \beta_b$ (p-value)	0.338		0.594	0.568	(ref.)	0.314
Control Variables	✓	✓	✓	✓	✓	✓
Individual FE	✓	✓	✓	✓	✓	✓
Region-year FE	✓	✓	✓	✓	✓	✓
First-stage SW F-Stat	16.44	20.06	18.00	17.32	17.96	10.25
Observations	23,943	7,920	9,203	15,637	5,962	1,245
R^2	0.564	0.571	0.611	0.533	0.569	0.594

Notes: The table reports the 2SLS estimation results of equation (3.1) by household income levels and registration type. Households with incomes above the 75th percentile level are classified as high-income, and the others are low-income. Rural-urban households are those that changed from a rural to urban household registration type during the sample period, and urban-rural households are those that changed from an urban to rural household registration type. All regressions include a full set of control variables as in column (4) of Table 3.1. Two-way cluster-robust standard errors at the individual and the province-year level are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

3.8.3.2 Hukou Registration Status

In China, the household registration (hukou) is classified into urban and rural types. The hukou system is closely linked to local social welfare and restricts the access to social welfare for the migrants without a local hukou. In order to pursue higher incomes and better quality social welfare, many rural individuals migrate to urban areas, wishing to obtain an urban hukou. However, one condition of applying for a local hukou is having fixed and stable residence in the city and many regions require house ownership as the pre-requisite.[3.C.1] Although housing appreciation mainly occurred in urban China (Fang et al., 2016), the health effect could be further extended to rural households that intend to move to urban areas.

To investigate potential differential impacts on households with various registration status, we classify households into four types: urban households, rural households, households that changed from a rural to urban type, and those changed from a urban to rural type.[3.C.2] The change from a rural to urban type could happen due to two reasons. First, the whole communities may be reclassified from a rural to urban type administratively during the process of urbanization (Gan et al., 2019). The registration type of households in the affected regions is changed accordingly. The second reason is rural-urban migration due to work or marriage purposes. While the first type does not necessarily induce household changes, the latter is more likely to involve a change of household, e.g. moving out of the parents' household and forming a new one in the city.

We estimate the health effects for each type of households separately, and the results are set out in columns $(3)-(6)$ of Table 3.C.1. The estimation results indicate that the negative health effects induced by rising housing prices are prevalent among individuals from both urban and rural households. The effect on rural households is similar in magnitude to that on urban households, although the housing price appreciation occurs mainly in urban areas. One potential explanation is related to rural-urban migrants who work and live in the urban areas holding a rural hukou. Non-local hukou holders are often constrained to or even prohibited from purchasing houses in many cities. Meanwhile, migrant workers earn an average lower income than

[3.C.1] The hukou system restricted rural to urban migration before the late 1990s when some regions implemented a hukou reform in small towns and cities that allowed rural migrant workers to obtain a local hukou if certain conditions were satisfied. Since 2001, the reform has been expanded to some medium and large cities, and in 2014, the hukou reform was expanded nationwide.

[3.C.2] In our sample, 28.71 percent of households are always registered as urban, 48.76 percent always rural, 18.76 percent changed from a rural to urban hukou type, and only 3.78 percent changed from an urban to rural type.

natives (e.g., Meng, 2012). With these two factors, the rise in housing prices imposes a high pressure on migrants who wish to purchase houses in the urban area (Li and Liu, 2018). Such effects could be further transmitted to family members in the rural area, in particular parents, as many Chinese parents consider purchasing a house for their son as an obligation, especially for marriage purposes, as evident in Table 3.6. Column (5) shows that the health impact on those changing from a rural to urban hukou is slightly stronger than urban and rural households, although this difference is not significant. An individual would experience a 4.57 percentage point increase in the probability of having chronic diseases with a 10 percent rise in residential housing prices. Rural to urban hukou holders can be affected due to different reasons. Those successfully changing their hukou from a rural to an urban type are relatively young without much saving but they often have to purchase a residence in order to obtain an urban hukou.[3.C.3] They are thereby more vulnerable. Not surprisingly, housing prices do not affect the health of urban to rural migrants. This could also be due to a much smaller sample size, though.

[3.C.3] Those who moved from rural to urban areas (average age 41) are on average two years younger than those who always stayed in urban areas (average age 43). In many cities, having a local residence is a pre-requisite to apply for a local hukou (Wang et al., 2021).

Chapter 4

Natural Disasters and Intimate Partner Violence: Evidence from Peru*

with:
Laura Barros

* We are grateful to Axel Dreher, Renate Hartwig, Krisztina Kis-Katos, Stephan Klasen, Lukas Menkhoff, Samuel Siewers, Maria Micaela Sviatschi, and participants at the Globalization and Development (GlaD) Workshop, the 2nd Workshop on Geodata in Economics, the 2021 Annual Meeting of the American Economic Association, the International Conference in Development Economics, the Spring Meeting of Young Economists 2021, and the 2nd Brazilian Meeting in Family and Gender Economics for comments and suggestions. We thank the Peruvian Ministry for Women and Vulnerable Populations (MIMP) for sharing relevant data. We acknowledge funding by the Deutsche Forschungsgemeinschaft (DFG), Project RTG 1723.

Abstract

This paper investigates if women are more vulnerable to intimate partner violence (IPV) in the aftermath of earthquakes in Peru. By combining household-level data on IPV and spatial data on all earthquakes that happened between 2000 and 2009, we show that exposure to very strong earthquakes increases the incidence of IPV by more than ten percentage points. We document that the effect is more pronounced for women residing in urban areas and in districts without protective institutions. Further evidence suggests that the increase in IPV following earthquakes is induced by an increase in male intra-household economic power, higher likelihood of the male partner staying in the household, and a rise in male alcohol consumption. Our findings shed light on a relatively neglected aspect of post-disaster settings and highlights the role of protective policies for women following large-scale disasters.

JEL Codes: J12, Q54, J16, I10.
Keywords: Intimate Partner Violence; Natural Disasters; Earthquake; Peru

4.1 Introduction

A large body of research shows that natural disasters and extreme weather events affect conflict (Miguel et al., 2004; Harari and Ferrara, 2018), crime (Kwanga et al., 2017), political instability (Dell, 2012; Jia, 2014; Almer et al., 2017), and political change (Brückner and Ciccone, 2011).[1] However, the effects of large-scale natural disasters on gender violence remain relatively neglected.[2] This paper fills this research gap by investigating if women are more vulnerable to intimate partner violence (IPV) in the aftermath of earthquakes in Peru. Additionally, we disentangle the role of external and internal factors explaining male violent behavior following large-scale natural shocks.

Violence against women is a pervasive phenomenon throughout the world. According to estimates from the World Health Organization (WHO), 35 percent of women experience some form of violence in their lifetime, with the majority of cases corresponding to IPV (WHO, 2013). Women who suffer violence are more likely to develop physical and emotional health problems (Campbell et al., 2002; Plichta, 2004). Sexual violence, in particular, has been shown to decrease educational attainment (Rees and Sabia, 2013) and female labor force participation (Sabia et al., 2013; Chakraborty et al., 2018). For children, exposure to violence results in higher levels of distress (Levendosky et al., 2013) and lower investments in human capital (Trako et al., 2018). Therefore, understanding the consequences of natural disasters for violence against women can have important implications for governments and international organizations in post-disaster settings.

We hypothesize that women might be more vulnerable to IPV in the aftermath of earthquakes for four main reasons. First, natural disasters are expected to have negative effects on income and wealth (Bui et al., 2014). Since women's (and men's) economic status are key determinants of intra-household violence, we expect large-scale disasters to affect IPV through an economic channel.[3] Second, exposure to large-scale

[1] For an overview of the literature, see Dell et al. (2014).

[2] A related literature has documented the relationship between rainfall shocks and violence against women in agricultural dependent societies, with the hypothesis that agricultural income or harvest would be affected in light of a rainfall shock. See, for example, Miguel (2005), and Sekhri and Storeygard (2014).

[3] From the theory side, intra-household bargaining models and male-backlash models have focused on the role of household resources and intra-household bargaining power in predicting violence against women. While bargaining models predict that violence decreases when women control more resources (Aizer, 2010; Anderberg et al., 2016), backlash models suggest the opposite, with males becoming more violent as a way of compensating for the loss in relative status (Macmillan and Gartner, 1999).

disasters can have profound psychological consequences that could lead to violent behavior, both directly and indirectly (e.g., through a loss in household wealth). These psychological changes include, for example, higher mental distress, frustration and alcohol consumption–well-known determinants of violence against women (Card and Dahl, 2011; Luca et al., 2015). Third, natural disasters can harm existing infrastructure and institutional capacity, potentially increasing women's vulnerability to violence (e.g., through a reduction in protective policies or law enforcement). Finally, disasters might alter the amount of time couples spend together, potentially affecting vulnerability to violence through a purely "mechanical" channel.

In this paper, we combine geo-referenced data on IPV from the Peruvian Demographic and Health Surveys (DHS) with ShakeMaps of all significant earthquakes that happened between 2000 and 2009. ShakeMaps consist of intensity polygons illustrating the ground motion and shaking intensity of earthquakes, as calculated by the United States Geological Surveys (USGS). Throughout our study period, Peru was hit by two very strong earthquakes (in 2001 and in 2007) affecting different regions of the country. We exploit the spatial and temporal distribution of earthquakes and the location of households to investigate if IPV is more pronounced in areas that are severely affected by the shock, relying on the exogenous nature of earthquakes. To reduce concerns on reporting bias of IPV in the DHS data, we additionally rely on administrative data on the number of domestic violence cases at the department level for robustness.

The empirical results show several relevant findings. First, we find robust evidence that exposure to very strong earthquakes six months prior to the survey increases the probability that women experience violence by more than ten percentage points on average.[4] Similar evidence is also found at the aggregate level, as documented by an increase in the number of registered cases of violence against women from six up to 36 months after the earthquake. In terms of underlying risk channels, our analysis indicates that the increase in violence is more pronounced in urban areas and in districts without protective institutions. Lastly, we find evidence suggesting that our results are partially explained by an increase in co-residence with the male partner, rising intra-household male economic dominance, as well as higher alcohol consumption by the partner.

This paper is related mainly to two strands of the literature. First, it adds directly to the literature documenting the socioeconomic effects of weather shocks and

[4] Following Gignoux and Menéndez (2016), we focus on earthquakes greater or equal to 7.0 in the MMI scale, referred to as very strong, due to their large potential of destruction.

natural disasters. It is widely documented that natural disasters can fuel civil conflict (Miguel et al., 2004; Harari and Ferrara, 2018), hinder long-run economic growth (Lackner, 2018), decrease human-capital accumulation (Caruso, 2017) and damage the economy (Cavallo et al., 2010). Additionally, one strand of the literature has documented increases in violence against women following weather shocks in agricultural-dependent societies (Miguel, 2005; Sekhri and Storeygard, 2014; Abiona and Koppensteiner, 2016). Our paper relates most closely to the work of Weitzman and Behrman (2016), which investigates the incidence of violence against women in Haiti following the 2010 earthquake. We contribute to this research by additionally investigating the temporal dynamics of the effect of earthquakes on IPV, and by disentangling the channels explaining male violent behavior.

Second, the paper contributes to the large body of papers studying the determinants of violence against women. For instance, IPV has been shown to depend on a number of factors, including female empowerment (Macmillan and Gartner, 1999; Aizer, 2010; Bhattacharyya et al., 2011; Anderberg et al., 2016; Bulte and Lensink, 2019), female wages (Aizer, 2010), transfers to women (Haushofer et al., 2019; Roy et al., 2019; Heath et al., 2020; Menon, 2020), social norms (Green et al., 2019), family structure (Khalil and Mookerjee, 2019; Tur-Prats, 2019; Heath et al., 2020), psychological factors (Card and Dahl, 2011), and formal and informal institutions (Stevenson and Wolfers, 2006; Brassiolo, 2016; Amaral et al., 2018; Cunningham and Shah, 2018; Trako et al., 2018; Bargain et al., 2019; Miller and Segal, 2019). We highlight the role of these underlying channels also in the context of large-scale disasters as triggers for violence. More specifically, we emphasize that, in addition to economic factors, protective institutions and psychological factors play a crucial role on women's vulnerability to IPV.

The rest of the paper is structured as follows. Section 4.2 provides background information on gender violence and earthquakes in Peru. Section 4.3 describes our data sources and Section 4.4 discusses our identification strategy and empirical model. The empirical results and robustness checks are presented in Section 4.5. Sections 4.6 and 4.7 discuss heterogeneity analyses and underlying channels, respectively. The last section concludes.

4.2 Peruvian Context

Peru is located at the border of the Nazca and the South American tectonic plates, in the so-called *Ring of Fire*, an area particularly prone to volcanic eruptions and major

seismic activity. According to the Global Facility for Disaster Reduction and Recovery (GFDRR), Peru has been affected by more than 30 major earthquakes in the last four centuries, out of which two of them occurred in our study period (2000-2009). Despite its risky location and the relative recurrence of major natural disasters, Peru still lacks in terms of disaster preparedness and response (PDC, 2015).[5] The relatively high levels of poverty and social vulnerability makes shock coping even more challenging, particularly for women and minorities who often lack resources and social networks.[6] In that respect, pre-existing historical inequalities are expected to be amplified in the aftermath of disasters (Mutter, 2015).

In Peru, gender inequalities are reflected in different dimensions. In addition to the well-documented economic disadvantages, women are underrepresented in political spheres and are more often subject to discrimination (Ñopo, 2009).[7] Violence against women, in particular, is a major public health problem in Peru, with more than a third of the respondents in our sample reporting having suffered some form of violence throughout their lifetime. According to Flake (2005), the high incidence of violence relates closely to rigid gender roles and strong norms of masculinity prevailing in the country. While men are traditionally expected to embody masculine stereotypes by being strong and aggressive, women are often expected to play a submissive role (Suárez Farfán et al., 2016). According to data from the Organisation for Economic Co-operation and Development (OECD, 2020b), one third of women in Peru agree that in some cases violence against women is justifiable.[8]

In contrast to these informal gender norms, the official legislation in Peru criminalizes violence against women, with many recent policies aiming at protecting women and girls and enforcing punishment.[9] Since the creation of the Ministry for the Promotion of Women and of Human Development (PROMUDEH), in 1996, the country

[5] The assessment of risk is calculated considering multiple dimensions, such as multi-hazard exposure, vulnerability and coping capacity (PDC, 2015).

[6] According to the World Development Indicators (WDI), in 2008 around 37 percent of the population in Peru was below the poverty line, as measured by the headcount ratio.

[7] According to the World Bank Gender Indicators (WBGI), in 2018 women held 27.7 percent of seats in national parliament and 27.8 percent of the ministerial positions in Peru. In economic terms, in 2019 only 8.6 percent of Perivian women were employed in industry while the male share corresponded to 21.1 percent.

[8] The precise definition of the indicator is the following: "The percentage of women who agree that a husband/partner is justified in beating his wife/partner under certain circumstances" (OECD, 2020b).

[9] According to the Ministry of Women and Vulnerable Populations (MIMP), two laws address violence against women and intra-family violence, more generally. First, the law 26260/1993 on the protection against intra-family violence and the law 30364/2015 on the prevention and eradication of violence against women and other family members (Suárez Farfán et al., 2016).

has put in place many policies to protect women and other vulnerable groups.[10] One of the most important policies to curb gender violence was the creation of the Women Justice Centers (WJC), that started in 1999 with the objective of providing legal and psychological support for victims of violence. It is estimated that the roll-out of these centers has contributed to a significant reduction in violence against women (Trako et al., 2018). While there has been an increase in the number and the coverage of policies to tackle violence against women in Peru, there is no clear agenda to curb violence following large-scale disasters. In the next section, we present the data sources we use to estimate empirically the relationship between earthquakes and IPV, as well as to investigate the role of internal and external mediating factors.

4.3 Data

4.3.1 Domestic Violence

We rely on two data sources to measure violence against women. First, individual-level data on IPV come from the domestic violence module contained in the Demographic and Health Surveys (DHS). All women aged 15-49 that have ever been married or cohabiting are eligible for the survey, and one woman per household is randomly selected for the interview. In addition to the questions about violence against women, the survey contains information on individual and household socioeconomic characteristics. We make use of four waves of the Peruvian DHS: the standard DHS 2000, the continuous DHS 2004-06, the continuous DHS 2007-08, and the standard DHS 2009. The data consist of repeated cross sections, with part of the survey clusters being interviewed in multiple rounds in the continuous DHS. Additionally, we have information on the GPS coordinates of the centroids of the survey clusters, which allows for the construction of our measure of exposure to earthquakes at a fine-grained level. In the domestic violence module of the DHS, there are four different measures of IPV: less severe violence, mental violence, severe violence, and sexual violence. In the 2000 wave of the Peruvian DHS, only the variable "less severe violence" is available in the dataset. In later waves, all four types of IPV incidence are available.

[10] The PROMUDEH was subsequently replaced by the Ministry of Women and Social Development (MIMDES), in 2002, and by the MIMP, in 2012.

Our second data source on violence against women come from MIMP. The data contain the monthly number of registered cases of violence against women in each Peruvian department between 2002 and 2009.[11]

Using two different sources of information on violence against women has the advantage of attenuating concerns on reporting bias, a frequently discussed issue in the gender violence literature. While the estimated numbers in both sources are likely to be underestimated due to the high emotional costs and barriers associated with reporting, there are reasons to believe that the sources of bias would be different for a survey in comparison to official statistics. Therefore, we would find it reassuring if the two datasets would show similar patterns. Additionally, as long as there is no differential reporting bias depending on exposure to earthquakes, we should be able to estimate the effect of exposure to earthquakes on changes in IPV.[12]

Table 4.1 shows the descriptive statistics for the individual outcome variable and the sociodemographic controls. More than a third of the women in our sample report having experienced some form of violence in their life. Women are on average 33 years old and four years younger than their partners. Seven percent of the women in our sample did not attain any formal education, 39 percent completed primary education, 38 percent completed secondary education and 16 percent attained tertiary education. The partners have, on average, relatively higher education attainment than the respondents.

4.3.2 Earthquakes

Information on all significant earthquakes that happened in Peru between 2000 and 2009 come from the USGS. The data consist of ShakeMaps, which portray the distribution and the spread of earthquakes by mapping out its ground motion and shaking intensity.[13] During our sample period, two very strong earthquakes happened in Peru (in 2001 and 2007, respectively), as shown in Figure 4.1.[14] The orange areas were affected by very strong earthquakes according to the MMI scale. The map also shows the centroids of the DHS clusters (in dark blue the clusters from the DHS 2000 and

[11] A Peruvian department, *Departamento*, is an administrative division which is analogous to a state.

[12] We acknowledge that official reporting of cases of violence against women might be disrupted following natural disasters. Nevertheless, we expect this to be less of a concern in the case of DHS, whose data collection is designed well in advance.

[13] Figure 4.A.5 in the Appendix (Section 4.10) shows examples of ShakeMaps of the two very strong earthquakes that happened in Peru over our study period obtained from the USGS.

[14] Following the USGS, we define earthquakes as *very strong* if the shaking intensity is equal to or above 7.0 in the MMI scale. Figure 4.A.6 shows the correspondence between the intensity categories and respective perceived shaking and potential damage, as defined by the USGS.

in forest green the clusters from the DHS 2009). We construct our exposure measure by intersecting the earthquake intensity polygons and the geographic coordinates of the DHS clusters. Therefore, we have a measure of exposure which based on inference, rather than actual reporting of earthquake experience. There is arguably a great amount of unpredictability in the locations where the earthquakes hit, which reduces concerns that earthquake exposure is spatially correlated over time (Lackner, 2018). Table 4.A.1 provides an overview of the two earthquakes that happened in Peru during the study period, with the most relevant information on the date, intensity, affected departments and estimated damage.

In our analysis, we exploit the temporal and spatial distribution of the earthquakes, as well as its severity, to construct our exposure variable. The next section discusses our identification strategy, our empirical model as well as the assumptions needed to obtain causal estimates.

4.4 Empirical Strategy

In this section, we discuss our estimation models and present the variables of interest. For our individual-level analysis, we have a pooled sample of 43,110 women distributed over the period 2000 and 2009. The sample is restricted to women aged 15 to 49 that have ever been married to or cohabiting with a male partner. Moreover, we restrict the sample to respondents that have been living in the same location for at least three years to reduce concerns on earthquake-driven in-migration. In our individual specification, we estimate the following model:

$$IPV_{wcdst} = \alpha_0 + \beta Earthquake_{c;(t-i,t)} + \lambda \mathbf{X}_w + \delta_t + \gamma_d + \eta_s \times t + \epsilon_{wcdst} \tag{4.1}$$

where IPV_{wcdst} is a binary outcome variable, denoting whether respondent w, living in DHS cluster c, district d, state (or *Department*) s, interviewed in year t, reports that she ever experienced IPV.[15] On the right-hand side, α_0 is a constant. $Earthquake_{c;(t-i,t)}$ is a dummy, indicating whether the women residing in cluster c has experienced an earthquake in the past i months up to the time of the interview. Thus, the average effect of earthquakes on IPV is captured by the β coefficient. By including district fixed effects we are essentially exploiting variation within districts.

[15] The DHS defines that a respondent suffers IPV if she reports that the husband has pushed her, shook her, thrown something at her, slapped her, punched her, kicked her or dragged her.

Our treatment is defined as one if respondents live in a DHS cluster that has been affected by an earthquake equal to or above 7.0 in the MMI scale up to i months before the interview. According to the MMI scale, earthquakes equal to or above 7.0 have very strong shaking intensity and high potential of destruction (Gignoux and Menéndez, 2016). We control for a vector of individual and household characteristics, $\mathbf{X}_w$, including respondent's age, squared age, age difference with the partner, educational level and partner's educational attainment to increase the precision of the estimates. We control only for sociodemographic characteristics of the respondents and their partners as these variables are less likely to be bad controls, *i.e.* to be affected by the earthquakes. We control for year fixed effects, δ_t, to account for changes in IPV that are common to all women in the country. District fixed effects, γ_d, account for constant factors that determine violence against women within the district. These include, for example, persistent gender values and norms that stay constant throughout the sample period. To account for region-specific trends in IPV, we also include department time trends, $\eta_s \times t$. We estimate the coefficients using a linear probability model (LPM) for ease of interpretation and to accommodate the different fixed effects. Standard errors are clustered at the district level.[16]

Identification Strategy In our main specification, we exploit the temporal and spatial variation in exposure to earthquakes to investigate if women living in regions affected by earthquakes are more vulnerable to IPV. The earthquake exposure is defined at the DHS cluster level and the outcomes are measured at the individual level. To obtain causal estimates, our main assumptions are that the spatial distribution and the timing of earthquakes are exogenous, which seem plausible given the nature of the shock. Our estimation strategy is closely related to that of Caruso (2017), which investigates the long-run inter-generational impacts of natural disasters in Latin America.

There are several potential threats to our identification strategy, which we try to circumvent in a number of ways. First, exposure to natural disasters might induce (in- and out-) migration, which could result in a selected sample of respondents. Assuming that more (less) violent individuals are more (less) likely to migrate after experiencing an earthquake, our estimated coefficients would be downward (upward) biased. While we can partially address in-migration by restricting our sample to individuals that have been living in the same location for a minimum of three years, we can not account for

[16] The level of exposure to the earthquake shocks is at the DHS cluster level. However, while the clusters in the DHS waves of 2000 to 2008 remained the same, the clusters of the 2009 wave were displaced. Thus, we cluster the standard errors at the district levels.

out-migration in our individual analysis. However, our aggregate-level analysis allows us to deal with intra-state migration, as the outcome variable is aggregated at the department level and should not be affected by within-state migration.

Second, quantitative studies on violence against women are typically plagued by reporting bias. Since intra-household violence is a sensitive topic, respondents often do not feel comfortable to truthfully report their experiences. In case exposure to the shock differentially affects reporting behavior, our results would be biased. By estimating our results using both survey and official data, we intend to reduce concerns on differential reporting bias following the earthquake.

Third, the occurrence of large-scale natural disasters often cause many casualties. In that regard, exposure to the earthquake makes it impossible to observe individuals following the disaster. While in the case of the 2001 earthquake the estimated number of deaths was relatively low (81), it was much higher for the 2007 earthquake, in which it was estimated that more than 500 people died, according to the USGS. Nevertheless, compared to the overall population size,[17] these casualties are most likely not large enough to bias the results.

Finally, in our setting we have more than one "treatment", and more than two time periods. Therefore, our estimate is essentially a weighted average of all two-by-two combinations between the "pure" control group, the early-treatment group (affected by the 2001 earthquake), and the late-treatment group (affected by the 2007 earthquake) (Goodman-Bacon, 2021; Cunningham, 2020). Assuming homogeneous treatment effects, we estimate unbiased coefficients. However, to reduce concerns on the estimation method, we implement alternative specifications with other treatment definitions and different time windows, as discussed in detail in the robustness section.

4.5 Results

This section starts with the results of the individual-level model, specified in equation (4.1). Then, it proceeds to discuss some robustness checks and the results of an aggregate-level analysis.

4.5.1 Main Results

Table 4.2 shows the estimated coefficients of our individual-level specification on the relationship between exposure to earthquakes and IPV. The dependent variable is

[17] According to WorldBank (2020), Peru's total population reached 28.6 million in 2008.

binary and captures whether the respondent has ever experienced IPV. In columns (1) and (2), exposure to earthquakes is a binary variable defined as one if respondents live in a DHS cluster that was hit by an earthquake equal to or above 7.0 in the MMI scale in the past six months. In columns (3) and (4) we measure the earthquake exposure in the past 12 months.

Columns (1) and (2) show a positive and highly significant contemporaneous effect of exposure to earthquakes on IPV. More specifically, we estimate that respondents that experienced a very strong earthquake six months prior to the interview were between 10.0 and 12.8 percentage points more likely to suffer IPV. When we change the earthquake exposure definition to 12 months, we still find a rise in IPV, although slightly smaller in magnitude and marginally less significant (columns (3) and (4)).

To investigate further how the effect of earthquakes on IPV changes over time, we construct alternative exposure measures by varying the length of the exposure window. For instance, we construct separate earthquake exposure variables from six up to 48 months prior to the interview.[18] Moreover, to test whether pre-trends exist, we construct leads of the exposure variable to estimate if future exposure to earthquakes affects present IPV. Empirically, we estimate equation (4.1) in separate regressions and plot the β coefficients (Figure 4.2a). We observe a positive impact of earthquakes on IPV in clusters that were hit by a very strong earthquake up to 18 months prior to the interview. This increase in violence is temporary, as seen by the insignificant coefficients after one and a half years. Lastly, none of the leads of the earthquake exposure variables exhibit a significant impact on IPV. This indicates that, overall, locations that are hit by earthquakes do not have systematically higher levels of violence against women prior to the shock.

Since our outcome variable captures if respondents *ever* experienced violence, we interpret the coefficients as changes at the extensive margin rather than at the intensive margin. This means that while we can not assess if there is an intensification of violence following disasters, we document increases in the probability that women *become* victims of violence in the aftermath of earthquakes. It is likely, therefore, that our coefficients are lower-bound estimates.

[18] We choose to vary the exposure window instead of constructing separate lags due to the low number of affected observations in each time window.

4.5.2 Robustness Checks

This section provides additional evidence showing that our results are not driven by a particular definition of earthquake exposure. Additionally, we show that aggregate-level results are consistent with our individual-level analysis.

Buffers around Clusters To ensure respondent's confidentiality, the DHS randomly displaces the GPS coordinates of part of the survey clusters. The displacement ranges from zero up to two kilometers for urban clusters, from zero up to five kilometers for 99 percent of rural clusters and ten kilometers for one percent of rural clusters in the sample. Our original earthquake exposure measure only takes into account if the DHS cluster's centroid intersects with earthquake polygons. Therefore, to reduce concerns that our results are affected by this particular definition, we construct an alternative exposure measure by intersecting the earthquake polygons with buffers around the centroids of the DHS clusters. Following the DHS displacement rule, we construct two-kilometer buffers around urban and five-kilometer buffers around rural clusters. The results using this alternative exposure measure are presented in Figure 4.2b.

Overall, the results remain very much comparable to our baseline model, shown in Figure 4.2a. We find a positive and significant impact of earthquakes on IPV when respondents experienced an earthquake up to 18 months prior to the interview. Additionally, we do not detect pre-trends in the effects, with all coefficients for the lead exposure variables being insignificant. Finally, the smaller significance of our coefficient is not entirely surprising, if we consider that with the buffer measure we are more likely to wrongly assign treatment to clusters that have not in fact been treated, which is expected to downward bias our coefficients.

Difference-in-difference Estimations

In our baseline regression model, we have multiple treatments (i.e., two independent earthquakes) occurring at different points in time and affecting different clusters. This implies that we do not have pure control and treatment groups and hence raises the concern that our results could be biased in case there are heterogeneous treatment effects (Cunningham, 2020; Goodman-Bacon, 2021). In this section, we implement alternative difference-in-difference specifications with varying treatments and time windows to gain further understanding on the results we document. Panel A of Table 4.A.3 shows the results of a "pooled" DID in which the treatment (*Earthquake*) is equal to one if a cluster was affected by one of the earthquakes (2001 or 2007) and zero otherwise (never treated group). Post is equal to one if interviews were conducted

after 2001 (or 2007) and zero otherwise. The interaction coefficient in this specification captures the average effect of being treated by either of the earthquakes over all years. In column (2), the interaction coefficient ($Post \times Earthquake$) is positive and significant, which indicates that our previous estimates are mainly driven by the "short-run" effects of the 2007 earthquake. The null results for the 2001 earthquake indicate that in the medium run the effect of earthquakes on IPV fades away. In this specification we can only capture medium to long run effects, since the first DHS wave after the earthquake was conducted in 2004 (three years after the shock).

Alternative Intensity Threshold

In our baseline specification, we define that women are exposed to a very strong earthquake if they reside in a cluster that was hit by an earthquake equal to or above 7.0 in the MMI scale. We now turn to the results of an alternative specification in which we vary the threshold of our exposure variable to include earthquakes equal to or above 6.0 in the MMI scale. The results are plotted in Figure 4.A.1 in the Appendix (Section 4.10). We do not find any evidence that being exposed to a strong earthquake ($\geq$ 6.0) affects the probability of women ever experiencing IPV, though the coefficients for 12 and 18 months are positive and close to standard significance levels. Compared to the main results shown in Figure 4.2a, the effects are smaller in magnitude. Finally, we do not detect any pre-trend either.

Alternative Earthquake Measure

As an additional robustness exercise, we vary the time window for the definition of the earthquake exposure variable. Instead of estimating the impacts of being exposed to earthquakes at different time ranges, as in equation (4.1), we run a regression model including all leads and lags jointly, where treatment is defined as one if a cluster has been affected by an earthquake at different time ranges: in the 1-6 months ([1,6]), 7-12 months ([7,12]), ... until 31-36 months ([31,36]) prior to the interview. In addition, we also include leads of the earthquake treatment dummies and check for the existence of pre-trends. As is seen in Table 4.1, less than one percent of the women in the sample are affected by a very strong earthquake six or 12 months prior to the interview. When including leads and lags far away from the interview year, the estimation of some coefficients is not feasible, since no cluster was exposed to the respective treatment. Thus, we include dummies indicating earthquake exposure from 36 up to 18 months after the interview date. Coefficient plots of the earthquake variables are shown in Figure 4.A.2 in the Appendix (Section 4.10). We find that the positive impact of earthquakes on IPV is driven by women who experience a very strong earthquake within six months prior to the interview. Moreover, no pre-trend

is observed.

Alternative IPV Measures

In our main specification, we use the less severe violence variable as the outcome of interest due to data availability reasons. In this section, we use the other three measures of IPV (mental violence, severe violence, and sexual violence) as alternative outcome variables to check whether exposure to earthquakes increases the probability of women suffering other types of IPV as well. It is worth noting that in these specifications, the estimation of the effects is only based on the survey waves collected between 2004 and 2009. The estimation results are plotted in Figure 4.A.3 in the Appendix (Section 4.10). We see from Figure 4.A.3a that being exposed to earthquakes results in a long lasting increase in the probability that women experience mental violence, although the coefficients are only marginally significant. Regarding vulnerability to severe violence (Figure 4.A.3b) and sexual violence (Figure 4.A.3c), we do not observe any corresponding change following very strong earthquakes.

Aggregate-level Results We now test empirically if aggregate patterns of violence in the aftermath of earthquakes are consistent with the individual-level results. As mentioned before, since reporting bias is a common issue in quantitative studies of violence against women, comparing individual survey information with official numbers can shed light on the strength of the results. For our aggregate analysis we have an unbalanced monthly panel covering all Peruvian states between 2002 and 2009. We estimate the following event study model:

$$IPV_{sm} = \alpha_0 + \sum_{i=-48}^{48} \beta_i Earthquake_{s;(m-i,m-i+1)} + \delta_m + \gamma_s + \theta_s \times t + \epsilon_{sm} \tag{4.2}$$

Where the outcome variable is the number of cases of domestic violence registered in state s in month m. Exposure to earthquake is binary and captures if a state was hit by an earthquake equal to or above 7.0 in the MMI scale between month i and $i-1$.[19] We control for month fixed effects δ_t, state fixed effects γ_s, and state trends $\theta_s \times t$.

Our results (Figure 4.3) indicate that exposure to earthquakes significantly increases the number of domestic violence cases registered at the state level, particularly after 6 months, which is consistent with our individual-level results. The increase in domestic violence cases is permanent, persisting for more than 36 months after the shock. The fact that in this specification we capture changes at the intensive mar-

[19] Figure 4.A.4 in the Appendix (Section 4.10) illustrates which states have been affected by the earthquakes.

gin might explain the differences compared with the individual-level results. Finally, in this specification we do not observe substantial pre-trends, as seen by the mostly insignificant lead coefficients.

4.6 Heterogeneity

In this section, we turn back to the individual-level analysis to investigate if there are heterogeneous effects of earthquakes on IPV depending on individual, household and district characteristics. We focus specifically on rural or urban status, availability of protective institutions, and women's economic status. Empirically, we interact the heterogeneity variables with the earthquake dummies, and include the interaction term as well as the heterogeneity dummy into equation (4.1). The estimation results are shown in Table 4.3.

Rural and Urban Locations Urban areas are typically characterized by higher population density and larger concentration of buildings. We hypothesize that earthquakes could have a higher potential of destruction in these areas in comparison to rural regions, as the loss of income and wealth would be larger in urban areas. Table 4.3 shows our heterogeneity results along this dimension. The estimated coefficients in column (1) indicate that women residing in urban areas affected by an earthquake in the past 12 months are about 10.7 percentage point more likely to suffer from IPV. For rural areas, on the other hand, we do not detect any effect of experiencing a strong earthquake on vulnerability to IPV.

External Protective Institutions Existing evidence documents the role of protective institutions in reducing violence against women, both in Peru (Trako et al., 2018) and in other developing countries (Amaral et al., 2018; Miller and Segal, 2019). In this part, we investigate the role of proximity to protective institutions following large-scale disasters. On the one hand, having protective institutions might be crucial to protect women both at the onset and in the aftermath of disasters. However, if access is temporarily or permanently restricted due to the disaster, violence might actually increase. We construct a measure of access to WJC at the district level and interact it with our measure of exposure to earthquakes. Results are presented in column (2), Table 4.3. Although our results should be interpreted with caution, due to the potential endogeneity of WJC location, we document interesting patterns. The interaction coefficient is negative and significant. This indicates that women living in districts with a WJC would be less likely to experience IPV, while women living in

districts without WJC would experience a large increase (of around 19.5 percentage points) in IPV.

Female Employment Status Female economic status is considered an important determinant of violence against women. While bargaining models predict that increasing women's economic status would lead to a decrease in violence through a change in women's outside options (Aizer, 2010; Anderberg et al., 2016), backlash models predict the opposite, with males becoming more violent as a way of compensating for the loss in relative status (Macmillan and Gartner, 1999). In this subsection, we investigate the role of economic status as a source of heterogeneity linking earthquakes and IPV.

We interact women's employment status with exposure to earthquakes. The results are presented in column (3), Table 4.3. Although we do not detect any significant heterogeneous impacts of earthquakes on exposure to IPV, we find that, at baseline, women who are employed are 4.8 percentage points more likely to report having ever experienced IPV, as compared to unemployed women. One should note, however, that the differences in the incidence of IPV could be potentially related to differential reporting behavior depending on economic status.

4.7 Underlying Channels

In this section, we investigate potential channels explaining the positive effects of natural disasters on IPV. We start by looking at the role of alcohol consumption by the respondent's partner, the co-residence with the partner and then we look at intra-household economic dominance following earthquakes. Empirically, we estimate a regression model similar to equation (4.1), in which the outcome variable captures the respective channel.

Alcohol Consumption Existing evidence shows that exposure to earthquakes causes long-term post-traumatic distress (Valenti et al., 2013). Additionally, individuals experience behavioral changes including the use of substances, such as tobacco, alcohol (Ayyagari and Sindelar, 2010). Since mental distress and substance abuse are shown to affect inter-personal violence, we expect that women might be more vulnerable to IPV following earthquakes. In this section, we use the DHS question containing information on the alcohol consumption of the respondent's partner to investigate whether this is an underlying mechanism explaining our previous results.

The results are presented in columns (1) and (2) of Table 4.4. We use two alternative measures for alcohol consumption, including one that captures changes in

consumption at the extensive margin (a dummy indicating whether the partner drinks alcohol or not) as well as the intensive margin (whether the partner frequently gets drunk). The results show that earthquake exposure increases the probability that the respondent's partner drinks by 15.1 percentage points, although there is no significant effect on the likelihood of them getting frequently drunk.

Male Partner in the Household

In our sample, about five percent of the women do not co-reside with their partner. In the absence of co-residence, the likelihood of women suffering from IPV is lower. We hypothesize that after a large-scale natural disaster, partners might be more likely to return home and engage in the re-construction of their homes, which could affect vulnerability to violence due to a purely mechanic channel. We estimate whether the male partner is more likely to be in the household after an earthquake hits the household. The result is presented in column (3), Table 4.4. We estimate that being exposed to a very strong earthquake in the past 12 months leads to an increase in the likelihood of the partner living in the household by 3.2 percentage points. This indicates that our effects could be partially explained through a physical channel – if the couples do not co-reside, there is hardly the occurrence of IPV.

Intra-household Economic Dominance In addition to individual economic status, intra-household dynamics on economic decisions might also determine women's vulnerability to IPV (Naved et al., 2018). In this section, we construct measures of intra-household responsibility for economic decisions to investigate if this is a channel explaining the rise in IPV following earthquakes. While our heterogeneity analysis using women's employment status only allows us to infer the role of female economic status on IPV, here we look at the relative power balance between spouses with respect to the economic decisions. We focus on two questions from the DHS data: (i) who has the final saying on large household purchases, and (ii) who has the final saying on what to do with the money the husband earns. We construct two binary variables measuring intra-household economic power: 1) *Male Dominance* equals to one if the male partner is responsible alone for either of the economic decisions; and 2) *Any Joint Decision*, which equals to one if the respondent and the partner share any decisions in the household. We use these variables as alternative outcomes in our main estimation model specified in equation (4.1). The results are presented in columns (4) and (5) in Table 4.4.

Our results show that when households experience very strong earthquakes, there is a rise of four percentage points in male dominance. However, we do not find any changes in cooperation between spouses with regards to household decisions.

In addition, we also check whether the earthquakes would cause wealth loss and the result is set out in column (6) in Table 4.4. The outcome variable is the standardized wealth index based on the ownership of multiple items of assets in the DHS dataset. We do not find any significant impact of earthquakes on household wealth, although the coefficient is negative. Finally, in Table 4.A.4 in the Appendix (Section 4.10), we check the association between the channel variables and IPV. We document that the incidence of male partner drinking alcohol and male dominance are positively associated with IPV, which are consistent with our hypothesis.

4.8 Conclusions

While the socioeconomic consequences of natural disasters and extreme weather events are well documented in the literature, the dynamics of intra-household violence in the aftermath of disasters has not been investigated extensively. This paper provides new empirical evidence on the effects of earthquakes on IPV in Peru. Exploiting the temporal and spatial distribution of earthquakes, our results show that earthquake exposure significantly affects women's vulnerability to violence. For instance, being affected by a very strong earthquake in the past six months increases the probability that women suffer violence by 12.8 percentage points. The effect persists up to 18 months after the shock, indicating that vulnerability to violence relates not only to the emergency state immediately after the earthquake, but is also extended to the reconstruction period.

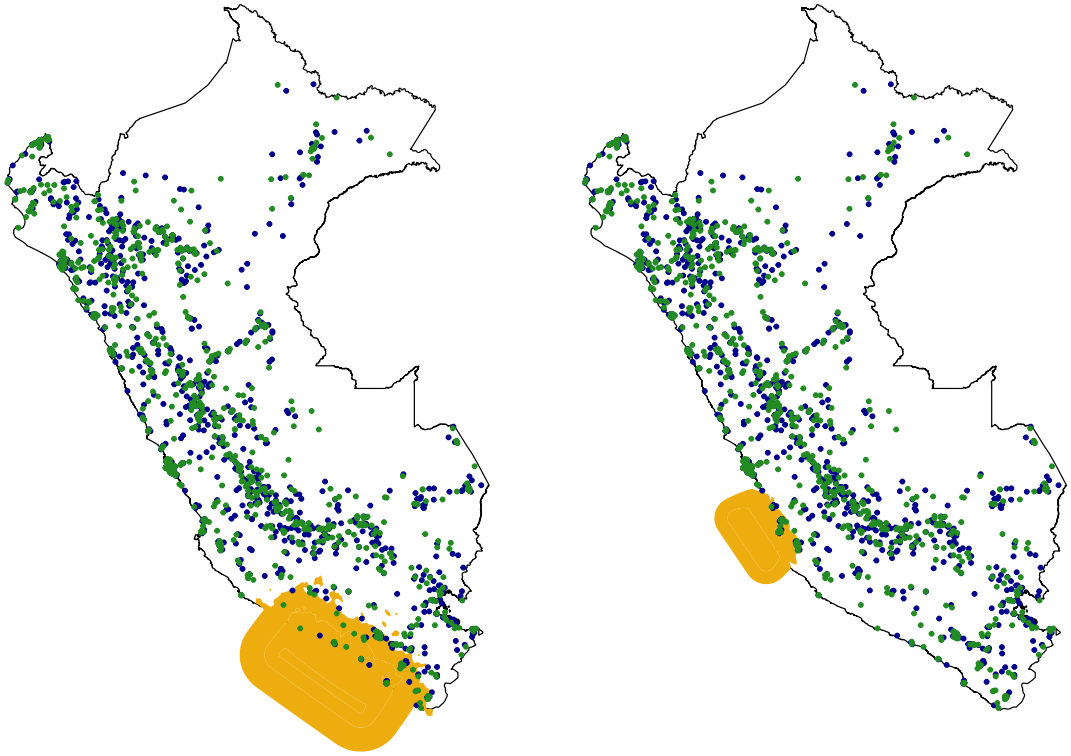
In terms of underlying channels and heterogeneities, we shed light on the role of individual and district factors affecting the incidence of IPV following earthquakes. For instance, our results show that the increase in violence following earthquakes is more pronounced in urban areas and in districts without protective institutions (WJC). Finally, channel analyses suggest that an increase in male intra-household economic dominance, higher likelihood of the partner being in the household, and a rise in male alcohol consumption are plausible underlying mechanisms explaining the rise in IPV following the earthquakes.

The nature of earthquakes is similar in many respects to other unexpected and highly destructive shocks, including other large-scale natural disasters (typhoons, tsunamis, etc.) as well as global or local epidemics. Therefore, the documented increase in women's vulnerability to IPV in the aftermath of natural disasters provides highly relevant policy implications that could be potentially extended to other settings. The vast majority of post-disaster relief programs focuses on the reconstruction

of infrastructure and economic transfers. This paper sheds light on an important but often neglected perspective, which also needs to be considered in post-disaster settings. In addition, our findings also relate to the growing literature documenting increases in IPV during the recent COVID-19 lockdowns (Agüero, 2021; Hsu and Henke, 2021). Along with tightening social distancing rules, policy makers should facilitate protective institutions for women.

4.9 Chapter 4 - Figures and Tables

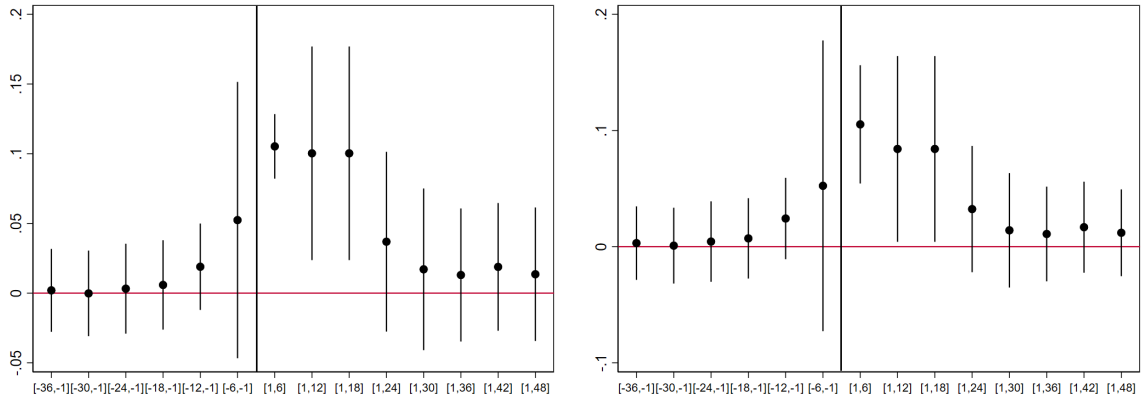
Figure 4.1: DHS survey clusters and 2001, 2005 and 2007 Earthquakes



(a) 2001 (b) 2007

Figures (a) and (b) show earthquake polygons equal to or above 7.0 in the MMI scale for the earthquakes that happened in Peru in 2001, and 2007. In addition, DHS clusters of the 2000 wave are shown in dark blue and DHS clusters of the 2009 wave are shown in forest green.

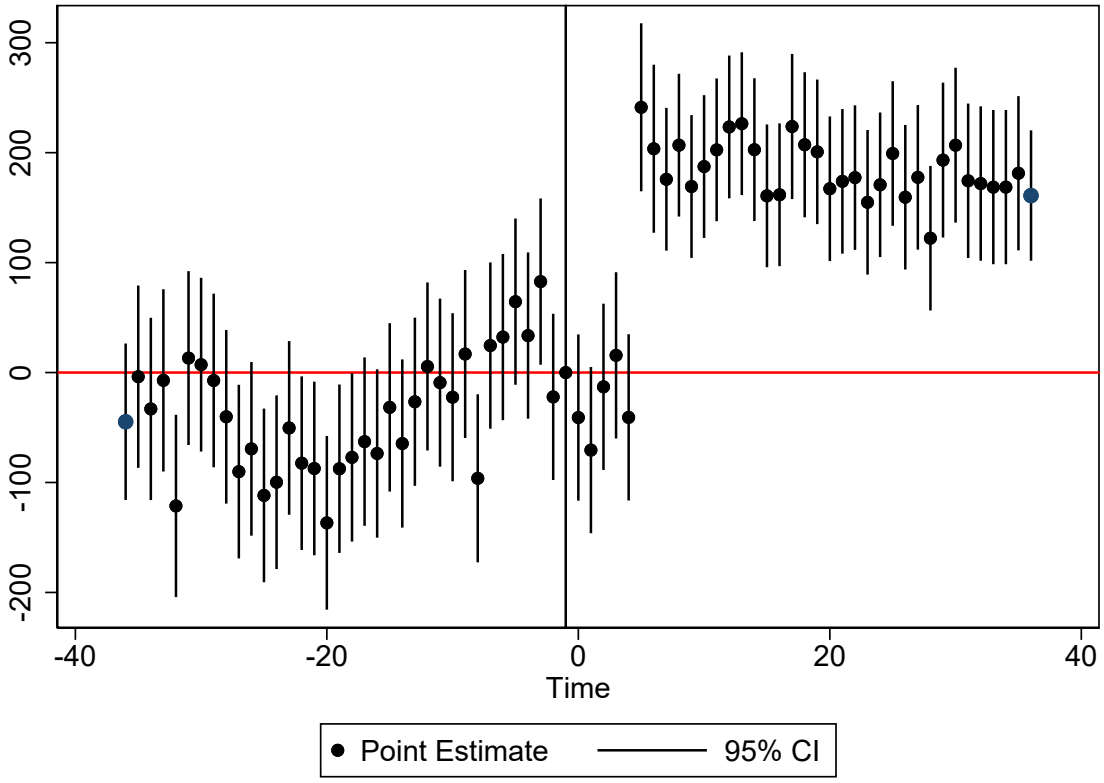
Figure 4.2: Earthquakes and Intimate Partner Violence



(a) No Buffer (b) With Buffer

Notes: No buffer indicates that the DHS cluster centers (as geographic points) are intersected with the ShakeMap directly. With buffer means that we first draw a buffer around the DHS cluster centers, and then intersect the buffers (as polygons) with the ShakeMap to measure whether the buffer has been hit by an earthquake in a certain year. Plot of leads and lags of earthquake coefficients estimated separately for earthquakes equal to or above 7.0 in the MMI scale. 90 percent confidence intervals.

Figure 4.3: Earthquakes and Intimate Partner Violence: Aggregate Results



Notes: Plot of leads and lags of earthquake coefficients estimated jointly for earthquakes equal to or above 7.0 in the MMI scale. 95 percent confidence intervals.

Table 4.1: Summary Statistics

	mean	sd	min	max
IPV - Less Severe Violence (0/1)	0.385	0.487	0	1
IPV - Mental Violence (0/1)	0.307	0.461	0	1
IPV - Severe Violence (0/1)	0.173	0.379	0	1
IPV - Sexual Violence (0/1)	0.084	0.277	0	1
Earthquake $\geq$ 7.0 in the Past 6 Months (0/1)	0.001	0.027	0	1
Earthquake $\geq$ 7.0 in the Past 12 Months (0/1)	0.003	0.055	0	1
Rural (0/1)	0.445	0.497	0	1
Women Justice Center in the District (0/1)	0.128	0.334	0	1
Currently Employed (0/1)	0.650	0.477	0	1
Partner Drinks (0/1)	0.741	0.438	0	1
Partner Frequently Drunk (0/1)	0.762	0.426	0	1
Partner in the Household (0/1)	0.958	0.202	0	1
Partner Economic Dominance (0/1)	0.295	0.456	0	1
hh Wealth (std.)	0.000	1.000	-2.202	2.514
Any Joint Decision (0/1)	0.886	0.317	0	1
Age	33.374	8.324	15	49
Squared age	1183.107	561.625	225	2401
Age difference between the respondent and the partner	4.064	5.802	-32	72
Respondent has no education	0.066	0.249	0	1
Respondent has primary education level	0.389	0.488	0	1
Respondent has secondary education level	0.376	0.484	0	1
Respondent has tertiary education level	0.169	0.375	0	1
Partner has no education	0.016	0.125	0	1
Partner has primary education level	0.313	0.464	0	1
Partner has secondary education level	0.489	0.500	0	1
Partner has tertiary education level	0.182	0.386	0	1
Observations	43,110			

Table 4.2: Earthquake and Intimate Partner Violence: Individual-level Results

	(1)	(2)	(3)	(4)
Earthquake $\geq$ 7.0 in the Past 6 Months (0/1)	0.128***	0.105***		
	(0.014)	(0.014)		
Earthquake $\geq$ 7.0 in the Past 12 Months (0/1)			0.079*	0.100**
			(0.043)	(0.047)
Year FE, District FE, Department Trend	✓	✓	✓	✓
Controls		✓		✓
Observations	43,110	43,110	43,110	43,110
R^2	0.056	0.074	0.056	0.074

Notes: The table reports the estimation results of equation (4.1) using linear probability model. Respondents are females aged 15 to 49 that have ever been married or cohabiting. The sample is restricted to females that have been living in the same location for at least 3 years. The outcome variable is a binary indicator which equals to one if the respondent reports having suffered any type of intimate partner violence and zero otherwise. Columns (1) and (2) report the effect of being exposed to earthquakes equal to or above 7.0 in the MMI scale in the past six months, while columns (3) and (4) show the effects of being exposed to earthquakes in the past 12 months. Control variables include respondents age, age squared, age difference with the partner, educational level and partner's educational level. Standard errors are clustered at the district level in all specifications.
*** $p < 0.01$, ** $p < 0.05$, *$p < 0.1$. Standard errors in parentheses.

Table 4.3: Earthquake and Intimate Partner Violence: Heterogeneity

Heterogeneity Variable (0/1):	Dep. Variable: Intimate Partner Violence (0/1)		
	Rural	WJC in District	Working
	(1)	(2)	(3)
EQ $\geq$ 7.0 12M (0/1)	0.107**	0.167***	0.114
	(0.046)	(0.044)	(0.072)
EQ $\geq$ 7.0 12M (0/1) $\times$ Hetero Var	-0.213***	-0.197***	-0.028
	(0.045)	(0.044)	(0.093)
Hetero Var	-0.071***	0.026*	0.048***
	(0.011)	(0.015)	(0.005)
Year FE, District FE, Department Trend	✓	✓	✓
Controls	✓	✓	✓
Observations	43,110	43,110	42,118
R^2	0.076	0.075	0.077

Notes: The table reports the estimation results of equation (4.1) using linear probability model. Respondents are females aged 15 to 49 that have ever been married or cohabiting. The sample is restricted to females that have been living in the same location for at least 3 years. The outcome variable is a binary indicator which equals to one if the respondent reports having suffered any type of intimate partner violence and zero otherwise. The three heterogeneity variables are all dummies, indicating whether the respondent resides in rural area or not, whether there was a Women Justice Center in the district, and whether the respondent was currently employed at the time of the interview. All regressions include a full set of control variables as in column (4) of Table 3.1. Standard errors are clustered at the district level in all specifications.
*** $p < 0.01$, ** $p < 0.05$, *$p < 0.1$. Standard errors in parentheses.

Table 4.4: Earthquake and Intimate Partner Violence: Channels

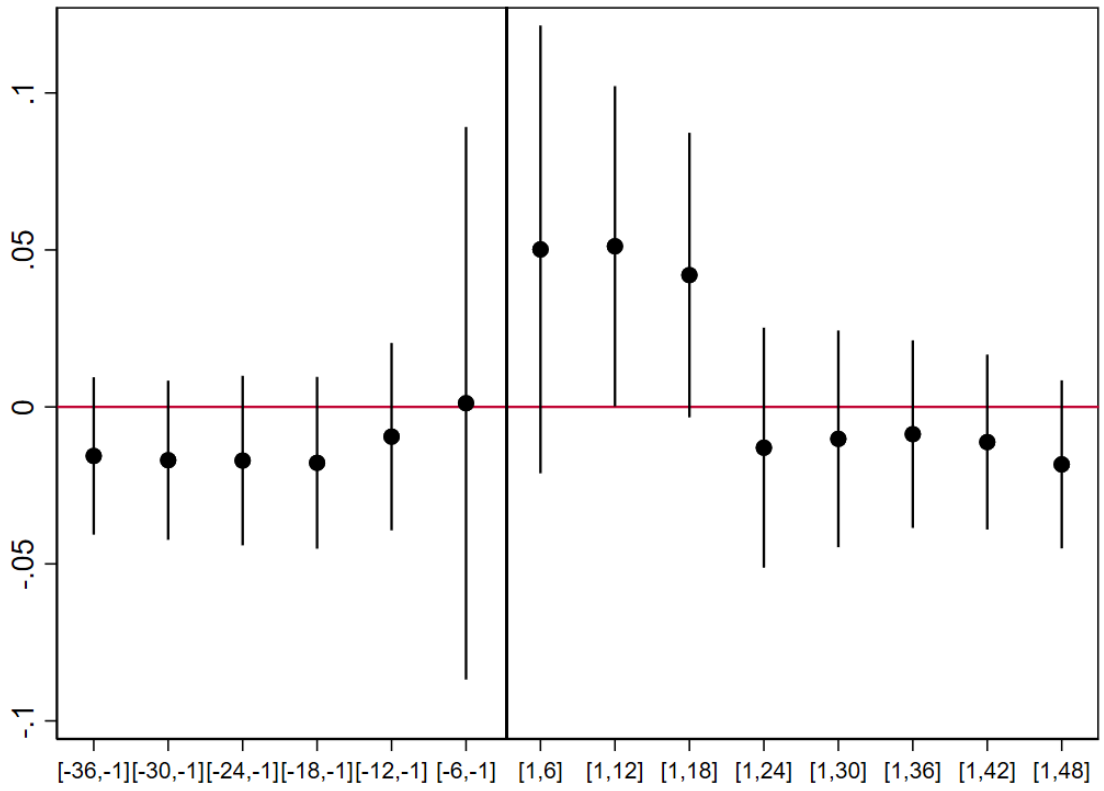
	Partner Drinks (0/1)	Frequent Drunk (0/1)	Male in the hh (0/1)	Male Dominance (0/1)	Joint Decision (0/1)	hh Wealth (continuous)
	(1)	(2)	(3)	(4)	(5)	(6)
EQ $\geq$ 7.0 12M (0/1)	0.151***	0.018	0.032**	0.094*	-0.039	-0.075
	(0.053)	(0.058)	(0.015)	(0.054)	(0.024)	(0.139)
Year FE, District FE, and Department Trend	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓
Observations	27,290	35,901	43,106	43,106	37,843	27,290
R^2	0.079	0.078	0.037	0.132	0.112	0.739

Notes: The table reports the estimation results of equation (4.1) using linear probability model. Respondents are females aged 15 to 49 that have ever been married or cohabiting. The sample is restricted to females that have been living in the same location for at least 3 years. The outcome variables in columns (1)-(5) are all dummies and are specified in each column. All regressions include a full set of control variables as in column (4) of Table 3.1. Standard errors are clustered at the district level in all specifications.

*** $p < 0.01$, ** $p < 0.05$, *$p < 0.1$. Standard errors in parentheses.

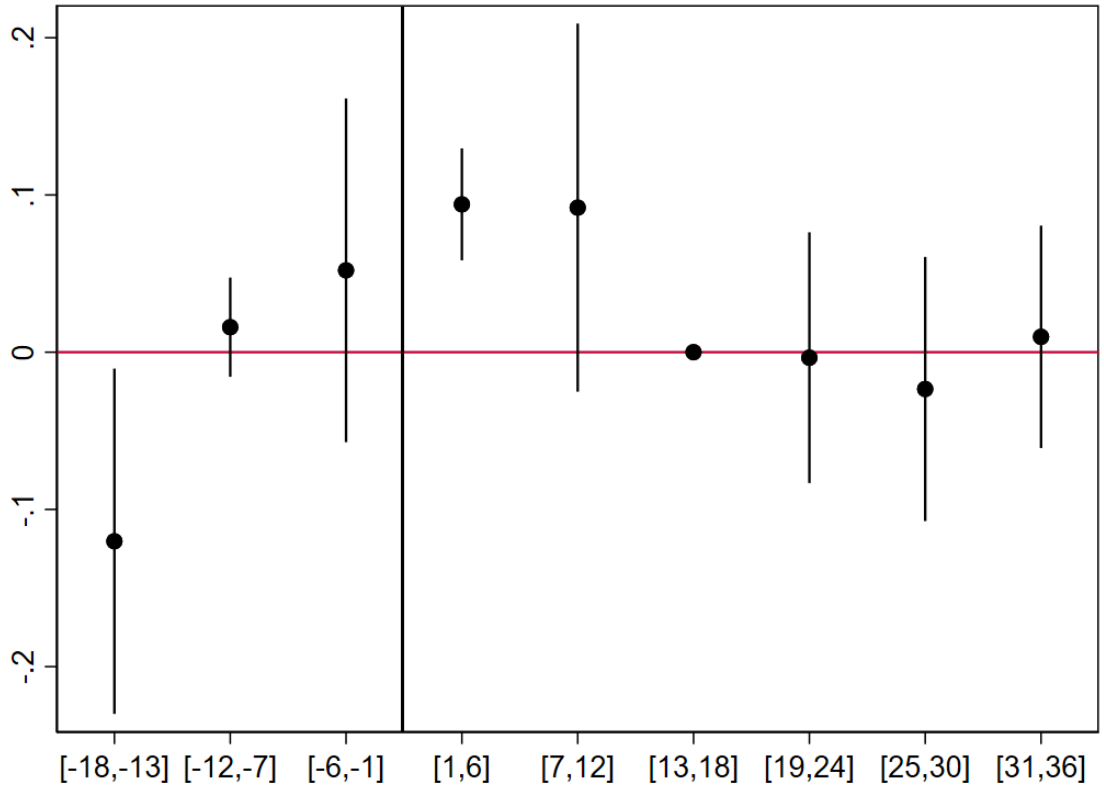
4.10 Chapter 4 - Appendix

Figure 4.A.1: Earthquakes and Intimate Partner Violence: Alternative Threshold at 6.0 MMI Scale



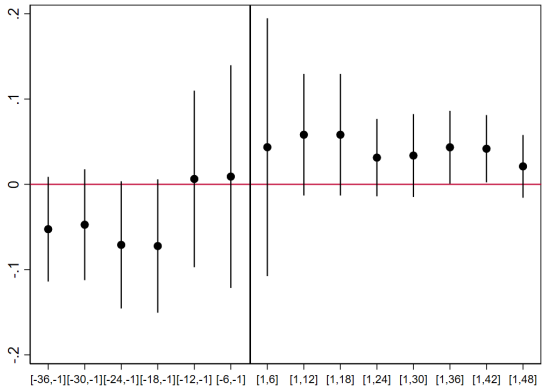
Notes: Plot of leads and lags of earthquake coefficients estimated separately for earthquakes equal to or above 6.0 in the MMI scale. 90 percent confidence intervals.

Figure 4.A.2: Earthquakes and Intimate Partner Violence: Alternative Definition

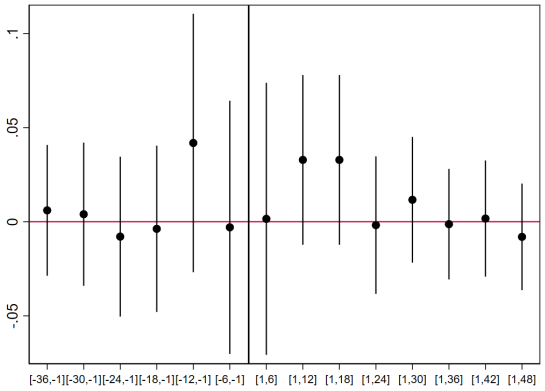


Notes: Plot of leads and lags of earthquake coefficients estimated separately for earthquakes equal to or above 7.0 in the MMI scale. 90 percent confidence intervals.

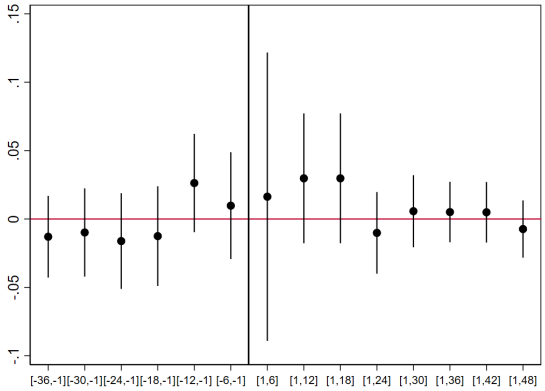
Figure 4.A.3: Earthquakes and Intimate Partner Violence: Alternative Outcomes



(a) Mental Violence



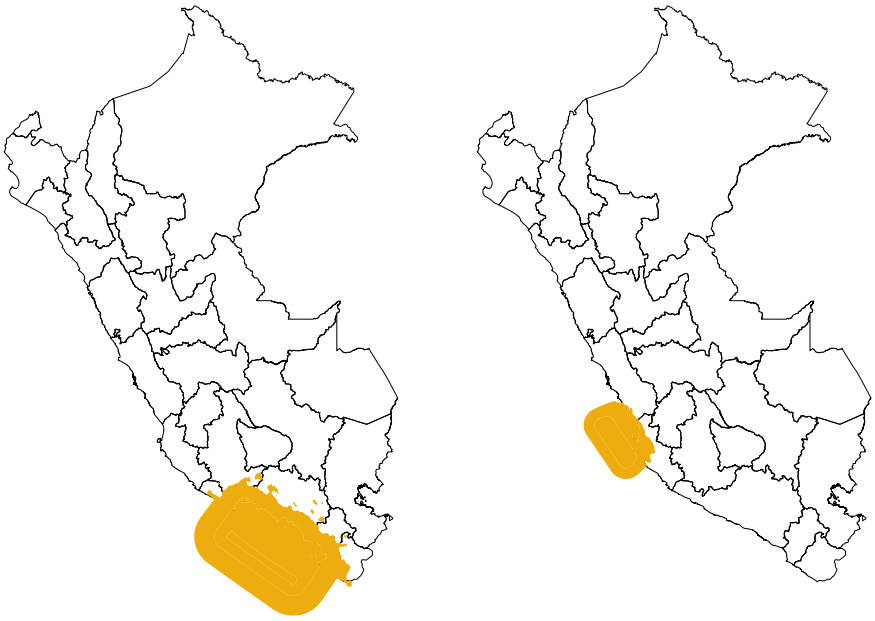
(b) Severe Violence



(c) Sexual Violence

Notes: Plot of leads and lags of earthquake coefficients estimated separately for earthquakes equal to or above 7.0 in the MMI scale. 90 percent confidence intervals.

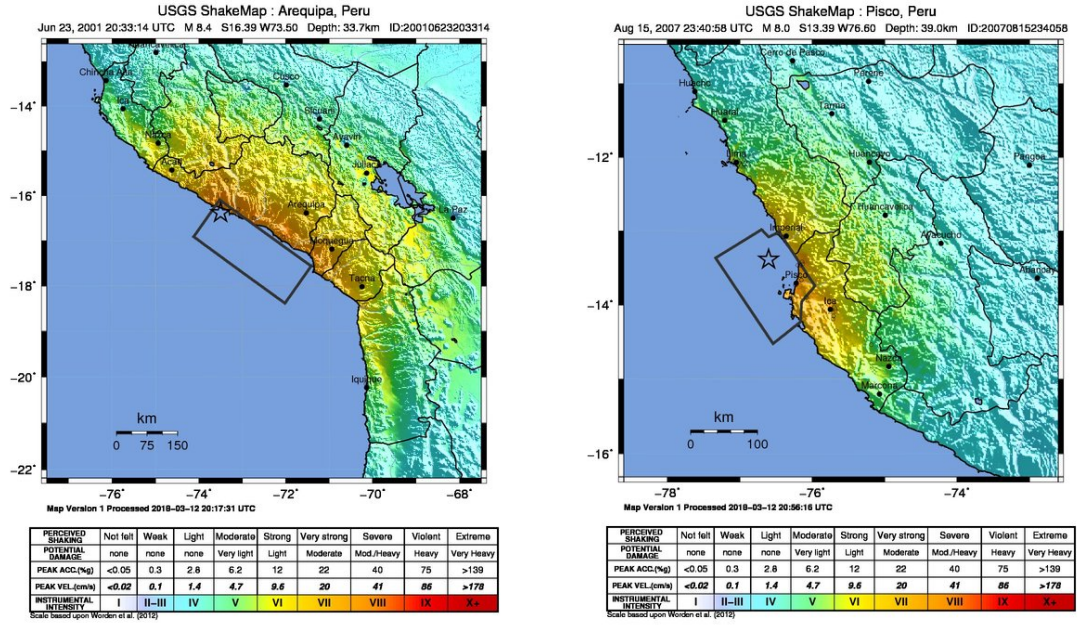
Figure 4.A.4: Peruvian Departments and 2001 and 2007 Earthquakes



(a) 2001 (b) 2007

The figures show the earthquake polygons equal to or above 7.0 for the earthquakes that happened in Peru in 2001 and 2007. Additionally, the borders of the Peruvian departments are shown.

Figure 4.A.5: ShakeMaps 2001 and 2007 Earthquakes



(a) 2001 (b) 2007

Source: United States Geological Survey (USGS). The figures show the ShakeMaps for the earthquakes that happened in Peru in 2001 and 2007. The legend shows the intensity and the correspondent perceived shaking and potential damage.

Figure 4.A.6: The Modified Mercalli Intensity Scale: USGS

Intensity	Shaking	Description/Damage
I	Not felt	Not felt except by a very few under especially favorable conditions.
II	Weak	Felt only by a few persons at rest,especially on upper floors of buildings.
III	Weak	Felt quite noticeably by persons indoors, especially on upper floors of buildings. Many people do not recognize it as an earthquake. Standing motor cars may rock slightly. Vibrations similar to the passing of a truck. Duration estimated.
IV	Light	Felt indoors by many, outdoors by few during the day. At night, some awakened. Dishes, windows, doors disturbed; walls make cracking sound. Sensation like heavy truck striking building. Standing motor cars rocked noticeably.
V	Moderate	Felt by nearly everyone; many awakened. Some dishes, windows broken. Unstable objects overturned. Pendulum clocks may stop.
VI	Strong	Felt by all, many frightened. Some heavy furniture moved; a few instances of fallen plaster. Damage slight.
VII	Very strong	Damage negligible in buildings of good design and construction; slight to moderate in well-built ordinary structures; considerable damage in poorly built or badly designed structures; some chimneys broken.
VIII	Severe	Damage slight in specially designed structures; considerable damage in ordinary substantial buildings with partial collapse. Damage great in poorly built structures. Fall of chimneys, factory stacks, columns, monuments, walls. Heavy furniture overturned.
IX	Violent	Damage considerable in specially designed structures; well-designed frame structures thrown out of plumb. Damage great in substantial buildings, with partial collapse. Buildings shifted off foundations.
X	Extreme	Some well-built wooden structures destroyed; most masonry and frame structures destroyed with foundations. Rails bent.

Source: United States Geological Survey (USGS). The figures shows the intensity and the correspondent perceived shaking and potential damage for each of the intensity categories.

Table 4.A.1: Earthquake Description

Date	**MMI Scale**	**Affected Departments**	**Damage**
23-Jun-01	VIII	Arequipa, Moquegua, Tacna, and Ayacucho	A tsunami followed the earthquake, spreading destruction along the Peruvian coast. According to the US Agency for International Development (USAID), 81 people were killed, more than 2,700 were injured and 220,000 were affected by the earthquake. The agency estimates that close to 36,000 houses were damaged and around 25,000 were destroyed.
15-Aug-07	VIII	Lima, and Ica	According to information from the USGS, more than 514 people were killed, more than a thousand were injured and around 39,000 buildings were destroyed. Additionally, transport routes were damaged and communications and power supply were compromised.

Notes: MMI stands for Modified Mercalli intensity, and USGS stands for United States Geological Survey.

Table 4.A.2: Description of variables

Variable	**Description**
IPV	Respondent reports that she ever experienced any type of IPV (partner pushed her, shook her, threw something at her, slapped her, punched her, kicked her or dragged her).
Age	Respondent's age
Squared age	Respondent's squared age
Age difference	Age difference between the respondent and the partner
Respondent has no education	Respondent has completed 0 years of education
Respondent has primary education	Respondent has completed between 1 and 6 years of education
Respondent has secondary education	Respondent has completed between 7 and 12 years of education
Respondent has tertiary education	Respondent has completed above 12 years of education
Partner has no education	Partner has completed 0 years of education
Partner has primary education	Partner has completed between 1 and 6 years of education
Partner has secondary education	Partner has completed between 7 and 12 years of education
Partner has tertiary education	Partner has completed above 12 years of education
Working	The respondent is employed at the time of the interview
Earthquake $\geq 6.0_{c,(t-i,t)}$	DHS cluster in which the respondent lives was affected by an earthquake above or equal to 6.0 in the MMI scale up to i months prior to the interview. For the aggregate analysis: the department was affected by an earthquake above or equal to 6.0 in the MMI scale within the past i months.
Earthquake $\geq 7.0_{c,(t-i,t)}$	DHS cluster in which the respondent lives was affected by an earthquake above or equal to 7.0 in the MMI scale up to i months prior to the interview. For the aggregate analysis: the department was affected by an earthquake above or equal to 7.0 in the MMI scale within the past i months.
Δ IPV (%)	Percentage change in the number of cases of violence against women registered in the Women Justice Centers at each department.
ln(Population)	logarithm of the population of each department

Table 4.A.3: Earthquakes and Intimate Partner Violence: Difference-in-difference

	Post 2001	Post 2007
	(1)	(2)
Post × Earthquake ≥ 7.0	-0.006	0.098**
	(0.029)	(0.043)
Earthquake ≥ 7.0	0.033	-0.040
	(0.042)	(0.076)
Year FE, District FE, Department Trend	✓	✓
Controls	✓	✓
Observations	43,110	21,566
R^2	0.081	0.086

Notes: The table reports the estimation results of equation (4.1) using linear probability model. Respondents are females aged 15 to 49 that have ever been married or cohabiting. The sample is restricted to females that have been living in the same location for at least 3 years. The outcome variable is a binary indicator which equals to one if the respondent reports having suffered any type of intimate partner violence and zero otherwise. The three heterogeneity variables are all dummies, indicating whether the respondent resides in rural area or not, whether there was a Women Justice Center in the district, and whether the respondent was currently employed at the time of the interview. All regressions include a full set of control variables as in column (4) of Table 3.1. Standard errors are clustered at the district level in all specifications.

*** $p < 0.01$, ** $p < 0.05$, *$p < 0.1$. Standard errors in parentheses.

Table 4.A.4: Channel Variables and Intimate Partner Violence

Channel Variable:	Dep. Variable: Intimate Partner Violence (0/1)					
	Partner Drinks (0/1)	Frequent Drunk (0/1)	Male in the hh (0/1)	Male Dominance (0/1)	Joint Decision (0/1)	hh Wealth (continuous)
	(1)	(2)	(3)	(4)	(5)	(6)
Channel Variable	0.171***	-0.032***	0.011	0.046***	-0.099***	-0.004
	(0.007)	(0.007)	(0.011)	(0.005)	(0.008)	(0.006)
Year FE, District FE, and Department Trend	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓
Observations	27,290	35,901	43,106	43,106	37,843	27,290
R^2	0.105	0.081	0.074	0.076	0.079	0.083

Notes: The table reports the estimation results of equation (4.1) using linear probability model. Respondents are females aged 15 to 49 that have ever been married or cohabiting. The sample is restricted to females that have been living in the same location for at least 3 years. The outcome variable in all columns is a binary indicator which equals to one if the respondent reports having suffered any type of intimate partner violence and zero otherwise. All regressions include a full set of control variables as in column (4) of Table 3.1. Standard errors are clustered at the district level in all specifications.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors in parentheses.

Chapter 5

Financial Education and Spillover Effects: Experimental Evidence from Uganda*

with:
Jana Hamdan
Tim Kaiser
Lukas Menkhoff

* We thank Mountains of the Moon University as well as Filder Aryemo and her team at Gaplink Uganda for support in data collection from 2018 onwards. We further thank Arianna Antezza, Kaja Rupieper and Markus Rafael Paffhausen for excellent research assistance. We further thank the seminar and conference participants at the University of Göttingen, the PhD Workshop of the German Economic Association (VfS) 2021, NEUDC 2021 for valuable discussions and comments. All remaining errors are ours. This research was funded by the German Research Foundation through the Research Training Group (RTG) 1723 "Globalization and Development" and CRC TRR 190 (project number 280092119). The study was registered in the AEA RCT Registry and the unique identifying number is AEARCTR-0006407.

Abstract

We study the impact of a business and financial literacy program and its spillover effects on about 2,000 micro-entrepreneurs in rural Uganda using a two-stage randomized saturation experiment. We first randomize the program at the trading center level, and then randomize the share of treated micro-entrepreneurs in each cluster. More than one year after the program, treated entrepreneurs in treated clusters are more likely to have formal savings, higher formal savings, more and larger business investments, more and higher savings in their mobile money account and more often using mobile money as a means of payment. At the same time, we do not find evidence of positive spillover effects on untreated entrepreneurs in treated clusters. Instead, the estimated spillover coefficients often show negative (albeit insignificant) signs.

JEL Codes: D14 (Personal Finance), C93 (Field Experiments), G53 (Financial Literacy), O12 (Microeconomic Analyses of Economic Development).
Keywords: Financial Education, Micro-entrepreneurs, RCT, Household Finance, Uganda, Mobile Money.

5.1 Introduction

The measurement of treatment effects keeps the treated group typically so distinct from a control group that there will be no interference. This design allows to infer a clear causal treatment effect. However, in many situations there are people who belong to neither group and still may be affected indirectly by the treatment. Such spillovers can occur, for example, when some people in a village get treated while others do not. The non-treated may indirectly profit from the intervention, they may remain unaffected or they may be even worse off. In any case, potential spillovers need to be considered when painting the full picture of an impact evaluation (Angelucci and Di Maro, 2016).

In this study we focus on potential spillovers of a five-hour financial education program with micro-entrepreneurs in rural Uganda. In general, the sign and size of spillovers is expected to be ambiguous: positive spillovers occur due to learning from neighbors, either because they support each other or because they observe each other as competitors. Negative spillovers occur if treated micro-entrepreneurs improve their business at the disadvantage of competing neighbors. The direction of spillover effects is a potentially important element of interventions as positive spillovers would create a positive externality and thus improve the overall effectiveness of a training and vice versa. Thus we do not only estimate the direction and size of spillover effects but also aim for better understanding whether there are potential determinants that influence spillovers. If there are such determinants, they could be considered and used *ex ante* in order to improve effectiveness.

Therefore we study the impact of an active learning financial education training on the financial behavior of micro-entrepreneurs covering five areas of outcomes, (1) budgeting and record keeping, (2) saving, (3) debt management, (4) business investment and (5) money transfer. The fifth area on money transfer puts an emphasis on the use of mobile money as this innovation plays a prominent role in our setting. In Uganda, more than 16 million active users make 240 million transactions per month in 2019, worth more than six trillion Ugandan Shilling (Bank of Uganda, 2021). Covering mobile money is quite new to financial education RCTs and reflects its increasing role in Uganda and the developing world in general (Suri, 2017).

In a field experiment with 2,177 micro-entrepreneurs who are located across 108 trading centers, we apply a two-stage random saturation design. First, we randomize the trading centers into 54 treated and 54 control clusters with 1,207 and 970 micro-entrepreneurs respectively, and we estimate the conventional causal Intent-to-Treat

(ITT) effects of the intervention, using an ANCOVA specification. Second, the share of micro-entrepreneurs invited to the training is varied randomly. Thus, not necessarily all micro-entrepreneurs get invited to the training in a treated trading center. This creates a spillover group in 36 treated clusters. In total, 861 micro-entrepreneurs are invited to our training, and considered targeted, while 346 are not, even though they are based in a treated trading center. As we also collect full information about this non-invited group within treated trading centers, spillover effects can be measured. Finally, we plan to make use of the random variation of the share of treated but this is still work in progress and not reported here.

This study provides three major results: first, we find that the relatively short intervention generates several intended significant changes; second, the treatment works to also impact the use of mobile money; and third, spillovers are largely insignificant with many negative coefficient signs. In detail: the financial education lasts about five hours and covers during this period in total five topic areas. This setting is motivated by relying on the "active learning" approach which has proved to be a relatively effective way of teaching and which shows promising effects also in the field of financial education (Kaiser and Menkhoff, 2018). We find significant effects on targeted micro-entrepreneurs in two of the four areas (1) to (4), i.e. in the areas of saving and investment in the 13-19 months time period between intervention and endline. This result is qualitatively in line with Kaiser and Menkhoff (2018) who also find effects in these two areas.

Regarding effects on the use of mobile money, this is a new outcome area and we are among the first to evaluate a financial literacy training that addresses this topic. As the use of mobile money has steadily increased over the last years, a change spurred by the COVID-19 pandemic (GSMA, 2021), it seems encouraging that we also find some significant results here. Those who are directly targeted are 5.4 percentage points more likely to have savings in the mobile money account compared to the control group, where 18.4 percent save with mobile money at endline. The amount of savings in their mobile money accounts also increases by 52 percent relative to control. Moreover, the targeted micro-entrepreneurs use several new payment functions of mobile money more actively, but not regarding using mobile money for money transfers.

The third major result regards spillovers, i.e. potential effects in treated clusters on the non-treated. The estimated spillover effects often show negative but small and insignificant coefficients. While some share of negative coefficients may be regarded largely as random, the design of the intervention would rather suggest positive coefficients. Clear effects of the training are in increased savings and a more intensive use

of mobile money, both of which should stimulate non-treated and does not set any incentive for the spillover group to save rather less and use mobile money less. Still, this is what we observe. Further effort seems justified to better understand whether this result is indeed random or whether there is some so far not understood determinant for negative spillovers.

Our study complements three strands of literature, i.e. the analysis of (1) financial education, (2) mobile money and (3) spillover effects. Financial education has been recognized as an important element of education for the whole population as evidenced by respective programs supported by the OECD worldwide (OECD, 2020a). Micro-entrepreneurs in developing countries are a particular target of these efforts as improved financial behavior may not only contribute to their family welfare but may also stimulate economic growth. While the use of improved financial behavior is undisputed, there was some discussion whether financial trainings are able to realize this ambition (Fernandes et al., 2014). However, most recent evidence clearly shows that financial education, even when evaluated via RCTs and when corrected for publication bias, has a positive causal effect on financial knowledge and downstream financial behaviors (Miller et al., 2015; Kaiser et al., 2021). However, the challenge remains to make best use of resources by increasing effectiveness of financial education. There are several ways that have been suggested, such as using (very costly) individual counseling (Carpena et al., 2019), to rely on goal setting (Carpena et al., 2019; Grohmann et al., 2022), to use an entertainment environment (Berg and Zia, 2017), to time education at a "teachable moment" (Doi et al., 2014; Kaiser and Menkhoff, 2017) or to refer to the "active learning" approach. Here we rely on the two latter elements, i.e. the training is designed to meet the purposes of these rural micro-entrepreneurs and the way of delivery follows the concept of "active learning".

This study also adds directly to the research on mobile money. We offer this as one out of five areas of the training here and see improved financial behavior also in this respect. While there is lot of evidence showing the impact of mobile money on risk sharing and household welfare (Jack and Suri, 2014; Munyegera and Matsumoto, 2016; Riley, 2018) and financial inclusion (Hamdan et al., 2021), there is little explicit evidence on the success of addressing mobile money in a financial education training. Chiwaula et al. (2020) combine a financial literacy and mobile money training with reminders among community savers in Malawi, finding a positive effect on mobile money savings four months after the intervention. However, they cannot differentiate the impact of the training from the reminders and focus on a selected group who already save.

Finally, this study contributes to the impact evaluations measuring spillover effects. An early and prominent example is Haushofer and Shapiro (2016) who find null effects for economic outcomes among non-recipients of unconditional cash transfers in treatment villages, and argue that they can thus focus on within-village treatment effects and use them as a control group. However, the authors document a negative spillover effect on consumption and food security in the long term (Haushofer and Shapiro, 2018). Haliassos et al. (2020) find positive spillovers of financially knowledgeable neighbors on stock market participation - they use a natural experiment, exploiting the variation of average financial literacy in Swedish neighborhoods, to study the long-term impact on refugees, who were randomly allocated to these neighborhoods. Closer to our design is McKenzie and Puerto (2021) who do not find spillovers from business education in their study on small female-led businesses in Kenya. The positive effects of the assignment to a training do not come along with significant spillover effects on the untreated businesses in the same markets.

The paper is organized as follows. In the next section, we describe the training, our study sample and design. We then discuss the empirical strategy, summary statistics, balance, program take-up and attrition in Section 5.3. In Section 5.4, we discuss our treatment outcomes and spillover findings. This is followed by some robustness exercises in Section 5.5. We conclude in Section 5.6.

5.2 Sample and Study Design

5.2.1 Financial Education Program

We evaluate the effectiveness of a five-hour financial education program. It includes a total of six parts – one introductory session and five main sessions with the following topics: (1) personal financial management, (2) saving, (3) debt management, (4) business investment and (5) money transfer. The training is organized in horizontal order, thus participants start with the introductory session, and then immediately move from session (1) through until session (5), spending up to one hour per session.

In all training sessions, we employ an active learning method (Kaiser and Menkhoff, 2018), where numerous interactive activities are included in the curriculum to engage the participants in the learning process and to encourage discussions. We include, for example, picture stories, real-life case studies as exercises, small-group discussions and the evaluation of individual situations. These features promote paying attention and learning from peers. The goals of the training include, for example, learning

how to budget and financially plan for themselves and their business, being able to distinguish needs from wants, promoting savings and knowing the different ways of saving, selecting and managing debt, making an investment plan and increasing business investments, and understanding the risks and costs associated with several ways to transfer money (e.g., mobile money). The participant materials and the trainer guide are made available in Appendix Section 5.8.1. A training team consists of one master trainer who holds the introductory session, and then five trainers who are responsible of one session each, further generating variety in the training.

We implement this program in function halls, school classrooms or suitable outside areas in or near the center of the trading centers (or villages) in our study region. On average, the training lasts five hours, with a group consisting of 12 participants. The training is delivered by one licensed financial literacy trainer, called the master trainer, as well as five local undergraduate students from the public Mountain of the Moon University (MMU), who act as trainers, being responsible for one session each. All trainers take part in a three-day workshop, where they become familiar with the curriculum and the learning goals, and equipped with the active learning method. All trainers' teaching qualifications are verified by holding mock training sessions and a pilot training.

The entire team that partly alternates during the intervention included two female and two male master trainers. It further consists of 16 female trainers, 11 male trainers, five female research assistants, and one male research assistant.

Participants of the training receive a cash payout of 10,000 UGX (equivalent to 2.80 USD) after completion of all training sessions. According to the baseline data, the average monthly income from the respondent's business and other sources was about 380,000 UGX (about 110 USD). Thus they earn an average of 12,700-15,200 UGX per day when assuming 25-30 working days. Therefore, the opportunity cost of attending a training was likely compensated for with the disbursed expense allowance, but not sufficient to cover a whole day of typical income.

5.2.2 Study Setting, Sample and Timeline

Uganda is similar to many developing countries in its stage of financial inclusion: mobile money services boosted basic access to financial accounts significantly in recent years, however, the access to formal financial institutions remains low and a gender gap persists (Demirgüç-Kunt et al., 2020). The 46-million population grows at an annual rate of 3.3 percent, 46 percent of Ugandans are below the age of 15 and 75

percent live in rural areas (WDI, 2020). Only one in four employed Ugandans is in some form of wage employment, most work for themselves or for their families in low-quality jobs; less than 10 percent are in formal employment (Merotto, 2019). Thus, policy efforts to support the growth of micro-enterprises are needed to create jobs and improve their quality – "whether as entrepreneurs or as employees, rural youth need access to finance, technology, skills, and assets" (Merotto, 2019). Meanwhile, the Ugandan National Financial Inclusion Strategy (NFIS) 2017-2022 unites efforts to improve financial security via financial inclusion. Its goals include increasing women's use of financial services, individual emergency savings and, more generally, consumer protection (Bank of Uganda, 2017).

In this context, we implement our study in the rural Kabarole district in the Western region of Uganda, see Figure 5.1. We take this district as it used to be known to people during baseline and still appears in several government statistics. Officially, the southern part of this district and, i.e. the former county of "Bunyangabu" has become a new independent district in 2017. The population of Kabarole, as we still use it, in 2020 is estimated to be about 528,000 (UBOS, 2021). The majority of the population reside outside the main town, Fort Portal.

Our sample is composed of micro-entrepreneurs from this area. They typically run their family enterprises, usually in the form of small shops, in so-called trading centers. These are shopping streets usually lined along the main road of villages and small towns. There is no official documentation of the existing trading centers in the district, and we mapped the 113 trading centers within Kabarole with the local authorities in November 2018, prior to the baseline survey. The baseline survey was implemented from February to April 2019 in all the mapped locations, and a total of 108 trading centers were included in the study. Five mapped trading centers were not included as no open business was found during the baseline visit. Figure 5.2 displays the locations of the 108 trading centers on a map by treatment status.

During the baseline survey, we conducted face-to-face interviews with a total of 2,177 micro-entrepreneurs[1], with an average of about 21 respondents in each trading center.[2] We estimate that this sample covers roughly 40 percent of all micro-enterprises

[1] About 62 percent of interviewed micro-entrepreneurs have small retail or wholesale shops, and 28 percent run a service business, mostly hair dressing salons and restaurants. The remaining 10 percent are in manufacturing, mostly in furniture making or metal goods production.

[2] Prior to the baseline survey, we piloted the interviews in Rugombe, a community in the adjacent Kyenjojo district. There, we also piloted the intervention, the financial education training, in August 2019. We originally interviewed 2,223 respondents, however, during our endline survey, it became clear that 46 interviews could not be verified *ex post*. These are excluded from this analysis.

in the rural areas of the Kabarole, as we counted 5,478 shop units that seemed to be active in some way. All those with an open small business were, in principal, eligible to participate in the study.

Our sampling procedure is designed to interview a sufficiently large share of micro-entrepreneurs in each trading center. As smaller trading centers are typically less busy, many business owners also work in subsistence farming in proximity to their shop, and close their business for part of the day. In order to interview enough micro-entrepreneurs in smaller trading centers, we approach every open business in trading centers with up to 100 businesses and invite them to participate in our survey.

This strategy results in covering 47 percent of shops in these areas, implying that the majority of shops (54 percent) were not open during our visiting time.

In the nine largest trading centers, those with more than 100 micro-enterprises, we randomly approach every third open business. This brings us to the outcome of interviewing 35 percent of micro-entrepreneurs in these clusters. Thus, in a big (small) trading centers, we under-(over-)sampled. Due to variation in the number of open businesses, there is substantial variation in the sampling rate across sizes, ranging from 20 to 100 percent across the trading centers. Overall, the average sampling rate is 46.1 percent per cluster. In order to control for these varying sampling probabilities by trading center size, we incorporate inverse sampling weights in our later analysis. However, in our robustness tests, we find that these do not meaningfully change any findings.

On average, a baseline survey takes about 45 minutes and is compensated with 4,900 UGX (equivalent to 1.40 USD). It collects extensive information on the financial situation and type of the businesses and the financial decision-making of the owners, particularly their saving, borrowing, transfer behaviors and the use of mobile money. Two survey experiments are incorporated into the baseline survey as well, one on the willingness to pay of mobile money for making transfers, the other measuring the risk preferences.

We re-visit the 54 treatment trading centers in August and September 2019 and implement our financial education intervention, thus four to six months after the baseline.[3] Targeted baseline respondents are invited in advance via phone calls. On the day of the intervention, those persons are called again prior to the beginning of the training. In addition, they are informed of and encouraged to take part in the training by the local council, elected officials who run villages (the lowest political

[3] During this period, we updated the contact details of those micro-entrepreneurs who participated in the training.

administrative units in Uganda), on the day of the training as well. The spillover group is not directly invited to take part in the training. In the ITT analysis, these participants belong to the spillover group because they are not invited. This may tentatively increase positive spillovers.

The endline survey is conducted in two steps, about 13-19 months after the intervention. The first round of collection is implemented using phone surveys from October to December 2020. Afterwards, in March and April 2021, we follow up with hitherto unreached participants via face-to-face interviews. On average, the endline survey takes on average about 30 minutes and is compensated with 6,400 UGX (equivalent to 1.80 USD). Overall, we reach 90.72 percent (N = 1,975) of baseline participants.[4]

In the baseline and the intervention phases, the study is implemented in cooperation with the our mentioned partner, the MMU. The endline data collection is implemented by Gaplink Uganda, a Kampala-based independent research company. Their team is kept unaware of the experimental design, and is provided with only the contact details of the participants and local councils. All surveys are conducted in the local language Rutooro or optional in the official language English. The responses are recorded on tablets using SurveyCTO.

5.2.3 Two-staged Randomization

The randomization is conducted in two stages, and the process is displayed in Figure 5.3. In the first stage, 54 out of a total of 108 clusters (trading centers) are randomly allocated to the treatment group and the other 54 are allocated to the control group. The randomization is conducted via a stratification strategy based on two variables: the mobile money account ownership rates and financial literacy levels at baseline.

We first create three strata based on the cluster-level average share of mobile money account owners, splitting the 108 clusters into low, middle and high mobile money ownership rates, each with 36 clusters. Afterwards, within each stratum, we further create six sub-strata depending on the average financial literacy score.[5] We obtain a total of 18 strata, with six clusters within each stratum. We perform the randomization into control and treatment group in a 1:1 ratio, using a random seed in Stata.

[4] The phone survey reaches 72.86 percent of the baseline sample.

[5] The two variables are chosen because participants with different baseline financial literacy levels and mobile money adoption may be significantly differently impacted by the financial education program. In order to understand possible heterogeneity effects, it is, for instance, necessary to have sufficient low financial literacy respondents in every group.

In the second step, among the 54 treated clusters, we further randomize the cluster-level treatment intensity with probability 1/3. In 18 clusters we invite 100 percent of the baseline respondents to participate in the financial education program, in another 18 clusters we invite 75 percent, and in the remaining 18 clusters we invite 50 percent. Within clusters, the subset of micro-entrepreneurs who are directly invited has been randomized as well. Overall, 861 respondents are directly targeted to the take part in the training, thus 39.55 percent of the baseline sample. The treatment intensity is varied to assess the magnitude of spillovers.

We registered our trial and pre-analysis plan in the American Economic Association's registry for randomized controlled trials (ID AEARCTR-0006407). Blinding of participants and the trainers is not feasible. However, it is possible to maintain blinding during the endline survey data collection for the evaluation of the training program.

5.3 Empirical Strategy and Summary Statistics

5.3.1 Empirical Strategy

Randomization of the treatment assignment allows us to establish a credible counterfactual condition and therefore allows for a causal estimate of the program's impact. In this sense, we first estimate the the Intent-to-Treat (ITT) effects of the treatment status at the individual level using the following ANCOVA regression:

$$Y_{isef} = \alpha_0 + \beta_1 Target_i + \beta_2 Spillover_i + \omega Y_{i0} + \lambda_s + \eta_e + \delta_f + \epsilon_{isef} \tag{5.1}$$

where, on the left hand side, Y_{ivsef} is the outcome of interest of micro-entrepreneur i located (at baseline) in trading center v and strata s, and interviewed by enumerator e using the interview method f (either phone or face-to-face) in the endline survey. On the right hand side, α_0 is a constant. $Target_i$ is a dummy, equaling to one if the micro-entrepreneur i is directly invited to the training. $Spillover_i$ is a dummy as well, indicating whether the individual i is in the treated clusters but was not directly invited to the training. The coefficients, β_1 and β_2, measure the impact of the program on those who are targeted and the spillover group, respectively, and the reference group is those individuals in the control trading centers. Y_{i0} is the outcome value at baseline. λ_s and η_e refer to the strata fixed effects and enumerator fixed effects, respectively. δ_f is a dummy indicating whether the individual is interviewed

via a phone or face-to-face interview in the endline survey. The standard errors, ϵ_{ivsef}, are clustered at the baseline trading center level. The results are weighted by the inverse probability of being sampled at baseline.

Second, we estimate the cluster-level ITT effects, thus the average effect of being assigned to the treated trading centers compared to being in the control trading centers, using the following equation:

$$Y_{ivsef} = \alpha + \beta Treatment_v + \omega Y_{i0} + \lambda_s + \eta_e + \delta_f + \epsilon_{ivsef} \tag{5.2}$$

Here, on the right hand side, $Treatment_v$ is a dummy variable, equaling to one if there is a financial training program delivered in trading center v during the intervention. Thus, the coefficient β represents the ITT estimator and measures the total average effect of the financial training program.

Lastly, we use an instrumental variable approach to provide the local average treatment effects (LATE) of training participation on our outcomes, instrumenting attendance with the invitation to participate. The first stage is simply:

$$Participation_{isef} = \alpha_0 + \beta_1 Target_i + \beta_2 Spillover_i + \omega Y_{i0} + \lambda_s + \eta_e + \delta_f + \epsilon_{isef} \tag{5.3}$$

and

$$No\,Participation_{isef} = \alpha_0 + \beta_1 Target_i + \beta_2 Spillover_i + \omega Y_{i0} + \lambda_s + \eta_e + \delta_f + \epsilon_{isef} \tag{5.4}$$

where $Participation_{isef}$ (or $No\,Participation_{isef}$) is a dummy equal to one if the micro-entrepreneur i participated (or did not participate) in the training. We then estimate the equation:

$$Y_{isef} = \alpha + \beta_1 \hat{Partici}pation_i + \beta_2 Non\,\hat{Partici}pation_i + \omega Y_{i0} + \lambda_s + \eta_e + \delta_f + \epsilon_{isef} \tag{5.5}$$

by two-stage least-squares, instrumenting the participation status using the invitation status in the treated clusters. The coefficients β_1 and β_2 measure the impact of the financial training on the compliers and defiers, i.e. those who participated because of being invited, and those who did not participate in the training but also was not

invited.

5.3.2 Summary Statistics and Balance

Overall, we evaluate the impact of our intervention on a total of 20 variables. Table 5.1 displays summary statistics of these 20 outcome variables as well as selected socioeconomic characteristics to verify the orthogonality of randomization at baseline. These variables are described in Appendix Tables 5.B.1 and 5.B.2. Table 5.1 reveals that 64 percent of our sampled micro-entrepreneurs are female. Moreover, the respondents are on average 34 years old and have some secondary education.

As can be seen from Table 5.1, for most variables there are no baseline difference across the three different treatment arms. Specifically, there are no significant differences in the areas of household characteristics and mobile money use. However, there are minor differences along other domains. For example, targeted respondents are more likely male, more financially literate and have more savings overall but less likely to have mobile money savings than spillover respondents from the treated clusters. Importantly, there are no significant differences between targeted respondents compared to the control group across the observed characteristics.

However, some variables differ between the spillover and the control group: the spillover individuals have lower financial literacy, and report on average lower savings, loans, and investments. Overall, among the 87 difference in means tests for outcome variables, only seven are statistically significant at the five percent level. However, ordered logit estimates and a Wald test indicate a joint significance of variables at p-value 0.069. In the Appendix, Table 5.B.3 shows balance at baseline for the endline sample, excluding the attriters.

5.3.3 Take-up and Attrition

The compliance rate upon invitation to the training among our baseline sample was 67.02 percent. Overall, few variables correlate with take-up of the training, see Appendix Table 5.C.1. Within the treated clusters, participation rates are lower on average among the more educated compared to those with fewer years of education. This may reflect that they think there is less to learn. Among those who are untargeted and thus serve as the spillover group, we find that wealth proxied by number of assets correlates negatively with training participation. This shows a tendency that those less well-off are more interested to participate in a financial training. Similar to McKenzie

and Puerto (2021), we find a positive coefficient of previous training participation for take-up, however, our coefficient is not statistically significant.

The endline survey tracks 90.72 percent of the baseline sample, making attrition relatively low between the two waves, apart by more than 1.5 years. However, respondents in the treatment clusters are 3.34 percent less likely to attrite than respondents in the control clusters (92.21 percent tracking rate in the treatment clusters against 88.87 percent in the control clusters). To understand why differential attrition occurs, we exploit the different reasons for attrition that have been documented in the endline survey, and we find suggestions that the differential attrition is due to the differences in business survival.

There are five reasons for attrition: relocation, death, decline, sickness and imprisonment, and cases where the reason is unknown. As shown in Table 5.B.4 in the Appendix, relocation causes about 71 percent of the attrition cases, and only ten participants declined to participate in the endline survey. Those participants who relocated closed down the original business as interviewed in the baseline, and migrated to another region. In addition, about half of them indicated directly that their business collapsed and they had to shut down the business and leave. According to the baseline data, those who relocated had lower business performance initially, compared to those participated in the endline survey, see Table 5.C.2.

When we regress the experimental design indicators on the attrition status and control for strata fixed effects, we find that participants in the control clusters are 3.8 percentage points more likely to attrite than those in the treatment clusters - see Table 5.2, column (1). According to column (2), there is no difference in coefficients between being directly targeted or in the spillover group. Thus, the differential attrition is not driven by treatment intensity, but appears to be occurring on the trading center level. Columns (3)-(8) show that the difference in attrition is almost completely driven by the difference in relocation (or business closure) by treatment status of the trading centers. As we can see from column (3), those in the treated clusters were 3.0 percentage points less likely to close down their business and relocate between baseline and endline, compared to those in the control clusters. There are only minor differences in the probability of declining to participate in the endline survey across different arms, or other reasons. Due to the overall small differences in attrition, we do not further consider this issue in the following analysis.

5.4 Results

We show that the relatively short financial training program has significant effects in several directions. Based on the positive effects, we further examine whether spillover effects exist.

5.4.1 Main Treatment Effects

The presentation of intervention outcomes follows the RCT-registration. The outcomes are grouped into five planned intervention areas: (i) savings, (ii) debt management, (iii) business investment, (iv) (budgeting and) record keeping, and (v) money transfer. Regarding areas (i) to (iv) we use two variables each, a binary and a numerical variable. An exception is the missing binary information formal debt (because information on volume is missing in the endline survey), and the binary information on business investment which was missing in the registration. Regarding money transfer, the training addresses the various functions that can be used via mobile money and emphasizes also the transaction costs, in particular for smaller money transfers. Overall, we have 20 outcomes for which we compare the changes in the group of the invited micro-entrepreneurs relative to the control group (the ITT). Results in Table 5.3 show for the first four intervention areas almost consistently positive coefficient signs. Exceptions are the outcomes "any debt taking" and "debt volume" where the training aims for controlling debt taking so that the negative insignificant coefficients are consistent with the expectation. The effects on formal savings are highly significant; being assigned to the training increases the likelihood of having savings by 4.6 percentages points (or 26 percent from 17.6 to 22.2 percentage points). In addition, being directly targeted leads to a 75 percent increase in formal savings. The effects on business investments are significant at the 10 percent level, increasing the likelihood of any investments by 4 percentage points and the investment amount by 64 percent. By contrast, we do not find significant results with regard to overall savings, debt taking, and record keeping.

Regarding money transfers, the coefficient signs for the variety of uses of mobile money are mostly positive (Table 5.4). Only the transfers to individuals at other places have a negative sign, which may be due to teaching cost awareness for relatively high transaction costs when making small transfers. However, the coefficients are not significant. Similarly, the active use of mobile money has increased but not to a significant degree. The remaining five out of the nine outcomes are affected significantly, i.e. two savings variables and three variables on using of mobile money for

business. The economic effect is sizable, and the incidence of having mobile money savings increases by 5.4 percentage points (or 29 percent from 18.4 to 23.8 percentage points). Being assigned to the training further increases mobile money savings by 52 percent. In addition, mobile money use for business purposes, such as payments in general, and specifically for suppliers increases in its likelihood by 4.2 and 5.2 percentage points, respectively. In the control group, respondents report that almost 4 percent of their customers paid on average using mobile money. This increases by 1.2 percentage point (i.e., 30 percent) for the targeted micro-entrepreneurs.

The positive result on formal savings and mobile money savings in combination with no significant increase in overall savings motivates a closer look. As informal savings increase slightly (but not significantly) more for the targeted than the control group, it is the semi-formal savings via SACCO's or ROSCA's where savings increase less for the targeted, that cause the insignificant result for total savings. As semi-formal institutions actually hold the majority of savings, we see in the first place a relative shift in savings from semi-formal to mobile money and informal savings, while formal savings decline in absolute volumes. The group being directly targeted, however, shifts more to mobile money and withdraws less from formal savings which is in line with the training objective, in order to avoid paying the withdraw fee, which could be high when withdrawing small amounts. There may be also some relative increase in total savings, but this is not significant in our main specification.

Overall, the evaluated financial training program causes several changes in financial behavior. In line with a similar training conducted in a neighboring area in rural Uganda (Kaiser and Menkhoff, 2018), there are effects on savings and investment, but no significant effects on debt and record keeping. The emphasis on mobile money in the current intervention increases the use of specific functions of mobile money. These results provide the basis to examine the potential effect of spillovers.

5.4.2 Spillovers of the Training Program

The spillover effects of this financial training program are measured as the effect on those individuals in the treated trading centers who are not directly invited to the training. This randomized saturation design is chosen purposely to being able to identify spillovers. Results are provided in Tables 5.3 and 5.4 in the rows below the treatment effects on those assigned to the training. We find that coefficients on the spillover effect are often negative; this applies to nine out of those 16 coefficients which are positive for the directly targeted, while the remaining four outcomes are negative

but in line with intervention expectations. This pattern of nine negative, seven positive and four debatable coefficients indicates that positive spillovers do not dominate in any case. Among all the 20 spillover coefficients in Tables 5.3 and 5.4, only two are significant, showing negative spillover effects of the intervention on the number and volume of transfers via mobile money. While we observe some negative spillovers, it is not fully clear why the spillover group reduces the use of mobile money transfers even stronger than the control group. In general it is not obvious why the coefficients are often negative in our setting. The most conventional explanation of a crowding out effect is not really convincing here. There may be two anecdotal explanations. First, it may be that we have different forms of reporting behavior across groups. Those who are not invited to the training and serve as the spillover group might feel discriminated and thus under-report their financial behaviors. Second, there may be a shift in financial transactions from the spillover to the targeted group within the treated trading centers. For example, when the spillover micro-entrepreneurs have a need of transferring money to friends or family in another village, they may ask those who are targeted to finish the transaction as they may be expected to know better how to make a mobile money transfer. However, this mechanism is hard to test with our data. Thus we mainly report the spillover effect but remain silent on a plausible explanation.

Another view on potential spillovers is the test of equivalence of the intervention effects for the targeted and the spillover group as shown in the third row of Tables 5.3 and 5.4. Largely in line with the earlier presented results, most significant treatment effects go along with a significant difference from the spillover group. Finally, the largely insignificant effect on the spillover group dilute the treatment effects so strongly that measuring the total causal intervention effect in treated markets relative to control markets eliminates all significant treatment effects (see robustness section).

In general, we interpret these results as clear evidence against positive spillovers. The many negative signs may even raise concerns whether under specific circumstances negative spillovers do occur. We leave this investigation to further research.

5.5 Robustness

We check the robustness of our results from three perspectives. First, we analyze the overall ITT effect by comparing both directly targeted and spillover individuals in the treated clusters with those in the control clusters (see equation (5.2)). Second, we estimate the impact on the compliers and defiers, i.e. those who were invited

to and participated in the training, and those who were not invited and did not participate (see equations (5.3)-(5.5)). Finally, we replicate the main analysis and the two robustness checks without weight adjustments, similar to McKenzie and Puerto (2021), and compare the results with those with weight adjustments.

The cluster-level ITT results are set out in Panel 1 in Appendix Tables 5.D.1 and 5.D.2. As we see from the previous section, the spillover effects are small and sometimes even negative. Thus, we expect that the combined effect on the targeted group and the spillover group should be smaller than the effect on the targeted group and may become insignificant. Indeed, we could see that most of the cluster-level ITT effects are insignificant. Only five coefficients remain significant at the ten percent level. These include that those micro-entrepreneurs are more likely to save formally and have larger amount of formal savings. Moreover, they would have smaller loans and increase the incidence of mobile money savings as well as using mobile money to pay suppliers.

In a second analysis we look at a more narrowly defined treatment group. We estimate the impact on the compliers and defiers only using a two-staged least squares. The remaining micro-entrepreneurs within the treated trading centers are thus excluded this analysis, i.e. those who were invited but did not participate in the training, and those who were assigned as spillover group but participated in the training. The expectation is that the program has even larger treatment effects on the complied participants than on the directly targeted (as covered in the main analysis, see Tables 5.3 and 5.4). As participation is partly endogenous due to self-selection, we follow Banerjee et al. (2007) and use an instrumental variable (IV) approach by using the invitation as instrument in the first stage regression.

The LATE results are shown in Panel 2 of Appendix Tables 5.D.1 and 5.D.2. Compared to the results from the main analysis, we could see that the program has a more significant and larger positive impact on the compliers. For instance, apart from the positive effect on formal savings and business investments, we also observe that participants profit from the training by significantly increasing their total amount of savings compared to the control group.

Finally, as the third robustness check we re-estimate our main analysis but do not perform re-weighting accounting for the variation in sampling probability and sampling probability at baseline. The results are set out in Appendix Tables 5.D.3 and 5.D.4. We find that the results remain mostly stable, and our results are robust to different weight adjustments.

5.6 Conclusion

This study evaluates the effect of a financial education training on three objectives. First, the main focus is assessing the potential spillover of the training within markets. This is an issue that is little researched but of high policy relevance. Second, the analysis of spillovers requires that there is any impact of the underlying treatment, so that we choose a treatment format, "active learning", which is known to generate relatively high impact. Third, we incorporate the proper use of mobile money into the training as this innovation increases rapidly in importance.

We get clear results on all three objectives. To start with the basis, we find that the financial education intervention has effects similarly to those found in a related training by Kaiser and Menkhoff (2018): the treatment significantly increases formal savings and investments more than one year after the intervention, but it does not improve record keeping. The insignificant result on debt taking may be in line with the objective to avoid "unnecessary" borrowing. Second, regarding mobile money we find that the training succeeds in increasing to use the savings function of mobile money and the use of several payment functions. The effect on (expensive) transfers is rather a reduction which is consistent with the training ambitions. Finally, regarding spillovers we do not find a significant effect. We are somewhat surprised, however, by the often negative coefficients among the spillover group, a result that deserves further attention.

5.7 Chapter 5 - Figures and Tables

Figure 5.1: The Study Setting – the Rural District Kabarole in Western Uganda

Figure 5.2: Distribution of Trading Centers by Treatment Status within the District

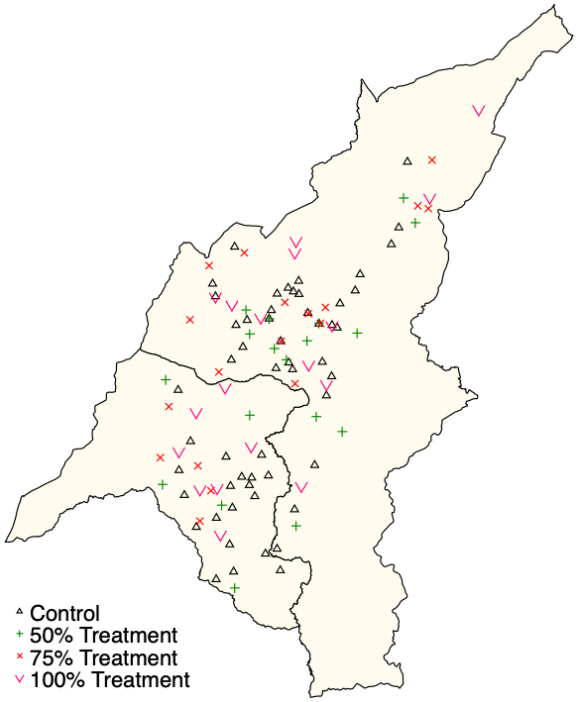
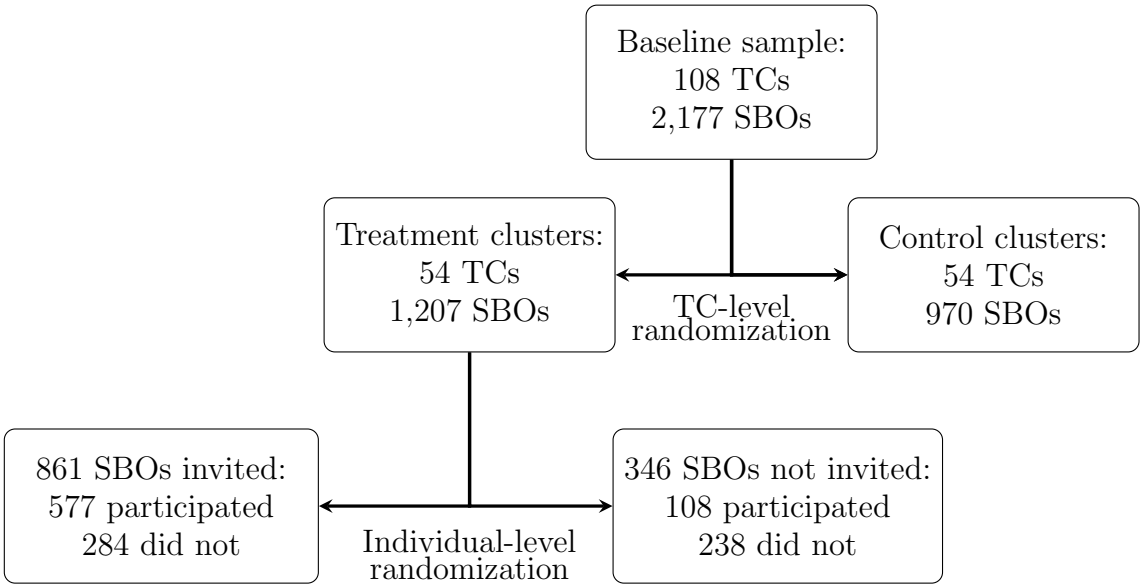


Figure 5.3: Randomization Process



Note: The figure displays the two-stage randomization procedure and the sample sizes of the treatment arms. SBO stands for small business owners and TC means trading center.

Table 5.1: Balance at Baseline

	Means (Std. Dev.)			Differences (p-values)		
	Control group (C)	Untargeted (T_1)	Targeted (T_2)	$T_1 - C$	$T_2 - C$	$T_2 - T_1$
Female (=1)	0.642	0.665	0.619	0.022	-0.023	-0.046*
	(0.480)	(0.473)	(0.486)	(0.252)	(0.661)	(0.053)
Age	33.364	34.258	34.410	0.894	1.046	0.152
	(11.424)	(11.561)	(11.966)	(0.281)	(0.135)	(0.804)
Married (=1)	0.486	0.494	0.499	0.009	0.014	0.005
	(0.500)	(0.501)	(0.500)	(0.953)	(0.608)	(0.489)
Years of Education	8.801	8.948	8.758	0.147	-0.043	-0.190
	(4.504)	(4.651)	(4.659)	(0.789)	(0.531)	(0.355)
Financial Literacy (0-7)	3.653	3.491	3.667	-0.161**	0.014	0.175*
	(1.628)	(1.639)	(1.655)	(0.033)	(0.772)	(0.064)
Risk Tolerance (0-10)	5.240	5.338	5.249	0.098	0.008	-0.090
	(2.757)	(2.781)	(2.715)	(0.986)	(0.656)	(0.798)
Household Size	3.933	4.055	4.033	0.122	0.100	-0.022
	(2.418)	(2.377)	(2.532)	(0.872)	(0.679)	(0.772)
# Assets	36.586	37.139	37.682	0.553	1.096	0.543
	(17.773)	(16.864)	(18.003)	(0.821)	(0.684)	(0.614)
Household Consumption (UGX)	481659.517	498349.871	497264.306	16690.354	15604.789	-1085.566
	(337331.357)	(331841.261)	(332508.241)	(0.830)	(0.862)	(0.873)
Saving (=1)	0.779	0.720	0.777	-0.060*	-0.002	0.057*
	(0.415)	(0.450)	(0.416)	(0.065)	(0.854)	(0.058)
ln Saving	9.763	9.049	9.836	-0.714**	0.072	0.786**
	(5.380)	(5.809)	(5.440)	(0.033)	(0.929)	(0.029)
Formal Saving (=1)	0.137	0.162	0.168	0.025	0.031	0.007
	(0.344)	(0.369)	(0.374)	(0.147)	(0.137)	(0.737)
ln Formal Saving	1.821	2.212	2.252	0.391	0.431	0.040
	(4.646)	(5.075)	(5.046)	(0.111)	(0.146)	(0.880)
Loan (=1)	0.346	0.301	0.351	-0.046**	0.004	0.050
	(0.476)	(0.459)	(0.477)	(0.042)	(0.749)	(0.260)
ln Loan	4.409	3.777	4.475	-0.632**	0.066	0.698
	(6.113)	(5.824)	(6.154)	(0.019)	(0.792)	(0.197)
Formal Loan (=1)	0.048	0.049	0.042	0.001	-0.007	-0.007
	(0.215)	(0.216)	(0.200)	(0.776)	(0.349)	(0.389)
Invest (=1)	0.883	0.868	0.880	-0.015*	-0.003	0.012
	(0.321)	(0.339)	(0.325)	(0.076)	(0.145)	(0.762)
ln Invest	11.854	11.701	11.870	-0.153**	0.016	0.169
	(4.478)	(4.731)	(4.551)	(0.028)	(0.210)	(0.653)
Record (=1)	0.223	0.214	0.220	-0.009	-0.003	0.006
	(0.416)	(0.411)	(0.414)	(0.920)	(0.753)	(0.698)
Separate Personal (=1)	0.239	0.223	0.237	-0.017	-0.002	0.014
	(0.427)	(0.417)	(0.425)	(0.436)	(0.712)	(0.621)
MM Active (=1)	0.486	0.457	0.504	-0.029	0.018	0.047
	(0.500)	(0.499)	(0.500)	(0.679)	(0.720)	(0.507)
# MM Active (0-4)	0.668	0.647	0.683	-0.021	0.015	0.036
	(0.794)	(0.811)	(0.787)	(0.994)	(0.731)	(0.862)
MM Saving (=1)	0.099	0.095	0.077	-0.004	-0.022	-0.019**
	(0.299)	(0.294)	(0.266)	(0.693)	(0.103)	(0.031)
ln MM Saving	1.114	1.024	0.882	-0.090	-0.232	-0.142*
	(3.397)	(3.250)	(3.083)	(0.897)	(0.134)	(0.071)
MM Transfer (=1)	0.291	0.263	0.288	-0.028	-0.003	0.025
	(0.454)	(0.441)	(0.453)	(0.912)	(0.587)	(0.980)
ln MM Transfer	3.327	2.919	3.239	-0.408	-0.088	0.320
	(5.257)	(4.946)	(5.146)	(0.681)	(0.457)	(0.934)
MM Payment (=1)	0.277	0.289	0.314	0.012	0.036	0.025
	(0.448)	(0.454)	(0.464)	(0.991)	(0.568)	(0.644)
MM Supplier (=1)	0.033	0.040	0.035	0.007	0.002	-0.006
	(0.179)	(0.197)	(0.183)	(0.817)	(0.808)	(0.743)
MM Customer Share	0.005	0.006	0.007	0.000	0.002	0.002
	(0.048)	(0.055)	(0.052)	(0.598)	(0.902)	(0.640)
Observations	970	346	861	1316	1831	1207

Note: The table displays the summary statistics for the control group and by individual treatment status within the treated trading centers (N=2177) at baseline in 2019. In addition, the differences in means and, in parentheses, the p-values of linear regressions are shown. These include inverse sampling probability weights, controlling for strata fixed effects and clustering standard errors at the trading center level. *** p<0.01, ** p<0.05, * p<0.1.

Table 5.2: Attrition, Relocation, Decline, and Other Reasons by Treatment Status

	Attrition (=1)		Relocation (=1)		Decline (=1)		Others (=1)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treated Villages	-0.038**		-0.030*		-0.003		-0.005*	
	(0.016)		(0.017)		(0.003)		(0.003)	
Assigned to Training		-0.038**		-0.030*		-0.001		-0.006**
		(0.017)		(0.017)		(0.003)		(0.003)
Spillover Group		-0.038*		-0.029		-0.006**		-0.003
		(0.019)		(0.019)		(0.003)		(0.004)
Observations	2177	2177	2177	2177	2177	2177	2177	2177
R^2	0.015	0.015	0.017	0.017	0.008	0.009	0.009	0.009
Strata FE	✓	✓	✓	✓	✓	✓	✓	✓

Note: The table shows linear regression results with the binary dependent variable attrition in columns (1)-(2) and reasons for attrition in columns (3)-(8). The regression models include strata fixed effects. Weighted by sampling weights and experimental design weights. "Others" include illnesses, imprisonment and death. Robust standard errors (clustered at the TC level) in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 5.3: Effects on Savings, Loans, Investment and Business Formality

	Saving (=1)	ln Saving	Formal Saving (=1)	ln Formal Saving	Loan (=1)	ln Loan	Formal Loan (=1)	Invest (=1)	ln Invest	Record (=1)	Separate Personal (=1)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Control Mean	0.849	10.869	0.176	2.171	0.791	10.078	0.085	0.763	10.149	0.506	0.589
ITT and Spillover Effects											
Assigned to Training	0.015	0.154	0.046**	0.560**	-0.023	-0.358	0.022	0.040*	0.495*	0.011	0.018
	(0.020)	(0.265)	(0.020)	(0.263)	(0.021)	(0.246)	(0.014)	(0.021)	(0.288)	(0.028)	(0.026)
Spillover Group	-0.017	-0.341	-0.002	0.006	-0.021	-0.455	0.000	0.001	-0.038	0.010	-0.013
	(0.022)	(0.314)	(0.021)	(0.279)	(0.024)	(0.290)	(0.018)	(0.028)	(0.376)	(0.031)	(0.022)
T = Spillover ($p-value$)	0.100	0.076	0.047	0.073	0.954	0.720	0.255	0.097	0.099	0.970	0.136
Observations	1975	1975	1975	1975	1975	1975	1975	1975	1975	1975	1975
R^2	0.075	0.083	0.101	0.088	0.061	0.071	0.059	0.157	0.163	0.118	0.112

Note: Table shows regression results, controlling for the lagged variable, dummies for missing values, enumerator FE, face-to-face interview dummy, and strata FE. Weighted by sampling weights and experimental design weights. Standard errors are clustered at the trading center level and displayed in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 5.4: Effects on Mobile Money Use

	MM Active (=1)	# MM Active (0-4)	MM Saving (=1)	ln MM Saving	MM Transfer (=1)	ln MM Transfer	MM Payment (=1)	MM Supplier (=1)	MM Customer Share
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Control Mean	0.912	1.839	0.184	2.030	0.777	9.348	0.789	0.379	0.037
ITT and Spillover Effects									
Assigned to Training	0.016	0.058	0.054***	0.519**	-0.019	-0.177	0.042*	0.052*	0.012*
	(0.016)	(0.052)	(0.020)	(0.234)	(0.027)	(0.332)	(0.025)	(0.028)	(0.007)
Spillover Group	-0.011	-0.094	-0.015	-0.157	-0.085***	-0.972***	0.023	0.026	0.005
	(0.019)	(0.058)	(0.027)	(0.313)	(0.030)	(0.364)	(0.024)	(0.036)	(0.008)
T = Spillover ($p-value$)	0.127	0.008	0.011	0.031	0.017	0.014	0.464	0.564	0.383
Observations	1975	1975	1975	1975	1975	1975	1975	1975	1975
R^2	0.101	0.137	0.070	0.072	0.078	0.086	0.162	0.144	0.131

Note: Table shows regression results, controlling for the lagged variable, dummies for missing values, enumerator FE, face-to-face interview dummy, and strata FE. Weighted by sampling weights and experimental design weights. Standard errors are clustered at the trading center level and displayed in parentheses. *** p<0.01, ** p<0.05, * p<0.1

5.8 Chapter 5 - Appendix

5.8.1 Training Materials

The training materials for the Financial Literacy Ring (FLIR) can be accessed here on a public Google Drive. FLIR was originally developed to promote financial inclusion in Uganda, designed by the Bank of Uganda (BoU) and German Development Cooperation (GIZ). We further developed this program in order to include up-to-date information and mobile money insights. Our materials comprise the explanatory trainer manual and participant learning materials for the introductory session and the five distinct stations. One stations' learning materials are presented below as an example.

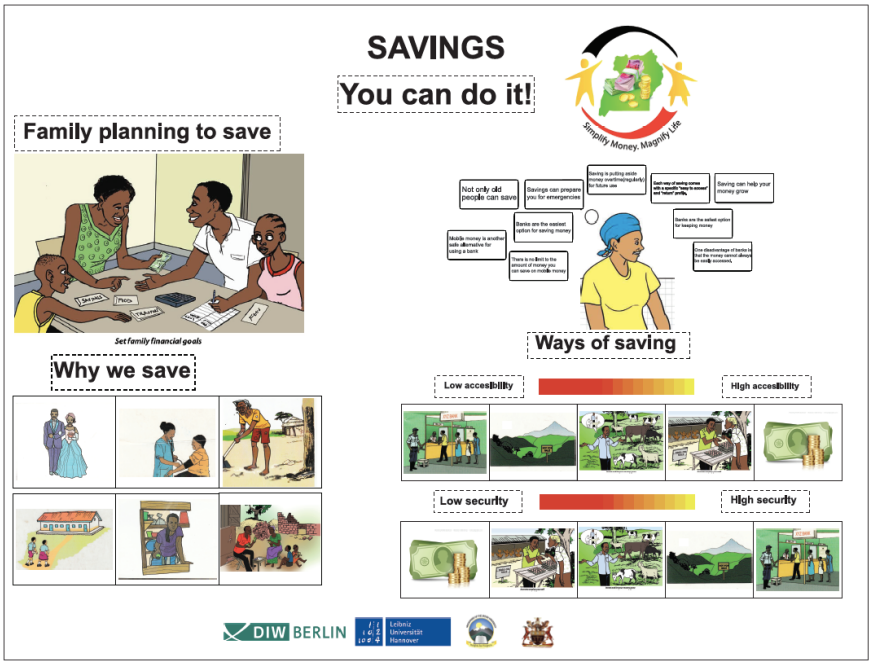


Figure 5.A.1: Learning materials for the FLIR station on savings

5.8.2 Additional Details

The following table describes the balance and outcome variables that are retrieved at baseline.

Table 5.B.1: Variable Description

Variable	Type	Level	Description
Female (=1)	Binary	Individual	1 if respondent is female; 0 if male
Age	Continuous	Individual	Age in years
Married (=1)	Binary	Individual	1 if respondent is married; 0 if not married
Years of Education	Continuous	Individual	Education in years
Financial Literacy (0-7)	Continuous	Individual	Number of correct financial knowledge questions.
Risk Tolerance (0-10)	Continuous	Individual	Likert scale item on "Are you generally a person who is fully prepared to take risks or do you try to avoid taking risk?" 0 if completely unwilling to take risks; 10 if fully prepared to take risk.
Household Size	Continuous	Household	Number of people in household including respondent. A household comprises all the people who normally live and eat meals together in a home.
# Assets	Continuous	Household	Sum over a list of standard items owned by the household without livestock.
Household Consumption (UGX)	Continuous	Household	The value of all consumption within past four weeks. Includes food inside and outside the household, toiletries, water, rent, electricity, clothing, school fees, medical costs, leisure expenses, other necessities etc. [In Ugandan Shilling; top coded at 99%]
Saving (=1)	Binary	Individual	Currently any money saved in informal places, in a SACCO or in a ROSCA, in a commercial bank account, or a mobile money account (for example MoKash). [Yes/No]
ln Saving	Continuous	Individual	Amount of money saved in total across aforementioned places at the moment. [In Ugandan Shilling; top coded at 99% and logarithmised]
Formal Saving (=1)	Binary	Individual	Currently any money saved in a commercial bank account.
ln Formal Saving	Continuous	Individual	Amount of money saved in a commercial bank account at the moment. [In Ugandan Shilling; top coded at 99% and logarithmised]
Loan (=1)	Binary	Individual	Currently any loan outstanding with a family member or friend or moneylender, an ASCA, SACCO or a microfinance institution, a commercial bank or mobile money.
ln Loan	Continuous	Individual	Amount of money currently outstanding in total across aforementioned places. [In Ugandan Shilling; top coded at 99% and logarithmised]
Formal Loan (=1)	Binary	Individual	Currently any loan outstanding with a commercial bank.

Table 5.B.2: Variable Description (contin.)

Variable	Type	Level	Description
Invest (=1)	Binary	Individual	Any money invested in business in total during the past 12 months. Investments could be new equipment, restocking done on top of regular restocking, new furniture or signs for advertising but do not include regular expenditures for buying new supplies or stock.
ln Invest	Continuous	Individual	Amount of money invested in business in total during the past 12 months. [In Ugandan Shilling; top coded at 99% and logarithmised]
Record (=1)	Binary	Individual	Keeping a log or record of expenses and revenues for any businesses. [Yes/No]
Separate Personal (=1)	Binary	Individual	Keeping money separate for business and personal finances. [Yes/No]
MM Active (=1)	Binary	Individual	1 if respondent qualifies as an active mobile money user; if not. We define active mobile money use via four conditions, out of which at least one must be fulfilled: current mobile money savings or loans; any mobile money transfers made in the past three months; or ever made mobile payments directed at other businesses.
# MM Active	Continuous	Individual	Number of different mobile money services used out of the aforementioned four conditions.
MM Saving (=1)	Binary	Individual	Currently any money saved in a mobile money account (for example MoKash). [Yes/No]
ln MM Saving	Continuous	Individual	Amount of money saved in a mobile money account (for example MoKash) in total at the moment. [In Ugandan Shilling; top coded at 99% and logarithmised]
MM Transfer (=1)	Binary	Individual	Transferred money to people outside the respondent's village using a mobile money account during the past three months. [Yes/No]
ln MM Transfer	Continuous	Individual	Amount of money transferred using a mobile money account in total over the past three months. [In Ugandan Shilling; top coded at 99% and logarithmised]
MM Payment (=1)	Binary	Individual	Ever made a payment using the mobile phone. This means a payment to another business, supplier, employee or utility provider, for example with MTN MoMoPay or Airtel Money Pay. [Yes/No]
MM Supplier (=1)	Binary	Individual	Uses mobile money to pay for supplies (for example raw materials, items for resale). [Yes/No]
MM Customer Share	Continuous	Individual	Percentage of customers paid using mobile money in the past three months. [0-100]

Table 5.B.3: Balance at Baseline of Endline Sample

	Means (Std. Dev.)			Differences (p-values)		
	Control group (C)	Untargeted (T_1)	Targeted (T_2)	$T_1 - C$	$T_2 - C$	$T_2 - T_1$
Female (=1)	0.642	0.659	0.614	0.018	-0.027	-0.045*
	(0.480)	(0.475)	(0.487)	(0.231)	(0.618)	(0.088)
Age	33.685	34.486	34.739	0.801	1.054	0.253
	(11.423)	(11.426)	(12.126)	(0.505)	(0.209)	(0.694)
Married (=1)	0.497	0.503	0.512	0.007	0.015	0.009
	(0.500)	(0.501)	(0.500)	(0.897)	(0.496)	(0.556)
Years of Education	8.839	9.009	8.748	0.171	-0.091	-0.262
	(4.536)	(4.646)	(4.644)	(0.807)	(0.313)	(0.246)
Financial Literacy (0-7)	3.667	3.541	3.694	-0.126**	0.027	0.153
	(1.636)	(1.622)	(1.651)	(0.043)	(0.884)	(0.114)
Risk Tolerance (0-10)	5.332	5.369	5.230	0.037	-0.102	-0.139
	(2.710)	(2.780)	(2.711)	(0.570)	(0.504)	(0.997)
Household Size	3.969	4.109	4.087	0.141	0.118	-0.022
	(2.410)	(2.421)	(2.554)	(0.887)	(0.684)	(0.834)
# Assets	37.158	37.650	37.972	0.492	0.814	0.322
	(17.940)	(16.941)	(17.741)	(0.843)	(0.956)	(0.740)
Household Consumption (UGX)	492362.546	509102.897	502875.741	16740.351	10513.194	-6227.157
	(340355.106)	(336025.522)	(338630.008)	(0.895)	(0.572)	(0.695)
Saving (=1)	0.796	0.741	0.779	-0.055*	-0.017	0.039
	(0.403)	(0.439)	(0.415)	(0.075)	(0.281)	(0.148)
ln Saving	10.014	9.368	9.885	-0.646**	-0.129	0.517*
	(5.264)	(5.708)	(5.428)	(0.042)	(0.286)	(0.096)
Formal Saving (=1)	0.147	0.169	0.170	0.021	0.023	0.001
	(0.355)	(0.375)	(0.376)	(0.240)	(0.322)	(0.869)
ln Formal Saving	1.949	2.314	2.267	0.365	0.318	-0.047
	(4.771)	(5.177)	(5.049)	(0.173)	(0.333)	(0.961)
Loan (=1)	0.355	0.319	0.348	-0.036*	-0.007	0.029
	(0.479)	(0.467)	(0.477)	(0.085)	(0.951)	(0.691)
ln Loan	4.522	4.010	4.449	-0.513**	-0.074	0.439
	(6.154)	(5.927)	(6.153)	(0.037)	(0.912)	(0.544)
Formal Loan (=1)	0.049	0.053	0.043	,004	-0.006	-0.010
	(0.215)	(0.225)	(0.203)	(0.751)	(0.317)	(0.315)
Invest (=1)	0.885	0.863	0.879	-0.022*	-0.006	0.016
	(0.319)	(0.344)	(0.326)	(0.056)	(0.118)	(0.648)
ln Invest	11.888	11.652	11.861	-0.236**	-0.027	0.209
	(4.457)	(4.804)	(4.578)	(0.018)	(0.159)	(0.570)
Record (=1)	0.229	0.216	0.211	-0.013	-0.018	-0.005
	(0.420)	(0.412)	(0.408)	(0.963)	(0.363)	(0.597)
Separate Personal (=1)	0.241	0.219	0.230	-0.023	-0.012	0.011
	(0.428)	(0.414)	(0.421)	(0.314)	(0.490)	(0.585)
MM Active (=1)	0.494	0.481	0.515	-0.013	0.020	0.033
	(0.500)	(0.500)	(0.500)	(0.829)	(0.722)	(0.646)
# MM Active (0-4)	0.682	0.681	0.700	-0.001	0.018	0.019
	(0.797)	(0.818)	(0.793)	(0.889)	(0.794)	(0.674)
MM Saving (=1)	0.099	0.103	0.078	0.005	-0.020	-0.025**
	(0.298)	(0.305)	(0.269)	(0.397)	(0.211)	(0.015)
ln MM Saving	1.103	1.108	0.901	0.005	-0.202	-0.207**
	(3.370)	(3.366)	(3.117)	(0.509)	(0.290)	(0.037)
MM Transfer (=1)	0.297	0.275	0.295	-0.022	-0.002	0.020
	(0.457)	(0.447)	(0.456)	(0.971)	(0.680)	(0.981)
ln MM Transfer	3.415	3.068	3.320	-0.347	-0.096	0.252
	(5.316)	(5.039)	(5.184)	(0.795)	(0.508)	(0.989)
MM Payment (=1)	0.285	0.303	0.322	0.018	0.036	0.018
	(0.452)	(0.460)	(0.467)	(0.841)	(0.640)	(0.762)
MM Supplier (=1)	0.035	0.041	0.037	0.006	0.002	-0.004
	(0.183)	(0.198)	(0.188)	(0.959)	(0.835)	(0.740)
MM Customer Share	0.006	0.006	0.008	0.000	0.002	0.002
	(0.051)	(0.057)	(0.054)	(0.520)	(0.990)	(0.682)
Observations	862	320	793	1182	1655	1113

Note: The table displays the summary statistics at baseline in 2019 for the control group and by individual treatment status within the treated trading centers for those respondents who were reached again during endline (N=1975). In addition, the differences in means and, in parentheses, the p-values of linear regressions are shown. These include inverse sampling probability weights, controlling for strata fixed effects and clustering standard errors at the trading center level. *** p<0.01, ** p<0.05, * p<0.1.

Table 5.B.4: Number of Attriters by Treatment Status

Reason	Control	Targeted	Spillover	Total
Relocated	92	61	24	177
Died	7	2	1	10
Declined	6	4	0	10
Sick	2	0	0	2
Imprisoned	1	1	1	3
Total	108	68	26	202

5.8.3 Additional Results

Table 5.C.1: Correlates of Take-Up

Variable	(1) Participation among all in treated TCs	(2) Participation among targeted in treated TCs	(3) Participation among untargeted in treated TCs
Female (=1)	-0.000	-0.001	0.067
	(0.037)	(0.039)	(0.051)
Age	0.001	0.001	-0.000
	(0.001)	(0.002)	(0.002)
Married (=1)	-0.002	-0.030	0.012
	(0.032)	(0.030)	(0.045)
Years of Education	-0.008**	-0.007	-0.007
	(0.004)	(0.005)	(0.004)
Previous FL Training (=1)	0.079	0.049	0.107
	(0.066)	(0.099)	(0.094)
Financial Literacy (0-7)	0.018	0.020	-0.004
	(0.013)	(0.016)	(0.012)
Risk Tolerance (0-10)	0.001	0.003	-0.000
	(0.005)	(0.007)	(0.008)
Household Size	-0.001	-0.003	0.001
	(0.009)	(0.009)	(0.010)
# Assets	0.001	0.002	-0.003*
	(0.001)	(0.002)	(0.001)
Household Consumption (UGX)	-0.000	-0.000	0.000
	(0.000)	(0.000)	(0.000)
Constant	0.891***	0.818***	1.082***
	(0.085)	(0.098)	(0.125)
Observations	1203	858	345
R^2	0.082	0.062	0.302
Strata FE	✓	✓	✓

Note: The table shows linear regression results with the binary dependent variable participation in the training. Column (1) includes the all respondents in treated clusters. Column (2) includes targeted respondents only, column (3) only untargeted respondents within the treated clusters. The regression models include strata fixed effects and is weighted by sampling weights and experimental design weights. Standard errors are clustered at the trading center (TC) level and displayed in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 5.C.2: Correlates of Relocation

Dependent Variable:	Relocation (=1)		
	(1)	(2)	(3)
Female (=1)	0.01	0.01	0.01
	(0.012)	(0.011)	(0.012)
Age	-0.00***	-0.00***	-0.00***
	(0.001)	(0.001)	(0.001)
Work Experience (Years)	-0.00	-0.00	-0.00
	(0.001)	(0.001)	(0.001)
Sales (UGX)	-0.00***	-0.00***	-0.00***
	(0.000)	(0.000)	(0.000)
Constant	0.18***	0.18***	0.17***
	(0.024)	(0.031)	(0.035)
Observations	2165	2165	2165
R^2	0.020	0.033	0.033
Strata FE	–	✓	✓
Inverse Sampling Weights	–	–	✓

Note: The table shows linear regression results with the binary dependent variable relocation as the reason for attrition. In column 2 onwards, the regression equation include strata fixed effects and, in column 3, is weighted by sampling weights and experimental design weights. Sales refer to average daily sales in the past month and are top coded at 99 percent. Standard errors are clustered at the trading center (TC) level and displayed in parentheses. *** p<0.01, ** p<0.05, * p<0.1

5.8.4 Robustness Checks

Table 5.D.1: Cluster-level ITT and LATE on Savings, Loans, Investment and Business Formality

	Saving (=1)	ln Saving	Formal Saving (=1)	ln Formal Saving	Loan (=1)	ln Loan	Formal Loan (=1)	Invest (=1)	ln Invest	Record (=1)	Separate Personal (=1)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Control Mean	0.849	10.869	0.176	2.171	0.791	10.078	0.085	0.763	10.149	0.506	0.589
Panel 1: Cluster-level ITT											
Treated vs Control TC	0.005	0.000	0.031*	0.387*	-0.022	-0.388*	0.015	0.028	0.329	0.011	0.008
	(0.018)	(0.252)	(0.018)	(0.228)	(0.019)	(0.229)	(0.013)	(0.021)	(0.282)	(0.025)	(0.023)
Observations	1975	1975	1975	1975	1975	1975	1975	1975	1975	1975	1975
R^2	0.073	0.081	0.099	0.086	0.061	0.070	0.058	0.156	0.162	0.118	0.112
Panel 2: LATE											
Participated in Training	0.028	0.436*	0.056**	0.729**	-0.028	-0.366	0.012	0.040*	0.523*	0.008	-0.001
	(0.020)	(0.264)	(0.023)	(0.302)	(0.021)	(0.249)	(0.015)	(0.023)	(0.303)	(0.030)	(0.027)
Non-Participation	-0.027	-0.630**	-0.005	-0.109	-0.014	-0.419	0.020	0.012	0.048	0.015	0.023
	(0.023)	(0.308)	(0.026)	(0.327)	(0.025)	(0.309)	(0.016)	(0.023)	(0.306)	(0.030)	(0.024)
T = Spillover ($p-value$)	0.007	0.000	0.076	0.058	0.611	0.870	0.637	0.142	0.055	0.836	0.302
Observations	1975	1975	1975	1975	1975	1975	1975	1975	1975	1975	1975
R^2	0.077	0.089	0.103	0.090	0.061	0.071	0.058	0.157	0.163	0.118	0.112

Note: Table shows regression results, controlling for the lagged variable, dummies for missing values, enumerator FE, face-to-face interview dummy, and strata FE. Weighted by sampling weights and experimental design weights. Standard errors are clustered at the trading center (TC) level and displayed in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 5.D.2: Cluster-Level ITT and LATE on Mobile Money Use

	MM Active (=1)	# MM Active (0-4)	MM Saving (=1)	ln MM Saving	MM Transfer (=1)	ln MM Transfer	MM Payment (=1)	MM Supplier (=1)	MM Customer Share
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Control Mean	0.912	1.839	0.184	2.030	0.777	9.348	0.789	0.379	0.037
Panel 1: Cluster-level ITT									
Treated vs Control TC	0.007	0.011	0.033*	0.307	-0.039	-0.424	0.036	0.044*	0.010
	(0.015)	(0.047)	(0.019)	(0.217)	(0.024)	(0.308)	(0.022)	(0.023)	(0.006)
Observations	1975	1975	1975	1975	1975	1975	1975	1975	1975
R^2	0.100	0.133	0.066	0.069	0.074	0.083	0.162	0.143	0.130
Panel 2: LATE									
Participated in Training	0.014	0.044	0.039**	0.349	-0.036	-0.345	0.056**	0.065**	0.015**
	(0.016)	(0.051)	(0.020)	(0.230)	(0.030)	(0.368)	(0.024)	(0.030)	(0.007)
Non-Participation	-0.002	-0.037	0.023	0.247	-0.044	-0.539	0.006	0.013	0.003
	(0.017)	(0.057)	(0.024)	(0.274)	(0.027)	(0.342)	(0.026)	(0.029)	(0.007)
T = Spillover ($p-value$)	0.228	0.133	0.457	0.694	0.804	0.610	0.059	0.180	0.143
Observations	1975	1975	1975	1975	1975	1975	1975	1975	1975
R^2	0.100	0.134	0.066	0.069	0.074	0.083	0.164	0.145	0.132

Note: Table shows regression results, controlling for the lagged variable, dummies for missing values, enumerator FE, face-to-face interview dummy, and strata FE. Weighted by sampling weights and experimental design weights. Standard errors are clustered at the trading center (TC) level and displayed in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 5.D.3: Without Weights: Individual- and Cluster-level ITT and LATE on Savings, Loans, Investment and Business Formality

	Saving (=1)	ln Saving	Formal Saving (=1)	ln Formal Saving	Loan (=1)	ln Loan	Formal Loan (=1)	Invest (=1)	ln Invest	Record (=1)	Separate Personal (=1)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Control Mean	0.849	10.869	0.176	2.171	0.791	10.078	0.085	0.763	10.149	0.506	0.589
Panel 1: ITT and Spillover Effects											
Assigned to Training	0.017	0.208	0.045**	0.534**	-0.017	-0.316	0.023*	0.045**	0.612**	0.007	0.015
	(0.019)	(0.268)	(0.020)	(0.253)	(0.018)	(0.228)	(0.013)	(0.022)	(0.285)	(0.028)	(0.025)
Spillover Group	-0.012	-0.271	0.004	0.046	-0.025	-0.514*	0.009	-0.001	-0.030	-0.000	-0.015
	(0.021)	(0.313)	(0.023)	(0.289)	(0.023)	(0.278)	(0.019)	(0.029)	(0.391)	(0.033)	(0.024)
T = Spillover ($p-value$)	0.123	0.103	0.086	0.109	0.721	0.424	0.460	0.060	0.057	0.815	0.205
Observations	1975	1975	1975	1975	1975	1975	1975	1975	1975	1975	1975
R^2	0.077	0.086	0.100	0.090	0.057	0.069	0.068	0.155	0.164	0.123	0.115
Panel 2: Cluster-Level ITT											
Treated vs Control Villages	0.010	0.079	0.034*	0.403*	-0.019	-0.369*	0.019	0.033	0.439	0.005	0.007
	(0.018)	(0.251)	(0.018)	(0.227)	(0.017)	(0.217)	(0.012)	(0.021)	(0.281)	(0.026)	(0.023)
Observations	1975	1975	1975	1975	1975	1975	1975	1975	1975	1975	1975
R^2	0.076	0.085	0.099	0.089	0.057	0.068	0.068	0.154	0.163	0.123	0.114
Panel 3: LATE											
Participated in Training	0.029	0.476*	0.051**	0.646**	-0.015	-0.235	0.014	0.042*	0.600**	-0.001	-0.001
	(0.019)	(0.262)	(0.022)	(0.279)	(0.019)	(0.238)	(0.013)	(0.023)	(0.298)	(0.030)	(0.027)
Non-Participation	-0.021	-0.540*	0.007	0.023	-0.026	-0.578*	0.027	0.018	0.189	0.014	0.020
	(0.022)	(0.302)	(0.025)	(0.305)	(0.024)	(0.303)	(0.017)	(0.024)	(0.317)	(0.031)	(0.023)
T = Spillover ($p-value$)	0.009	0.000	0.146	0.098	0.670	0.291	0.454	0.231	0.118	0.629	0.381
Observations	1975	1975	1975	1975	1975	1975	1975	1975	1975	1975	1975
R^2	0.078	0.090	0.100	0.091	0.058	0.069	0.068	0.154	0.164	0.123	0.115

Note: Table shows regression results, controlling for the lagged variable, dummies for missing values, enumerator FE, face-to-face interview dummy, and strata FE. Standard errors are clustered at the trading center level and displayed in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table 5.D.4: Without Weights: Individual- and Cluster-level ITT and LATE on Mobile Money Use

	MM Active (=1)	# MM Active (0-4)	MM Saving (=1)	ln MM Saving	MM Transfer (=1)	ln MM Transfer	MM Payment (=1)	MM Supplier (=1)	MM Customer Share
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Control Mean	0.912	1.839	0.184	2.030	0.777	9.348	0.789	0.379	0.037
Panel 1: ITT and Spillover Effects									
Assigned to Training	0.009	0.037	0.050***	0.508**	-0.022	-0.174	0.026	0.035	0.011*
	(0.016)	(0.053)	(0.018)	(0.213)	(0.026)	(0.334)	(0.025)	(0.025)	(0.006)
Spillover Group	-0.011	-0.093	-0.001	-0.009	-0.089***	-0.953**	0.016	0.028	0.007
	(0.021)	(0.062)	(0.027)	(0.303)	(0.032)	(0.401)	(0.025)	(0.034)	(0.007)
T = Spillover ($p-value$)	0.258	0.025	0.057	0.081	0.013	0.017	0.697	0.858	0.586
Observations	1975	1975	1975	1975	1975	1975	1975	1975	1975
R^2	0.112	0.150	0.075	0.078	0.076	0.084	0.185	0.147	0.123
Panel 2: Cluster-Level ITT									
Treated vs Control Villages	0.004	0.002	0.037**	0.368*	-0.040	-0.383	0.023	0.033	0.010*
	(0.016)	(0.049)	(0.017)	(0.201)	(0.025)	(0.323)	(0.022)	(0.022)	(0.006)
Observations	1975	1975	1975	1975	1975	1975	1975	1975	1975
R^2	0.112	0.148	0.074	0.076	0.073	0.081	0.185	0.147	0.123
Panel 3: LATE									
Participated in Training	0.007	0.032	0.036**	0.342	-0.033	-0.285	0.040	0.052**	0.013**
	(0.017)	(0.053)	(0.018)	(0.209)	(0.029)	(0.366)	(0.025)	(0.027)	(0.007)
Non-Participation	-0.001	-0.044	0.037*	0.410	-0.050*	-0.536	-0.003	0.003	0.004
	(0.018)	(0.057)	(0.022)	(0.260)	(0.029)	(0.364)	(0.026)	(0.025)	(0.006)
T = Spillover ($p-value$)	0.605	0.135	0.968	0.775	0.568	0.491	0.083	0.104	0.186
Observations	1975	1975	1975	1975	1975	1975	1975	1975	1975
R^2	0.112	0.148	0.074	0.076	0.073	0.082	0.186	0.149	0.123

Note: Table shows regression results, controlling for the lagged variable, dummies for missing values, enumerator FE, face-to-face interview dummy, and strata FE. Standard errors are clustered at the trading center (TC) level and displayed in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Chapter 6

Shutdown: The Impact of the COVID-19 Restrictions on Micro-entrepreneurs in Rural Uganda*

with:
Jana Hamdan

* This project is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation), project RTG 1723. We are grateful to Gaplink Uganda, in particular Filder Aryemo, for her tireless effort in this project. Christopher Kißling provided valuable research assistance. We further thank the seminar and conference participants at the University of Göttingen, the PEGNet PhD Workshop, and the Annual Meeting of the German Economic Association (VfS) 2021 for valuable discussions and comments. All remaining errors are ours.

Abstract

We provide timely evidence on the impact of the COVID-19 lockdown on the financial well-being of micro-entrepreneurs in a low-income country. The analysis is based on regionally representative panel data on 1,975 micro-entrepreneurs from rural Uganda. We first show that several business characteristics predict compliance to the national lockdown, including running a service business. Further, we document a sharp increase in overall household informal savings, loans, as well as the use of mobile money compared to pre-pandemic levels. Moreover, we find a substantial drop in business investments and profits. Difference-in-difference estimation results suggest a decrease in overall financial well-being more than six months after the national lockdown. In addition, we find that the longer the individual business shutdown due to COVID-19, the smaller the *ex-post* business profits alongside reduced investments and financial safety. These findings point towards a need to design policy instruments aiding small businesses and, in particular, service businesses to help them overcome their ongoing weakened economic situation.

JEL Codes: I18 (Government Policy), I31 (General Welfare, Well-Being), O12 (Microeconomic Analyses of Economic Development).
Keywords: COVID-19, Lockdown, Micro-entreprises, Mobile Money, Coping Strategies, Uganda.

6.1 Introduction

Starting in early 2020, many governments have put in place containment measures to slow the spread of the COVID-19 pandemic. While targeting at flattened infection curves and saving lives, the lockdowns disrupted consumption and global economic operations (Battistini et al., 2021; Chen et al., 2021; Ozili and Arun, 2020). Among the global poor, the situation was particularly challenging. From surveys among 30,000 respondents in April-June 2020, Egger et al. (2021) find significant immediate declines in employment, income, and food security in Bangladesh, Burkina Faso, Colombia, Ghana, Kenya, Rwanda, and Sierra Leone. Should these effects continue, the risk of poverty for vulnerable households is high. However, it remains unclear how the past lockdowns continue to affect on the people and businesses in low income countries. In Uganda, micro-entrepreneurs account for the majority of businesses and play a crucial role in the economy. Understanding the impact of lockdowns on this group beyond the immediate effects is important to design better policies and development assistance.

We provide evidence on the medium-term impact of COVID-19 measures on rural micro-entrepreneurs in Western Uganda, utilizing survey data from both one year before and several months after the immediate lockdown period. The sample is based on an extensive in-person baseline survey of 2,177 respondents in spring 2019. The endline survey reaches 91 percent of the baseline sample using both phone and in-person interviews between October 2020 and April 2021, thus at least 6 months after the stay-at-home national lockdown.[1]

The survey data provides us with unique information on the individual duration of business closure due to COVID-19. Moreover, we collect detailed information on business performance, financial behaviors, and mobile money use, both before and after the imposed shutdown. This allows us, first, to investigate the variation in business closure duration and illustrate several correlating characteristics to understand what drives compliance to the national lockdown regulations. In a second step, we exploit the variation of the length of shutdown across businesses and estimate the effect of shutdown length on business performance, household finance, and mobile money use. We apply a difference-in-difference estimation by exploiting the variation in business shutdown length across businesses and across time.

The analysis reveals several findings. First, the variation in reported business closure is substantial. Service businesses are particularly affected by the decree. On av-

[1] We utilize the first two rounds of a panel data collected for a randomized controlled trial, evaluating the impact of a financial education program implemented in 2019 (Hamdan et al., 2021). The program has no effect on business shutdown due to COVID-19.

erage, the less educated are more likely to close their business longer relative to those with higher education; however, the illiterate are least likely to do so. In addition, those with less access to mobile money, as well as those more independent from mobility restrictions, shut down more shortly as well. At the village level, we find that a higher share of local service industry and a longer distance to the main town are linked to a longer shutdown.

Second, we document that financial stress increases after the COVID-19 lockdown but, at the same time, the adoption of financial innovation increases. We document that micro-entrepreneurs carry a higher debt burden and are more likely to be late with loan payments. Meanwhile, although they generally increase savings, most of this rise was from informal savings. Business performance, measured from profits and investments, is on average worse than pre-pandemic levels. In addition, we find that the sampled micro-entrepreneurs considerably increase their use of mobile money after the national lockdown, both at the extensive and the intensive margins. This could be driven by the temporary fee waivers in spring 2020, changes in awareness and preferences, and restrictions to physical mobility. The rise in mobile money use also speaks for risk sharing within social networks during an economic crisis.

Third, we find from the difference-in-difference estimation results that longer business shutdown leads to a lower probability of business survival as well as lower profits and smaller investments in the medium term. The incidence of business shutdown causes a higher debt burden and lower financial safety. On the upside, we show positive effects of business shutdown on active mobile money use, especially for business purposes.

Taken together, these findings suggest that the financial downturn experienced by micro-entrepreneurs in Uganda lasts beyond the imposed lockdown period. Moreover, compliance with lockdown rules is quite heterogeneous. It is important for policy makers to consider the determinants of this heterogeneity. For example, targeting information campaigns to the illiterate population could be fruitful for efforts to increase compliance with regulations, while service businesses could be especially aided via financial support programs.

This paper adds directly to studies investigating the impact of COVID-19 lockdowns on livelihoods in low income economies. In Uganda, negative immediate impacts on business profits, household income and consumption are documented both in the capital (Hartwig and Lakemann, 2020) and rural areas (Mahmud and Riley, 2021). Moreover, similar findings are reached via an online survey study in Uganda and Kenya (Kansiime et al., 2020). COVID-19 lockdowns are found to cause huge

immediate financial stress for households in Bangladesh (Rahman and Matin, 2020), India (Gupta et al., 2021), Pakistan (Malik et al., 2020), and Senegal (Le Nestour et al., 2020). Most evidence highlights that poorer households are more vulnerable to the negative shock.

Compared to the existing evidence, this paper provides insights in four ways. First, this study differs from the previous studies by providing medium-term impact analysis about one year after the national lockdown started. This is relevant as the negative impact on livelihoods may extend beyond the immediate shutdown period. Long lasting effects could not only be different from the immediate impact but also heterogeneous, with households and businesses recovering from the shock at different speeds and their individual experiences shaping their long-run behaviors.

Second, we discuss the compliance with COVID-19 restrictions, a topic predominantly studied in high income countries. This relates to existing evidence from 92 interviews across six communities in Bangladesh on the compliance with the national six-week lockdown (Ali et al., 2021). They argue that a lack in state capacaity and public critisism of the relief regime caused the low compliance in the South Asian country. Our analysis relies on a large regionally representative sample, and provides novel evidence from rural Uganda.

Third, our study is built on an extensive panel of 1,975 regionally representative respondents, enabling the, to date, largest study on the economic impact of COVID-19 in East Africa. In addition, the attrition between the baseline and follow up surveys is small – less than ten percent. Both increase the internal validity of the findings. Moreover, in contrast to previous studies, we focus on micro-entrepreneurs and their finances, as this group is large and relevant in low income economies.

Fourth, this study also adds to the emerging literature on individual behaviors after an extreme and negative shock in a rural setting. For example, Blumenstock et al. (2016) point out that mobile money use and transfers increased after an earthquake in Rwanda. Mobile money is shown to have profound impacts regarding informal risk sharing (Jack et al., 2013; Jack and Suri, 2014). Similarly, we provide consistent evidence that the use of mobile money increases in the aftermath of an unexpected shock.

The structure of this paper is as follows. In Section 6.2, we outline the course and implications of the enforced COVID-19 measures in Uganda. Section 6.3 describes our data and empirical strategy. Section 6.4 provides suggestive evidence that the observed variation in business closure length is correlated with several business and respondent characteristics. In Section 6.5, we present our main results on the impact of business

closure on business performance, financial well-being, and financial behaviors. Section 6.6 provides policy suggestions and concludes.

6.2 Background on the Ugandan Lockdown

The Ugandan government reacted to the COVID-19 pandemic with strict restrictions to social interactions and mobility. Already on March 18, 2020, public gatherings were suspended, including church services, weddings, cultural meetings, and rallies. Moreover, pubs were ordered to close. On March 22nd, measures were tightened with a strict stay-at-home lockdown. This lasted for six weeks until May 5th. After this date, a few businesses were allowed to operate, including factories, garages, hardware and metal-working businesses, and restaurants for take away. However, most businesses and schools still had to remain closed. These restrictions were prolonged by another 12 weeks through July 27th, after which most of the businesses were gradually allowed to re-open. However, even in November 2021, many restrictions still remain: most schools are still closed; some businesses such as bars and night clubs have not been permitted to operate; buses and taxis have to operate at half capacity; and there is a night curfew. In a nutshell, the national lockdown imposed a shutdown for most micro-enterprises that lasted between six and 18 weeks according to official regulations; however, even 20 months after lockdowns were initiated, some still cannot operate as usual at the time of the follow up survey.

To compare the Ugandan measures to those of other Sub-Saharan countries, the Oxford COVID-19 Government Response Tracker's Stringency Index (Hale et al., 2021) is a useful measure. The index ranks the strictness of government pandemic responses on a scale of zero to 100. The composite measure includes nine metrics: school closures, workplace closures, cancellation of public events, restrictions on public gatherings, closures of public transport, stay-at-home requirements, public information campaigns, restrictions on internal movements, and international travel controls. Figure 6.A.1 in the appendix shows that Uganda exhibits a very high index value until September 2020, dropping down thereafter until increasing again in July 2021, compared to 43 other sub-Saharan countries. It is noteworthy that African lockdowns are generally stringent compared to other continents while economic support to citizens is simultaneously lower (Lakemann et al., 2020).

Due to limited testing, it is difficult to ascertain the extent of COVID-19 diffusion and health effects within the Ugandan population. By October 4, 2021, there have

been a total of 123,976 confirmed cases and 3,160 deaths.[2] This is in a context of a population of 46 million in 2020. Roughly three out of four Ugandans are younger than 25, making it one of the youngest countries on the planet. Less than two percent were older than 65 in 2019 and 74 percent live in rural areas.[3] With regard to the economy, the World Bank (2020) reports severe effects of the COVID-19 pandemic in Uganda. These come from the long lockdown, border closures, and further side effects of disruption in global demand and supply chains. Real GDP growth is expected to contract by up to one percent in 2020, compared to 7.5 percent growth in 2019.[4]

The imposed strict lockdown in Uganda did not just affect business directly through their opening ban, but also indirectly via social distancing and reduced mobility. Thus, the length of each business's shutdown and the economic consequences may not only be attributed to the lockdown directly, but also indirectly. For example, businesses might have fewer customers during the lockdown and afterwards due to reduced demand. They might also be short of goods such as products for resale and tools during and after the lockdown. In fact, Mahmud and Riley (2021) find that 58 percent of their Ugandan sample (1,075 rural households in Kagadi and Kyenjojo districts) in May 2020 report that the main reason for business closure was regulation. However, 29 percent say that they mainly closed not because of the imposed lockdown, but because no customers came or no stock was available for them to sell. See Figure 6.A.4 in the Appendix for all reported main reasons. This information makes a high variation in the realized length of business closures likely − both within and beyond the actual lockdown period.

In our follow up survey, we collected qualitative evidence on what caused the variation in business shutdown length. First, there is variation across different villages. The respondents indicated that the strictness of the enforcement to the national lockdown orders by local leaders and the resident district commissioners may vary. Second, within villages, different micro-entrepreneurs had different levels of compliance to the

[2] These numbers are from Worldometer, retrieved October 04, 2021, 09:20 GMT.

[3] See Worldometer, retrieved October 4, 2021, 09:20 GMT.

[4] The World Bank (2020) further reports that "the COVID-19 crisis is threatening to reverse some of the gains made on structural transformation and the declining poverty trend of the past decade. This transformation was characterized by a reduction in the workforce employed in on-farm agriculture and a take-off in industrial production, largely in agro-processing. However, following the COVID-19 shock, there have already been widespread firm closures, permanent layoffs in industry and services, a rapid slowdown of activity particularly in the urban informal sector, and a movement of labor back to farming. At the same time, household incomes have fallen which is concerning given the high levels of vulnerability to poverty, limited social safety nets, and impacts this might have on human capital development and Uganda's capacity to benefit from its demographic transition."

national lockdown regulations. For example, some respondents closed down the businesses due to the lockdown rules directly or indirectly. There is a higher cost of running a business during the lockdown period. Some reported suffering from high transport fares during and after the lockdown, meaning they could not afford to go to towns to buy goods. Local corruption was reported to occur more frequently, increasing the cost of business operation. Some respondents reported goods being confiscated when a business operated illegally and then had to pay a bribe in order to released the goods. During the lockdown, some respondents reported that they focused fully on agricultural cultivation of their land. Further, some business premises were broken into during the lockdown and trade items were stolen, therefore causing even more losses to the owners. There were also some reported non-compliance with the order of the strict lockdown. For instance, some businesses continued to operate illegally from the backdoor. Some respondents stated that they would open briefly, keeping their eyes open for authorities. Others would get orders from clients, then open briefly or deliver the goods to customers without having to open the business.

6.3 Method

6.3.1 Data

The study region for our panel data spans 108 rural trading centers in the Kabarole district of Western Uganda. Trading centers are shopping streets along village roads where almost all shops are located and where the main business activity occurs. Their distribution is shown in Figure 6.1. These locations were mapped in cooperation with local officials and the staff at the Mountains of the Moon University in November 2018. The average distance to the district's main town, Fort Portal, is about 28 minutes by car or 19 kilometers (ranging between three and 50 kilometers). In addition, trading centers vary in size, ranging from about five to 392 settled micro-entrepreneurs, with an average of 51 (median 35). Trading centers with 100 or fewer small businesses tend to be less busy, with businesses typically closed for parts of the day. In these small trading centers, we target all open shops at the time of the visit to interview sufficient business owners in these locations. In the nine larger trading centers, we randomly approach every third open business.[5] The share of micro-entrepreneurs declining to participate is negligible. Overall, we sample a large, regionally representative, fraction

[5] This sampling strategy produces a mean interviewee fraction of 47 percent in the smaller trading centers and 35 percent in the larger centers.

of micro-enterprises in our study region. We interview about 41 percent of the estimated total, thus 2,177 rural micro-entrepreneurs, representing an average of 21 per trading center.

Our baseline survey is implemented via in-person interviews from February to April 2019 with trained local university students as interviewers, supervised by local researchers and the authors. We follow up with our sample via phone surveys from October to December 2020 and face-to-face interviews in March and April 2021, ultimately reaching about 91 percent of the baseline sample.[6] Thus, after more than 1.5 years, we successfully re-interview 1,975 of the micro-entrepreneurs.[7]

We document the reasons for the moderate attrition during the follow up survey. 88 percent of the attrition is due to a relocation of a micro-entrepreneur out of our study region, with only 25 cases of attrition due to death, refusal, imprisonment, or sickness. The relocated are characterized by worse business performance at baseline. Therefore, we suspect that they are more likely to close down the original business permanently and move away as their businesses might have a lower chance of surviving the COVID-19 crisis. Our study results on the negative lockdown impact on business performance may be downward biased since all effects are conditional on the fact that the micro-enterprises survived the COVID-19 crisis.

Both baseline and follow up surveys cover the financial situation of the businesses and the owners in detail, particularly their profits, savings, investing, borrowing, and transfer behaviors, as well as the use of mobile money. In Appendix Tables 6.A.3-6.A.5 in Section 6.8, we provide variable definitions. Importantly, the follow up survey includes a question on the length of business closure due to COVID-19.

Summary statistics of our panel sample at baseline are displayed in Table 6.1. Almost two-thirds of all business owners are female. They average 34 years of age and have seven years of work experience in a business. About 62 percent run a retail or wholesale business, while 28 percent manage a service business, with the remaining 10 percent working in the manufacturing sector. While 81 percent are formalized with a trading license, only 22 percent keep business records. While almost all respondents have access to mobile money services, only 20 percent own a bank account. On average they make a monthly profit of US$ PPP 133, and have outstanding loans

[6] The follow-up is implemented in cooperation with Gaplink Uganda, an independent research company.

[7] The phone survey conducted in winter 2020 successfully interview 73 percent of the baseline sample, making up to eight attempts to reach respondents via the contact information collected at baseline. We then follow up with the micro-entrepreneurs who are not reached on the phone via face-to-face interviews in spring 2021.

almost twice that amount (34 percent had a loan at baseline). Before COVID-19, the micro-entrepreneurs had, on average, 2.8 times their monthly consumption as savings. However, it is noteworthy that the median total savings are only 162 US$ PPP, meaning actually less than one month's consumption in savings. Every third business is not located along a main road in the district and, as such, can be considered as being remote.

6.3.2 Empirical Strategy

As a first step, we estimate the correlates of the length of business closure because of COVID-19. As described in Section 6.2, the business shutdown length may vary across the trading centers and as well as within trading centers across micro-entrepreneurs. Both are of interest to policy makers to understand on who and where it is important to focus. We estimate what characteristics at the business and trading center levels predicts the shutdown length. First, at the individual business level, we use the following model for the analysis:

$$Shutdown_i = \alpha + \beta' X_{i0} + \epsilon_i \tag{6.1}$$

where $Shutdown_i$ is the self-reported length of business shutdown by the micro-entrepreneur i. On the right hand side, α is a constant and X_{i0} includes a set of observed characteristics of micro-entrepreneur i at baseline. The standard errors, ϵ_i, are clustered at the trading center level. With a large number of possible explanatory variables X_{i0}, a model could be easily over-fitted. Therefore, we use the machine learning technique, LASSO (Tibshirani, 1996), to select the variables which best explain the variation in the shutdown length. The procedure reduces the complexity of the model and solves the covariate selection problem. The chosen characteristics are presented in Section 6.4.

Second, we study the correlates of trading center characteristics and the average length of business shutdown within them, as follows:

$$Shutdown_v = \alpha + \beta_1' Z_{v0} + \beta_2' \hat{X_{i0}} + \epsilon_v \tag{6.2}$$

where $Shutdown_v$ captures the mean shutdown duration within trading center v and Z_{i0} includes two location characteristics. These are trading center size and the distance to the district main town. In addition, the model includes $\hat{X_{i0}}$, the local average

of selected respondent characteristics. ϵ_v stands for the standard errors. Similarly, we use LASSO to select the trading center level characteristics that explain the between trading center variation in the average business closure length.

Finally, we study the effects of COVID-19 related business closure on the micro-entrepreneurs. We first perform descriptive analysis by estimating the mean of each outcome of interest by the two survey waves (before and after the COVID-19 shock). Then, we use a difference-in-difference estimation to tackle the causal effects of the shutdown on micro-entrepreneurs. Specifically, the model takes the following form:

$$Y_{it} = \alpha + \beta Shutdown_{it} + \theta_i + \gamma_t + \epsilon_{it} \tag{6.3}$$

where Y_{it} is the outcome of interest of the business or micro-entrepreneur i surveyed at the baseline ($t = 0$) or follow-up ($t = 1$) survey. α is a constant. $Shutdown_{it}$ is the self-reported length of business closure due to COVID-19 restrictions, measured in number of weeks or a binary variable indicating the incidence of the business shutdown. The business shutdown length or incidence takes the value of zero for all the observations in the baseline survey. θ_i refers to the business fixed effects, controlling for all the business-specific characteristics that may affect the outcome directly. Moreover, γ_t denotes the year-month fixed effects, which further adjust for changes in the outcomes due to seasonality or any macro shock affecting all micro-entrepreneurs in the same month. The year-month fixed effects also take into account potential differences due to interview method, since all face-to-face follow up interviews took place in March and April 2021. The estimation of β relies on the comparison of the same business before and after the shutdown, as well as the variation of shutdown length or incidence across businesses. All standard errors are clustered at the baseline trading center level.

6.4 The Correlates of the Business Shutdown

In our follow up survey, respondents were asked "Did you have to shut your business due to the corona virus (COVID-19)? If yes, for how many weeks?". We show the distribution of answers (in percentages) in Figure 6.A.2 in the appendix. Generally, it shows a shorter shutdown than the expected six to 18-weeks imposed by the Ugandan government. Around every fourth respondent did not close their business at all. In turn, 72 percent report that they closed their business due to COVID-19, and 50 percent shut for six weeks or longer. Conditional on any closure, they shut down on

average for 11 weeks (median 12), with closures ranging between one and 52 weeks. Thus, only half of the respondents closed their business in accordance with the lower bound of the lockdown regulations, while others did so for much longer than required. To better understand why business closure is so heterogeneous, we discuss some factors that could be at play and provide suggestive evidence.

First, there is likely local variation in information dissemination and the level of enforcement of lockdown measures by local officials. This could affect compliance throughout our sample and also cause differential knowledge about COVID-19 and imposed regulations. Second, micro-entrepreneurs possibly closed their business longer than being obliged to due to individual health and safety concerns, or any other personal issues. For example, this could be the case due to low profitability, low sales and inventory, and higher transport fares, making outside options such as cultivating land more attractive. Actually, business income was more than 20 percentage points more likely to fall than farm income in Uganda between May and August 2020 (Josephson et al., 2021). While we cannot disentangle the exact reasons for individual business closure related to the pandemic, we are able to assess the general correlates of business shutdown. This is important for understanding what business types and which micro-entrepreneurs comply more to the policy and are most affected by consequences of longer business closure, as discussed in the next section.

We use LASSO to select the determinants for business shutdown among 128 possible explanatory variables from the baseline survey data. For the binary dependent variable of any business closure, three variables are selected by the approach as relevant determinants. For the continuous variable, length of business closure measured in number of weeks, 15 variables are chosen. After selecting these substantial variables, we conduct the ordinary least squares regression (OLS post-LASSO) in a multivariate analysis. We plot the resulting coefficients in Figures 6.2a and 6.2b, respectively. This presentation accentuates the relative impact of the coefficients on the shutdown probability. Most significant are the business type characteristic, an indicator of extreme remoteness, mobile money use, and illiteracy. According to the LASSO approach, variables such as debt, financial literacy, and risk preferences are not important determinants of business closure due to COVID-19. The related full regression results are shown in Appendix Tables 6.A.1 and 6.A.2.

It is striking that businesses in one sector were more likely to be closed and shut down longer: owning a service business is associated with both higher likelihood and longer business closure. On average, they closed almost three weeks longer than retail or manufacturing businesses. Figure 6.A.3 visualizes these differences by business

type. As we control for factors such as profits, investments, and work experience,[8] we speculate that the driving factor of this distinction is the higher contact intensity and resulting perceived COVID-19 risks associated with this sector compared to other types of micro-enterprises. Moreover, they could be perceived as less of a necessity by their customers, driving down demand. These possible explanations most apply to two specific service businesses: hair and beauty salons, and restaurants. Compared to all others, they even closed an additional five and two weeks, respectively. In turn, older and more experienced business owners were less likely to close their business, possibly due to higher opportunity costs.

Finally, three other factors stand out: living in a very remote location where one needs to take a motorcycle taxi (boda) to the nearest mobile money agent (this is the case for only five percent of the panel), is on average associated with closing the business for about two weeks less. A similar effect is observed with the financial behavior of using mobile money to pay suppliers (just 3.5 percent of the panel did so at baseline). This could be due to a higher independence from mobility restrictions. Lastly, being illiterate also correlates negatively with the number of weeks a business closed. This attribute applies to 12 percent of the panel. It could be that COVID-19 awareness campaigns and the lockdown declaration did not reach them due to their limitation and also failed to reach those in very remote locations.

In sum, we find suggestive evidence that service businesses are more likely to have longer business closure. These should be specifically targeted by aid programs to compensate for their substantial losses. At the same time, those very remote, those independent from mobility restrictions to reach suppliers, and the illiterate appear to have complied less with (or known less about) the lockdown regulations and could be less affected by the national shutdown.

Next, we investigate whether two trading center characteristics and the micro-entrepreneur composition correlates with local average business shutdown. This would be relevant to design any regional policy focus. In Table 6.2, we show the results of a linear regression across the 108 trading centers. Similar to before, we use LASSO to select the relevant determinants for business shutdown on the trading center level. The approach selects seven and three variables for the incidence and extent of business closure, respectively. We find that the local probability of business closure increases with distance to the district main town: ten additional kilometers to the main town

[8] On average, service businesses made 19 percent less profits, were three percentage points more likely to have made investments in the past year before baseline, and their owners have about 1.5 years less work experience.

are, on average, associated with a three percentage point higher likelihood of local business closure. In addition, trading centers not located along a main road have a five percentage point higher shutdown incidence. We argue that this suggests a larger impact of opportunity costs rather compared to lockdown enforcement, since those trading centers closer to the city are larger and along main roads. More rural micro-entrepreneurs may retreat more easily to farming (at baseline, there is a strong positive correlation between distance to the main town and non-business income sources). However, lack of easy access to mobile money relates to much shorter shutdown. This could be due to two reasons. First, low access to mobile money could be an indicator of extreme remoteness, causing a lack of information about the regulations. Second, having low access to mobile money services also affects the risk sharing ability of the micro-entrepreneurs in this location.

In addition, we find that a larger share of service businesses correlate with longer business shutdown at the trading center level. However, the trading center size does not appear to play a role. Overall, these findings indicate that the more rural locations and those with a larger service industry are more affected by business shutdown due to COVID-19.

6.5 The Impact of Business Shutdown on Micro-entrepreneurs

6.5.1 Descriptive Findings

Descriptive figures, illustrating the considerable changes over time across financial outcomes are shown in Figure 6.3. First, regarding the business performance, we document that micro-entrepreneurs have smaller business profits and invest less compared to the pre-COVID period. Moreover, prior to the COVID-19 crisis, only about 50 percent of micro-entrepreneurs use mobile money actively. However, in late 2020 and early 2021, almost 90 percent of them are qualified as active mobile money users.Moreover, they are also more likely to use mobile money to transfer money to business partners than before.

In addition, we observe that the micro-entrepreneurs are ten percent more likely to save and they increase their savings. However, the increase is mainly driven by the rise in informal savings. This implicates lower financial safety among the micro-entrepreneurs.

Despite the generally increasing savings, a sharp increase in individual debt is observed. After the COVID-19 crisis, almost 80 percent of the respondents have an outstanding loan. However, prior to the crisis, less than 40 percent of them have an existing loan. The average loan amount in Ugandan Shilling has increased by over 2.4 times. In addition, the probability of the micro-entrepreneurs being punctual with loan repayment decreases from 92 percent to only 49 percent.

6.5.2 Main Results

The causal effects of the extent of reported business closure based on the difference-in-difference estimation on business performance, household finance, the use of mobile money, and money transfer behaviors are set out in Tables 6.3, 6.5, 6.4, and 6.6, respectively.

Regarding business performance (Table 6.3), a longer business shutdown would lead to a lower probability of business survival, smaller business profits, and reduced investments. Specifically, one additional week of business shutdown would lead to 5.0 and 9.7 percent decrease in profits and investments, respectively.

When it comes to the use of mobile money (Table 6.4), we observe an increased use both at the extensive and intensive margins. Specifically, the coefficient in column (1) suggests that one additional week of business shutdown would cause a 0.6 percentage point increase in the active use of mobile money and experiencing any business closure would lead to a 8.3 percentage point increase in the general use of mobile money. Moreover, the micro-entrepreneurs are more likely to use mobile money for making payments and for business purposes, like paying suppliers, bills, and receiving customer payments.

For loans and transfers, we show results in Table 6.5. Credit burdens significantly increase with the incidence of business closure but is not affected by the length of shutdown. While we report overall lower repayment punctuality across the sample, this is not affected by the individual length of business shutdown, but appears to be a trend. For transfers, we observe that the probability of losing a transfer increases with the length of business shutdown.

Lastly, Table 6.6 reports the effects on savings. We do not find any evidence showing that individual savings decreases in the medium-term as a consequence of COVID-19 business shutdown. However, the safety of savings drops and this could be due to the increase in informal savings. Overall, the regression analysis on the causal

impacts of the COVID-19 shutdown yields consistent results with the descriptive analysis.

6.6 Conclusion

So far, the attention of COVID-19 research in low income countries has been focused on documenting increased poverty and food insecurity. In addition to these documented negative immediate effects of lockdowns, we provide unique evidence on their persistent impact on financial well-being and behaviors from a large panel of micro-entrepreneurs in rural Uganda.

In this paper, we present two major contributions: one on the heterogeneity in compliance to lockdown regulations and the other on the impact of the business shutdown on business outcomes, mobile money use, savings, loans, and financial security.

Firstly, we find that only about half of the sampled micro-entrepreneurs comply with the minimum national lockdown regulations. Several business owner and location characteristics correlate with the reaction to COVID-19. Most important for policy makers should be that service businesses (especially those with high contact intensity) and the more rural were more likely to shut down, while those with less access to mobile money, the illiterate, and micro-entrepreneurs more independent from mobility would comply less. Our findings suggest important roles for opportunity costs and information dissemination.

Secondly, we show that while business closure appears to be a driver of innovation and business practices (boosting the use of mobile money and record-keeping practices), it worsens business performance in terms of profits, investments, and financial stress, as it causes higher debt levels for the micro-entrepreneurs.

The findings in this paper have several important policy implications. In these exceptional circumstances, micro-entrepreneurs have a higher need of financial services, especially mobile money services and loans. This suggests that governments in low income countries need to ensure that the corresponding costs are provided affordably even when demand increase. In Uganda, mobile money transfer fees were suspended temporarily and voluntarily by the main providers. This likely supported risk sharing among social networks and could be partially responsible for the higher usage rates today. By further promoting the adoption of mobile financial services and ensuring consumer protection, specifically with regard to loans and repayment flexibility, governments in low income setting may improve general welfare.

As certain micro-entrepreneurs are more likely to comply with the lockdown rules, these should be targeted by aid programs by providing, for instance, subsidies for service businesses. Low compliance among half the sample and specifically for certain groups underlines that policy makers need to give out special support and target information campaigns to these groups during a lockdown period to achieve higher levels of compliance.

6.7 Chapter 6 - Figures and Tables

Figure 6.1: The Study Region and Sampling Trading Centers in Western Uganda

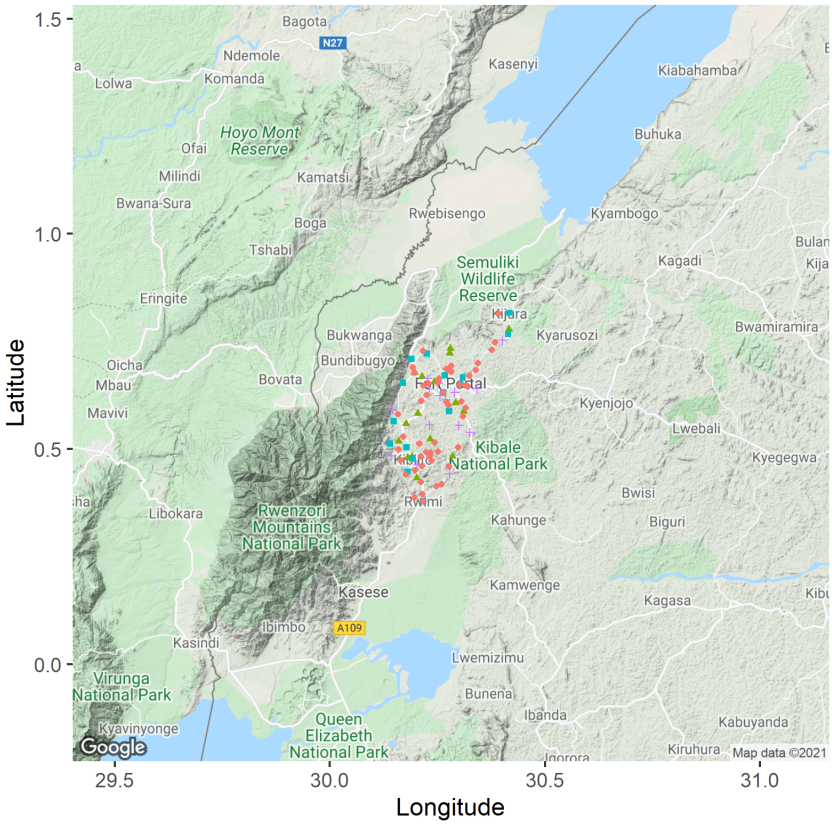
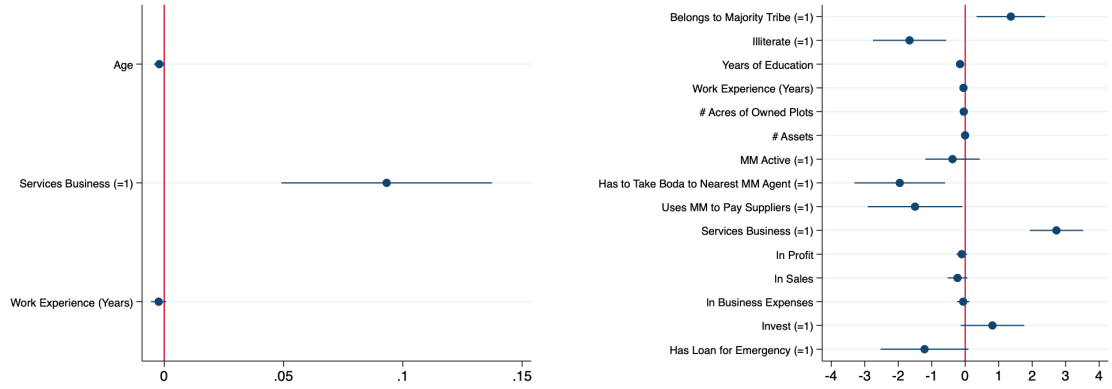


Figure 6.2: Post-LASSO Correlates of the Business Shutdown



(a) Incidence of Shutdown (b) Length of Shutdown

Note: The figure displays the point estimates and their confidence intervals of post-lasso (OLS) regression results. These are shown in Appendix Table 6.A.1 and 6.A.2.

Figure 6.3: Financial Outcomes Before and After the 2020 COVID-19 Lockdown

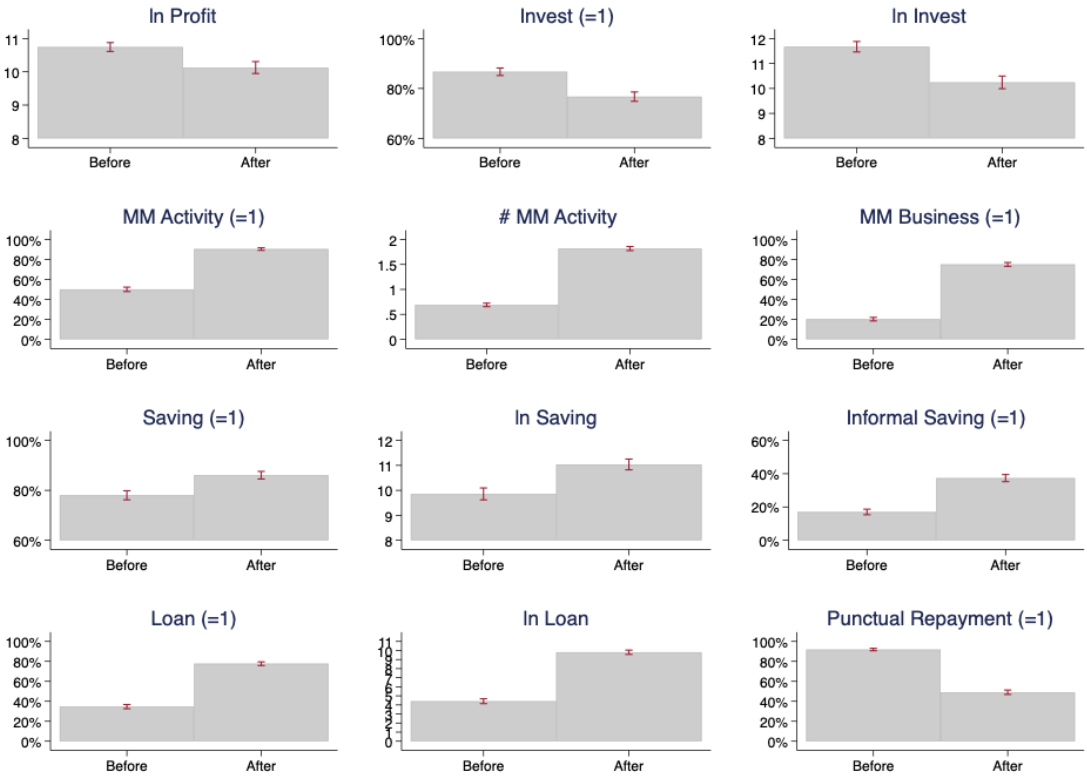


Table 6.1: Summary Statistics

Variables	Mean	Std. Dev.	N
Individual and household characteristics			
Female (=1)	0.633	–	1975
Age	34.238	11.716	1967
Married (=1)	0.504	–	1975
Illiterate (=1)	0.117	–	1975
Years of Education	8.830	4.596	1975
Previous Financial Education (=1)	0.374	–	1975
Financial Literacy (Std.)	0.013	0.767	1975
Risk Tolerance (Std.)	0.019	0.994	1975
Work Experience (Years)	7.147	7.809	1975
Household Size	4.039	2.470	1975
# Assets	37.565	17.697	1975
Consumption per adult equivalent (US$ PPP)	199.302	123.732	1975
Business characteristics			
Retail/Wholesale Business (=1)	0.625	–	1975
Services Business (=1)	0.277	–	1975
Manufacturing Business (=1)	0.098	–	1975
Trading License (=1)	0.813	–	1975
Record (=1)	0.219	–	1975
Separate HH Record (=1)	0.233	–	1975
# Workers	0.430	0.715	1975
Financial characteristics			
MM Account (=1)	0.867	–	1975
Bank Account (=1)	0.219	–	1975
Monthly Profit (in US$ PPP)	133.420	178.860	1938
Investments in Past Year (in US$ PPP)	1182.489	2288.314	1947
Current Loans (in US$ PPP)	226.675	632.492	1975
Total Savings (in US$ PPP)	564.654	1229.620	1975
Location characteristics			
Remote Location (=1)	0.311	–	1975
MM Agent Density	3.941	3.600	1975
Network Quality (0-10 Scale)	7.246	2.151	1975

Note: The table shows summary statistics for 1975 small business owners in the rural Kabarole district in Western Uganda at baseline (27th February-11th April 2019). The variables consumption, profits, investments and loans are winsorized at 99% and converted from Ugandan Shillings (UGX) using the 2019 PPP conversion factor for private consumption for Uganda: 1 US$ = 1235.95 UGX (Source: The World Bank). To calculate the adult equivalent of household-level consumption, we use the OCED-modified scale, first proposed by Hagenaars et al. (1994) which assigns the weight 1 to the household's first adult, 0.5 to each additional one, and 0.3 to each child. Remote location indicates businesses being not along a main road.

Table 6.2: Village-Level Correlates of Business Shutdown

Dep. Variable: Trading Center Avg. Business Closure	Any (=1)	# Weeks
	(1)	(2)
Distance to District Main Town (km)	0.003***	
	(0.001)	
Remote Location (=1)	0.047*	
	(0.027)	
Services Business (=1) *avg.*		5.755***
		(1.936)
# Children in Household *avg.*	-0.092***	-0.827**
	(0.017)	(0.349)
Has to Take Boda to Nearest MM Agent (=1) *avg.*	-0.245***	-3.639***
	(0.061)	(1.189)
ln Profits *avg.*	0.026*	
	(0.015)	
Able to Manage Money (1-4 Scale) *avg.*	0.076	
	(0.050)	
Missing: Father Education (=1) *avg.*	0.133	
	(0.107)	
Constant	0.630***	11.795***
	(0.208)	(1.670)
Observations	108	108
R^2	0.494	0.237

Note: The table shows post-LASSO (OLS) regression results with the binary dependent variable being mean of any reported business closure due to COVID-19 on the trading center (TC) level in column 1, and the continuous dependent variable being the mean number of weeks of such business closure (not dependent on any closure) in column 2. The sample consists of the 108 trading centers where the respondents were based at baseline (95% of our sample did not move after the baseline survey). The ability to manage money well over time is elicited from the question "If you get money, do you tend to spend it too quickly? Answer options: Ofen, Sometimes, Rarely, Never." *** p<0.01, ** p<0.05, * p<0.1.

Table 6.3: The Effects of the Shutdown on Business Outcomes

	Survival (=1)	ln Profit	Invest (=1)	ln Investment	Record (=1)	# Workers (=1)
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Length of Business Closure						
Shutdown (Weeks)	-0.003***	-0.050***	-0.008***	-0.097***	0.003*	0.000
	(0.001)	(0.016)	(0.002)	(0.023)	(0.002)	(0.003)
Observations	3950	3950	3950	3950	3950	3950
R^2	0.956	0.562	0.540	0.552	0.619	0.604
Panel B: Incidence of Business Closure						
Shut Down (=1)	0.007	0.030	-0.024	-0.321	0.110***	0.007
	(0.014)	(0.293)	(0.032)	(0.448)	(0.031)	(0.085)
Observations	3950	3950	3950	3950	3950	3950
R^2	0.956	0.559	0.534	0.547	0.621	0.604

Note: The table shows regression results for six variables linked to business performance and behaviors. Controls for business fixed effects, and interview month-year dummies. Weighted by sampling weights and follow-up attrition probability. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 6.4: The Effects of the Shutdown on Mobile Money Use

	MM Active (=1)	# MM Active	MM Transfer (=1)	ln MM Transfer	MM Pay (=1)	MM Business (=1)
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Length of Business Closure						
Shutdown (Weeks)	0.006***	0.007*	0.003	0.032	0.006**	0.005**
	(0.002)	(0.004)	(0.002)	(0.024)	(0.002)	(0.002)
Obserations	3950	3950	3950	3950	3950	3950
R^2	0.659	0.728	0.650	0.665	0.668	0.699
Panel B: Incidence of Business Closure						
Shutdown (=1)	0.083**	0.146**	0.066*	0.603	0.067**	0.076**
	(0.034)	(0.067)	(0.036)	(0.420)	(0.031)	(0.029)
Obserations	3950	3950	3950	3950	3950	3950
R^2	0.658	0.729	0.650	0.665	0.667	0.699

Note: The table shows regression results for six variables linked to mobile money services. Controls for business fixed effects, and interview month-year dummies. Weighted by sampling weights and follow-up attrition probability. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 6.5: The Effects of the Shutdown on Loans and Transfers

	Loan (=1)	# ln Loan	Never Late (=1)	Any Transfer (=1)	ln Transfer	Any Transfer Stolen (=1)
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Length of Business Closure						
Shutdown (Weeks)	0.003	0.029	-0.001	0.002	0.023	0.004**
	(0.002)	(0.031)	(0.002)	(0.002)	(0.022)	(0.001)
Obserations	3950	3950	3950	3950	3950	3950
R^2	0.646	0.653	0.638	0.641	0.679	0.523
Panel B: Incidence of Business Closure						
Shutdown (=1)	0.080**	0.928**	-0.042	0.047	0.496	0.031
	(0.036)	(0.436)	(0.031)	(0.030)	(0.378)	(0.025)
Obserations	3950	3950	3950	3950	3950	3950
R^2	0.647	0.654	0.638	0.642	0.679	0.521

Note: The table shows regression results for six variables linked to loans and transfers. Controls for business fixed effects, and interview month-year dummies. Weighted by sampling weights and follow-up attrition probability. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 6.6: The Effects of the Shutdown on Savings

	Saving (=1)	ln Saving	Formal (=1)	ln Formal	Informal (=1)	ln Informal	Savings Stolen (=1)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Length of Business Closure							
Shutdown (Weeks)	-0.002	-0.027	-0.000	-0.009	-0.001	-0.012	0.002
	(0.002)	(0.027)	(0.001)	(0.017)	(0.002)	(0.018)	(0.001)
Obserations	3950	3950	3950	3950	3950	3950	3950
R^2	0.560	0.581	0.620	0.614	0.563	0.555	0.531
Panel B: Incidence of Business Closure							
Shutdown (=1)	0.022	0.185	-0.029	-0.434	0.059**	0.639**	0.089***
	(0.033)	(0.431)	(0.024)	(0.317)	(0.029)	(0.321)	(0.021)
Obserations	3950	3950	3950	3950	3950	3950	3950
R^2	0.560	0.581	0.620	0.614	0.564	0.556	0.534

Note: The table shows regression results for seven variables linked to savings. Controls for business fixed effects, and interview month-year dummies. Weighted by sampling weights and follow-up attrition probability. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

6.8 Chapter 6 - Appendix

Figure 6.A.1: The Stringency of COVID-19 measures across sub-Saharan African Countries

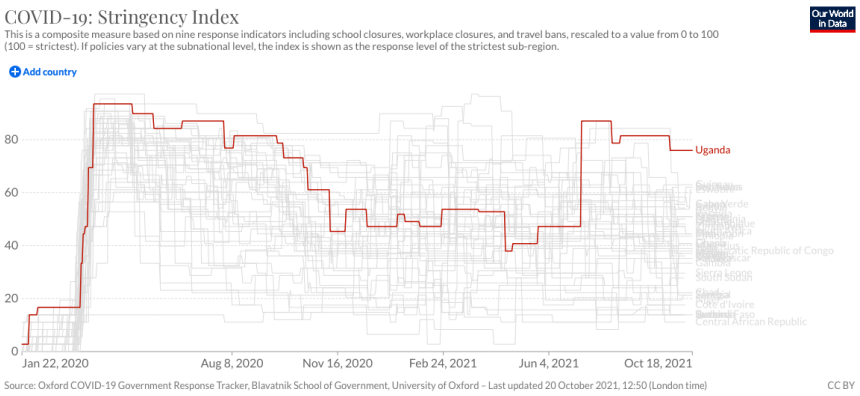
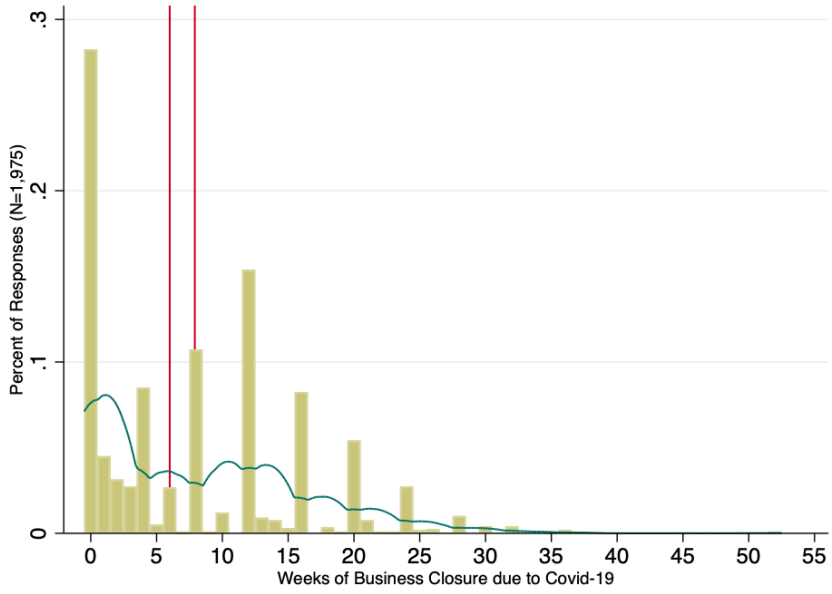


Figure 6.A.2: Distribution of Business Closure Length



Note: The blue line indicates the Kernel density estimate. The red vertical lines show the median (=6) and the mean (=7.91) of the number of weeks a business shutdown due to Covid-19 restrictions in 2020.

Figure 6.A.3: Average Length of Business Shutdown by Business Type

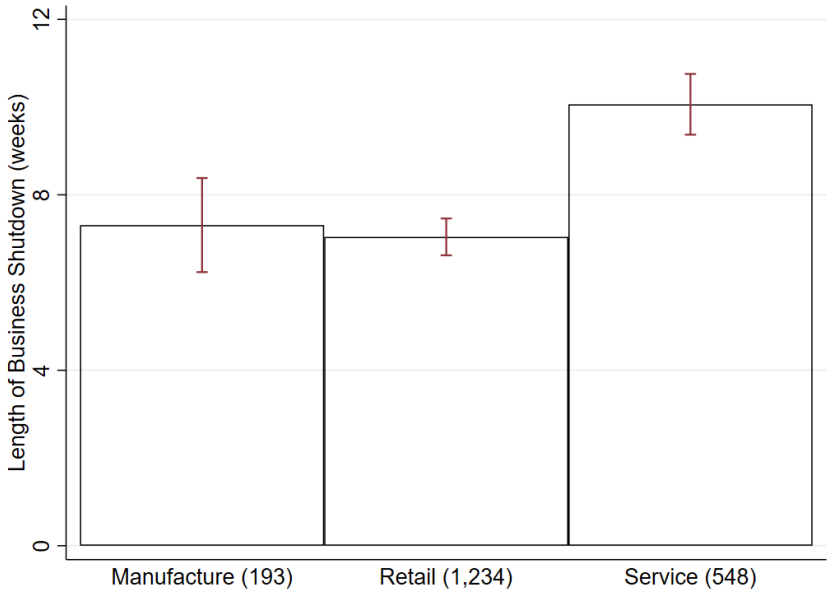


Figure 6.A.4: Self-Reported Reasons for Business Shutdown in Uganda

Main reason for business closure	
Not allowed to open due to lockdown	58%
No customers due to lockdown	22%
No stock available due to lockdown	7%
Unrelated to lockdown: making a loss	5%
Seasonal closure or usually closed in this season	5%
Other	3%
Unrelated to lockdown: could not give it time	0%
Unrelated to lockdown: did not have required expertise	0%
Total	100%

Source: Mahmud and Riley (2021) based on a phone survey conducted among 1,075 respondents from 12th to 23rd May 2020 in the Kagadi and Kyenjojo districts in Western Uganda.

Table 6.A.1: Correlates of Any Business Shutdown

	Any Closure (1)	Any Closure (2)	Any Closure (3)
Age	-0.002*	-0.002*	-0.001
	(0.001)	(0.001)	(0.001)
Services Business (=1)	0.093***	0.098***	0.100***
	(0.022)	(0.023)	(0.022)
Work Experience (Years)	-0.002	-0.003*	-0.005**
	(0.002)	(0.002)	(0.002)
Constant	0.779***	0.778***	0.773***
	(0.036)	(0.037)	(0.040)
Observations	1975	1975	1975
Inverse Sampling Weights	–	✓	✓
Inverse Attrition Weights	–	–	✓
R^2	0.017	0.019	0.024

Note: The table shows post-LASSO (OLS) regression results with the binary dependent variable being any reported business closure due to COVID-19 in 2020, clustering standard errors at the trading center-level in parentheses. Column 2 includes inverse sampling probability weights, and column 3 additional inverse weights to account for attrition. *** p<0.01, ** p<0.05, * p<0.1.

Table 6.A.2: Correlates of Length of Business Shutdown

	# Weeks of Closure (1)	# Weeks of Closure (2)	# Weeks of Closure (3)
Belongs to Majority Tribe (=1)	1.363***	1.277**	1.170**
	(0.516)	(0.544)	(0.525)
Illiterate (=1)	-1.662***	-1.468***	-1.890***
	(0.551)	(0.558)	(0.589)
Years of Education	-0.153***	-0.139***	-0.175***
	(0.042)	(0.043)	(0.043)
# Assets	-0.003	-0.005	-0.002
	(0.011)	(0.011)	(0.011)
# Acres of Owned Plots	-0.042***	-0.041***	-0.046***
	(0.010)	(0.010)	(0.008)
Work Experience (Years)	-0.051*	-0.049*	-0.068**
	(0.026)	(0.027)	(0.029)
Services Business (=1)	2.725***	2.790***	2.842***
	(0.402)	(0.414)	(0.399)
ln Profit	-0.104	-0.141*	-0.084
	(0.080)	(0.085)	(0.084)
ln Sales	-0.230	-0.307**	-0.255
	(0.149)	(0.154)	(0.187)
ln Business Expenses	-0.060	-0.029	0.025
	(0.092)	(0.093)	(0.105)
Invest (=1)	0.814*	0.807	0.840
	(0.479)	(0.505)	(0.526)
Has Loan for Emergency (=1)	-1.214*	-1.220*	-1.209
	(0.662)	(0.693)	(0.736)
MM Active (=1)	-0.378	-0.379	-0.510
	(0.411)	(0.430)	(0.460)
Has to Take Boda to Nearest MM Agent (=1)	-1.954***	-2.081***	-1.948***
	(0.683)	(0.669)	(0.686)
Uses MM to Pay Suppliers (=1)	-1.498**	-1.476*	-1.727**
	(0.713)	(0.761)	(0.728)
Constant	14.167***	15.167***	13.750***
	(2.103)	(2.128)	(2.244)
Observations	1975	1975	1975
Inverse Sampling Weights	–	✓	✓
Inverse Attrition Weights	–	–	✓
R^2	0.064	0.066	0.073

Note: The table shows post-LASSO (OLS) regression results with the dependent variable being the number of weeks of business closure (not dependent on any closure) due to COVID-19 in 2020, clustering standard errors at the trading center-level in parentheses. Column 2 includes inverse sampling probability weights, and column 3 additional inverse weights to account for attrition. *** p<0.01, ** p<0.05, * p<0.1.

Table 6.A.3: Variable Definition (1/3)

Variable	Type	Level	Description
Female (=1)	Binary	Individual	1 if respondent is female; 0 if male.
Age	Continuous	Individual	Age in years
Married (=1)	Binary	Individual	1 if respondent is married; 0 if not married.
Illiterate	Binary	Individual	1 if respondent neither able to read or write, or read only; 0 if able to read and write.
Years of Education	Continuous	Individual	Education in years
Previous Financial Education (=1)	Binary	Individual	1 if respondents reports to have received financial literacy training before baseline or participated in the intervention.
Financial Literacy (Std.)	Continuous	Individual	Number of correct financial literacy questions out of seven; standardized via item response theory.
Risk Tolerance (Std.)	Continuous	Individual	Likert scale item on "Are you generally a person who is fully prepared to take risks or do you try to avoid taking risk?" 0 if completely unwilling to take risks; 10 if fully prepared to take risk. Standardized z-score.
Able to Manage Money (1-4 Scale)	Continuous	Individual	Scale item on "If you get money, do you tend to spend it too quickly?". [1=Often, 2=Sometimes, 3=Rarely, 4=Never]
Record (=1)	Binary	Individual	Keeping a log or record of expenses and revenues for any businesses. [Yes/No]
Separate Personal (=1)	Binary	Individual	Keeping money separate for business and personal finances. [Yes/No]
Work Experience (Years)	Continuous	Individual	Number of years respondent has been working in any shop in general.
# Workers	Continuous	Individual	Number of people regularly working in the shop (excludes respondent).
Household Size	Continuous	Household	Number of people in household including respondent. A household comprises all the people who normally live and eat meals together in a home.
# Children in Household	Continuous	Household	Number of children below the age of 18 in the household.
# Assets	Continuous	Household	Sum over a list of standard items owned by the household without livestock.
Household Consumption (US$ PPP)	Continuous	Household	The value of all consumption within past four weeks. Includes food inside and outside the household, toiletries, water, rent, electricity, clothing, school fees, medical costs, leisure expenses, other necessities etc. [In US$ PPP; top coded at 99%]
Bank Account (=1)	Binary	Individual	1 if respondent has an individual bank account at a commercial bank; 0 if not.
MM Account (=1)	Binary	Individual	1 if respondent has a mobile money account; 0 if not.
MM Agent Density	Continuous	Individual	Number of reported mobile money agents within a 15-min walking distance from the business.
Has to Take Boda to Nearest MM Agent (=1)	Binary	Individual	1 if respondents needs to take a boda or motorvehicle to get to the nearest mobile money agent from the business; 0 if not.

Table 6.A.4: Variable Description (2/3)

Variable	Type	Level	Description
Retal/Wholesale Business (=1)	Binary	Individual	Main sector of business is retail or wholesale (e.g., retail of food, beverages, tobacco, motor vehicle parts, construction materials, clothes, shoes, or pharmacy).
Services Business (=1)	Binary	Individual	Main sector of business is services (e.g., hair dressing and beauty, hotels, restaurants, bars, repair shops for motor vehicles).
Manufacturing Business (=1)	Binary	Individual	Main sector of business is manufacturing (e.g., food processing, metal products, furniture making).
Trading Licence (=1)	Binary	Individual	1 if respondent has a trading license for any business, 0 if not.
Network Quality (0-10 Scale)	Continuous	Individual	Scale item on "From a scale from 0 to 10, how do you rate the quality of the network coverage at your shop during the past 7 days? 0 means very bad, no communication (phone calls, sending/receiving SMS) possible and 10 means very good quality of network coverage".
Remote Location (=1)	Binary	Location	1 if the trading center is not located along one of the paved main roads through the district; 0 if it is.
Distance to District Main Town (km)	Continuous	Location	Road distance from trading center to the center of the district main town, Fort Portal, in kilometers.
Survival (=1)	Binary	Individual	1 if respondent still runs a business during the follow-up survey, 0 if not.
ln Profit	Continuous	Individual	Total profit of the business made in the last four weeks after paying all expenses. [In Ugandan Shilling; top coded at 99% and logarithmised]
Invest (=1)	Binary	Individual	Any money invested into the business during the past 12 months. Investments could be new equipment, restocking done on top of regular restocking, new furniture or signs for advertising but do not include regular expenditures for buying new supplies or stock.
ln Invest	Continuous	Individual	Amount of money invested in the business in total during the past 12 months. [In Ugandan Shilling; top coded at 99% and logarithmised]
MM Active (=1)	Binary	Individual	1 if respondent qualifies as an active mobile money user; 0 if not. We define active mobile money use via four conditions, out of which at least one must be fulfilled: current mobile money savings or loans; any mobile money transfers made in the past three months; or ever made mobile payments directed at other businesses.
# MM Active	Continuous	Individual	Number of different mobile money services used out of the aforementioned four conditions.
MM Transfer (=1)	Binary	Individual	Transferred money to people outside the respondent's village using a mobile money account during the past three months. [Yes/No]
ln MM Transfer	Continuous	Individual	Amount of money transferred using a mobile money account in total over the past three months. [In Ugandan Shilling; top coded at 99% and logarithmised]

Table 6.A.5: Variable Description (3/3)

Variable	Type	Level	Description
MM Pay (=1)	Binary	Individual	Ever made a payment using the mobile phone. This means a payment to another business, supplier, employee or utility provider, for example with MTN MoMoPay or Airtel Money Pay. [Yes/No]
MM Business (=1)	Binary	Individual	Uses mobile money to pay for supplies (for example raw materials, items for resale). [Yes/No]
Loan (=1)	Binary	Individual	Currently any loan outstanding with a family member or friend or moneylender, an ASCA, SACCO or a microfinance institution, a commercial bank or mobile money.
ln Loan	Continuous	Individual	Amount of money currently outstanding in total across aforementioned places. [In Ugandan Shilling; top coded at 99% and logarithmised]
Never Late (=1)	Binary	Individual	1 if respondent was late with the repayment of any loan in the past five years (*in follow-up:* in the past year); 0 if not.
Transfer (=1)	Binary	Individual	During the past three months, any money transferred to people living outside own village using personal travel, asking a friend or relative to travel there, giving money to the bus driver, a SACCO, ROSCA or microfinance institution, a commercial bank or a mobile money account. [Yes/No]
ln Transfer	Continuous	Individual	Amount of money transferred in total across aforementioned places over the past three months. [In Ugandan Shilling; top coded at 99% and logarithmised]
Transfer Stolen (=1)	Binary	Individual	1 if respondent reports that some or all of the money transferred in the past 12 months got stolen or lost; 0 if not.
Saving (=1)	Binary	Individual	Currently any money saved in informal places, in a SACCO or in a ROSCA, in a commercial bank account, or a mobile money account (for example MoKash). [Yes/No]
ln Saving	Continuous	Individual	Amount of money saved in total across aforementioned places at the moment. [In Ugandan Shilling; top coded at 99% and logarithmised]
Formal Saving (=1)	Binary	Individual	Currently any money saved in a commercial bank account.
ln Formal Saving	Continuous	Individual	Amount of money saved in a commercial bank account at the moment. [In Ugandan Shilling; top coded at 99% and logarithmised]
Informal Saving (=1)	Binary	Individual	Currently any money saved in informal places (for example somewhere at home, in the garden, in the pocket, or with a friend or relative).
ln Informal Saving	Continuous	Individual	Amount of money saved informally at the moment. [In Ugandan Shilling; top coded at 99% and logarithmised]
Savings Stolen (=1)	Binary	Individual	1 if the respondent reports that some or all the money saved in any of the aforementioned places in the past 12 months was stolen or lost; 0 if not.

Bibliography

Abadie, A., S. Athey, G. W. Imbens, and J. Wooldridge (2017, November). When should you adjust standard errors for clustering? NBER Working Papers 24003, National Bureau of Economic Research, Inc.

Abiona, O. and M. F. Koppensteiner (2016). The impact of household shocks on domestic violence: Evidence from Tanzania. Technical report.

Abramitzky, R., A. Delavande, and L. Vasconcelos (2011). Marrying up: The role of sex ratio in assortative matching. *American Economic Journal: Applied Economics 3*(3), 124–57.

Agüero, J. M. (2021). Covid-19 and the rise of intimate partner violence. *World Development 137*, 105217.

Aizer, A. (2010). The gender wage gap and domestic violence. *American Economic Review 100*(4), 1847–59.

Aladangady, A. (2017). Housing wealth and consumption: Evidence from geographically-linked microdata. *American Economic Review 107*(11), 3415–46.

Ali, T. O., M. Hassan, and N. Hossain (2021). The moral and political economy of the pandemic in Bangladesh: Weak states and strong societies during Covid-19. *World Development 137*, 105216.

Almer, C., J. Laurent-Lucchetti, and M. Oechslin (2017). Water scarcity and rioting: Disaggregated evidence from Sub-Saharan Africa. *Journal of Environmental Economics and Management 86*, 193–209.

Almond, D., H. Li, and S. Zhang (2019). Land reform and sex selection in China. *Journal of Political Economy 127*(2), 560–585.

Amaral, S., P. Nishith, and S. Bhalotra (2018). Gender, crime and punishment: Evidence from women police stations in India. *Cornell University: NEUDC Retrieved from http://barrett. dyson. cornell. edu/NEUDC/paper_32. pdf*.

Anderberg, D., H. Rainer, J. Wadsworth, and T. Wilson (2016). Unemployment and domestic violence: Theory and evidence. *The Economic Journal 126*(597), 1947–1979.

Anderson, S. (2007a). The economics of dowry and brideprice. *Journal of Economic Perspectives 21*(4), 151–174.

Anderson, S. (2007b). Why the marriage squeeze cannot cause dowry inflation. *Journal of Economic Theory 137*(1), 140–152.

Angelucci, M. and V. Di Maro (2016). Programme evaluation and spillover effects. *Journal of Development Effectiveness 8*(1), 22–43.

Atalay, K., R. Edwards, and B. Y. Liu (2017). Effects of house prices on health: New evidence from Australia. *Social Science & Medicine 192*, 36–48.

Ayyagari, P. and J. L. Sindelar (2010). The impact of job stress on smoking and quitting: Evidence from the HRS. *The BE Journal of Economic Analysis & Policy 10*(1).

Banerjee, A. V., S. Cole, E. Duflo, and L. Linden (2007). Remedying education: Evidence from two randomized experiments in India. *The Quarterly Journal of Economics 122*(3), 1235–1264.

Bank of Uganda (2017). National financial inclusion strategy 2017-2022. Retrieved from `https://www.bou.or.ug/bou/bouwebsite/bouwebsitecontent/publications/special_pubs/2017/National-Financial-Inclusion-Strategy.pdf`.

Bank of Uganda (2021). Mobile money statistics. data and statistics. Retrieved from `https://www.bou.or.ug/bou/bouwebsite/PaymentSystems/dataandstat.html`.

Bao, X., S. Galiani, K. Li, and C. Long (2019). Where have all the children gone? An empirical study of child abandonment and abduction in China. Technical report, National Bureau of Economic Research.

Barcellos, S. H., L. S. Carvalho, and A. Lleras-Muney (2014). Child gender and parental investments in India: Are boys and girls treated differently? *American Economic Journal: Applied Economics 6*(1), 157–89.

Bargain, O., D. Boutin, and H. Champeaux (2019). Women's political participation and intrahousehold empowerment: Evidence from the Egyptian Arab Spring. *Journal of Development Economics 141*, 102379.

Battistini, N., G. Stoevsky, et al. (2021). The impact of containment measures across sectors and countries during the COVID-19 pandemic. Economic Bulletin Boxes, European Central Bank, Vol. 2.

Becker, G. (1991). A treaties on the family, enlarged edition, cambridge: Harvard University.

Berg, G. and B. Zia (2017). Harnessing emotional connections to improve financial decisions: Evaluating the impact of financial education in mainstream media. *Journal of the European Economic Association 15*(5), 1025–1055.

Bhalotra, S., A. Chakravarty, and S. Gulesci (2020). The price of gold: Dowry and death in India. *Journal of Development Economics 143*, 102413.

Bhattacharyya, M., A. S. Bedi, and A. Chhachhi (2011). Marital violence and women's employment and property status: Evidence from north Indian villages. *World Development 39*(9), 1676–1689.

Blumenstock, J. E., N. Eagle, and M. Fafchamps (2016). Airtime transfers and mobile communications: Evidence in the aftermath of natural disasters. *Journal of Development Economics 120*, 157–181.

Bombardini, M. and B. Li (2020). Trade, pollution and mortality in China. *Journal of International Economics 125*, 103321.

Botticini, M. and A. Siow (2003). Why dowries? *American Economic Review 93*(4), 1385–1398.

Brassiolo, P. (2016). Domestic violence and divorce law: When divorce threats become credible. *Journal of Labor Economics 34*(2), 443–477.

Brückner, M. and A. Ciccone (2011). Rain and the democratic window of opportunity. *Econometrica 79*(3), 923–947.

Bui, A. T., M. Dungey, C. V. Nguyen, and T. P. Pham (2014). The impact of natural disasters on household income, expenditure, poverty and inequality: Evidence from Vietnam. *Applied Economics 46*(15), 1751–1766.

Bulte, E., N. Heerink, and X. Zhang (2011). China's one-child policy and "the mystery of missing women": Ethnic minorities and male-biased sex ratios. *Oxford Bulletin of Economics and Statistics 73*(1), 21–39.

Bulte, E. and R. Lensink (2019). Women's empowerment and domestic abuse: Experimental evidence from Vietnam. *European Economic Review 115*, 172–191.

Cameron, A. and D. Miller (2015). A practitioner's guide to cluster-robust inference. *Journal of Human Resources 50*(2), 317–372.

Cameron, A. C., J. B. Gelbach, and D. L. Miller (2011). Robust inference with multiway clustering. *Journal of Business & Economic Statistics 29*(2), 238–249.

Campbell, J., A. S. Jones, J. Dienemann, J. Kub, J. Schollenberger, P. O'Campo, A. C. Gielen, and C. Wynne (2002). Intimate partner violence and physical health consequences. *Archives of Internal Medicine 162*(10), 1157–1163.

Card, D. and G. B. Dahl (2011). Family violence and football: The effect of unexpected emotional cues on violent behavior. *The Quarterly Journal of Economics 126*(1), 103–143.

Carpena, F., S. Cole, J. Shapiro, and B. Zia (2019). The ABCs of financial education: Experimental evidence on attitudes, behavior, and cognitive biases. *Management Science 65*(1), 346–369.

Caruso, G. D. (2017). The legacy of natural disasters: The intergenerational impact of 100 years of disasters in Latin America. *Journal of Development Economics 127*, 209–233.

Cavallo, E., A. Powell, and O. Becerra (2010). Estimating the direct economic damages of the earthquake in Haiti. *The Economic Journal 120*(546), F298–F312.

Chakraborty, T., A. Mukherjee, S. R. Rachapalli, and S. Saha (2018). Stigma of sexual violence and women's decision to work. *World Development 103*, 226–238.

Chamon, M. D. and E. S. Prasad (2010). Why are saving rates of urban households in China rising? *American Economic Journal: Macroeconomics 2*(1), 93–130.

Chaney, T., D. Sraer, and D. Thesmar (2012). The collateral channel: How real estate shocks affect corporate investment. *American Economic Review 102*(6), 2381–2409.

Chen, H., W. Qian, and Q. Wen (2021). The impact of the Covid-19 pandemic on consumption: Learning from high-frequency transaction data. *AEA Papers and Proceedings 111*, 307–11.

Chen, P.-F., M.-S. Chien, and C.-C. Lee (2011). Dynamic modeling of regional house price diffusion in Taiwan. *Journal of Housing Economics 20*(4), 315–332.

Chen, Y., H. Li, and L. Meng (2013). Prenatal sex selection and missing girls in China: Evidence from the diffusion of diagnostic ultrasound. *Journal of Human Resources 48*(1), 36–70.

Chiwaula, L., M. Matita, L. Cassim, M. Agurto, et al. (2020). Combining financial-literacy training and text-message reminders to influence mobile-money use and financial behavior among members of village savings and loan associations: Experimental evidence from Malawi. Partnership for Economic Policy Working Paper No. 2020-10.

Corno, L., N. Hildebrandt, and A. Voena (2020). Age of marriage, weather shocks, and the direction of marriage payments. *Econometrica 88*(3), 879–915.

Cunningham, S. (2020). Causal inference. *The Mixtape 1*.

Cunningham, S. and M. Shah (2018). Decriminalizing indoor prostitution: Implications for sexual violence and public health. *The Review of Economic Studies 85*(3), 1683–1715.

Cygan-Rehm, K. and C. Wunder (2018). Do working hours affect health? Evidence from statutory workweek regulations in Germany. *Labour Economics 53*, 162–171.

Dell, M. (2012). Path dependence in development: Evidence from the Mexican Revolution. Technical report.

Dell, M., B. F. Jones, and B. A. Olken (2014). What do we learn from the weather? The new climate-economy literature. *Journal of Economic Literature 52*(3), 740–98.

Demirgüç-Kunt, A., L. Klapper, D. Singer, S. Ansar, J. Hess, et al. (2020). The Global Findex Database 2017: Measuring financial inclusion and opportunities to expand access to and use of financial services. *World Bank Economic Review 34*(Supplement_1), 2–8.

Deschenes, O., H. Wang, S. Wang, and P. Zhang (2020). The effect of air pollution on body weight and obesity: Evidence from China. *Journal of Development Economics 145*(C).

Dettling, L. J. and M. S. Kearney (2014). House prices and birth rates: The impact of the real estate market on the decision to have a baby. *Journal of Public Economics 110*, 82–100.

Doi, Y., D. McKenzie, and B. Zia (2014). Who you train matters: Identifying combined effects of financial education on migrant households. *Journal of Development Economics 109*, 39–55.

Duflo, E. and A. Banerjee (2011). *Poor economics*, Volume 619. PublicAffairs.

Ebenstein, A. (2010). The "missing girls" of China and the unintended consequences of the one child policy. *Journal of Human Resources 45*(1), 87–115.

Ebenstein, A. Y. and E. J. Sharygin (2009). The consequences of the "missing girls" of China. *The World Bank Economic Review 23*(3), 399–425.

Edlund, L. (2005). Sex and the city. *Scandinavian Journal of Economics 107*(1), 25–44.

Edlund, L., H. Li, J. Yi, and J. Zhang (2013). Sex ratios and crime: Evidence from China. *Review of Economics and Statistics 95*(5), 1520–1534.

Egger, D., E. Miguel, S. S. Warren, A. Shenoy, E. Collins, D. Karlan, D. Parkerson, A. M. Mobarak, G. Fink, C. Udry, et al. (2021). Falling living standards during the COVID-19 crisis: Quantitative evidence from nine developing countries. *Science Advances 7*(6), 1–12.

Engel, J. W. (1984). Marriage in the People's Republic of China: Analysis of a new law. *Journal of Marriage and the Family*, 955–961.

Fan, C. C. and Y. Huang (1998). Waves of rural brides: Female marriage migration in China. *Annals of the Association of American Geographers 88*(2), 227–251.

Fan, H., F. Lin, and S. Lin (2020). The hidden cost of trade liberalization: Input tariff shocks and worker health in China. *Journal of International Economics 126*, 103349.

Fan, H., F. Lin, and L. Tang (2018). Minimum wage and outward FDI from China. *Journal of Development Economics 135*(C), 1–19.

Fang, H., K. N. Eggleston, J. A. Rizzo, S. Rozelle, and R. J. Zeckhauser (2012, June). The returns to education in China: Evidence from the 1986 compulsory education law. Working Paper 18189, National Bureau of Economic Research.

Fang, H., Q. Gu, W. Xiong, and L.-A. Zhou (2016). Demystifying the chinese housing boom. *NBER Macroeconomics Annual 30*(1), 105–166.

Fernandes, D., J. G. Lynch Jr, and R. G. Netemeyer (2014). Financial literacy, financial education, and downstream financial behaviors. *Management Science 60*(8), 1861–1883.

Flake, D. F. (2005). Individual, family, and community risk markers for domestic violence in Peru. *Violence Against Women 11*(3), 353–373.

Fu, S., Y. Liao, and J. Zhang (2016). The effect of housing wealth on labor force participation: Evidence from China. *Journal of Housing Economics 33*, 59–69.

Gan, L., Q. He, R. Si, and D. Yi (2019). Relocating or redefined: A new perspective on urbanization in China. NBER Working Papers 26585, National Bureau of Economic Research, Inc.

Gignoux, J. and M. Menéndez (2016). Benefit in the wake of disaster: Long-run effects of earthquakes on welfare in rural indonesia. *Journal of Development Economics 118*, 26–44.

Glaeser, E., W. Huang, Y. Ma, and A. Shleifer (2017). A real estate boom with Chinese characteristics. *Journal of Economic Perspectives 31*(1), 93–116.

Gong, Y., J. Hu, and P. J. Boelhouwer (2016). Spatial interrelations of chinese housing markets: Spatial causality, convergence and diffusion. *Regional Science and Urban Economics 59*, 103–117.

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*.

Green, D. P., A. Wilke, and J. Cooper (2019). Countering violence against women at scale: A mass media experiment in rural Uganda. *Unpublished Manuscript, Columbia University*.

Green, F. (2011). Unpacking the misery multiplier: How employability modifies the impacts of unemployment and job insecurity on life satisfaction and mental health. *Journal of Health Economics 30*(2), 265–276.

Green, F. (2020). Health effects of job insecurity. *IZA World of labor*.

Grohmann, A., L. Menkhoff, and H. Seitz (2022). The effect of personalized feedback on small enterprises' finances in Uganda. *Economic Development and Cultural Change 70*(3). Forthcoming.

GSMA (2021). State of the Industry Report on Mobile Money 2021. Retrieved from `https://www.gsma.com/sotir/`.

Guariglia, A., M. Monahan, K. Pickering, and T. Roberts (2021). Financial health and obesity. *Social Science & Medicine*, 113665.

Guerrieri, V., D. Hartley, and E. Hurst (2013). Endogenous gentrification and housing price dynamics. *Journal of Public Economics 100*, 45–60.

Gupta, A., H. Zhu, M. K. Doan, A. Michuda, and B. Majumder (2021). Economic impacts of the COVID-19 lockdown in a remittance-dependent region. *American Journal of Agricultural Economics 103*(2), 466–485.

Hagenaars, A. J., K. De Vos, M. Asghar Zaidi, et al. (1994). Poverty statistics in the late 1980s: Research based on micro-data. Luxembourg: Office for Official Publications of the European Communities.

Hale, T., N. Angrist, R. Goldszmidt, B. Kira, A. Petherick, T. Phillips, S. Webster, E. Cameron-Blake, L. Hallas, S. Majumdar, et al. (2021). A global panel database of pandemic policies (Oxford COVID-19 government response tracker). *Nature Human Behaviour 5*(4), 529–538.

Haliassos, M., T. Jansson, and Y. Karabulut (2020). Financial literacy externalities. *The Review of Financial Studies 33*(2), 950–989.

Hamdan, J. S., T. Kaiser, L. Menkhoff, and Y. Xu (2021). Financial education and spillover effects: Experimental evidence from Uganda. *Unpublished Manuscript*.

Hamdan, J. S., K. Lehmann-Uschner, and L. Menkhoff (2021). Mobile money, financial inclusion, and unmet opportunities. Evidence from Uganda. *The Journal of Development Studies*. Forthcoming.

Harari, M. and E. L. Ferrara (2018). Conflict, climate, and cells: A disaggregated analysis. *Review of Economics and Statistics 100*(4), 594–608.

Hartwig, R. and T. Lakemann (2020). When the going gets tough: Effects of the COVID-19 pandemic on informal entrepreneurs in Uganda. In brief, Hamburg: GIGA. https://www.giga-hamburg.de/en/news/when-the-going-gets-tough/.

Haushofer, J., C. Ringdal, J. P. Shapiro, and X. Y. Wang (2019). Income changes and intimate partner violence: Evidence from unconditional cash transfers in Kenya. Technical report, National Bureau of Economic Research.

Haushofer, J. and J. 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.

Haushofer, J. and J. Shapiro (2018). The long-term impact of unconditional cash transfers: Experimental evidence from Kenya. *Busara Center for Behavioral Economics, Nairobi, Kenya*.

He, G., T. Liu, and M. Zhou (2020). Straw burning, PM2.5, and death: Evidence from China. *Journal of Development Economics 145*(C).

Heath, R., M. Hidrobo, and S. Roy (2020). Cash transfers, polygamy, and intimate partner violence: Experimental evidence from Mali. *Journal of Development Economics 143*, 102410.

Hilber, C. A. L. and W. Vermeulen (2016, March). The impact of supply constraints on house prices in England. *Economic Journal 126*(591), 358–405.

Holly, S., M. H. Pesaran, and T. Yamagata (2011). The spatial and temporal diffusion of house prices in the UK. *Journal of Urban Economics 69*(1), 2–23.

Hsu, L.-C. and A. Henke (2021). Covid-19, staying at home, and domestic violence. *Review of Economics of the Household 19*(1), 145–155.

Hu, M. and X. Wang (2019). Homeownership and household formation: No homeownership, no marriage? *Journal of Housing and the Built Environment*, 1–19.

Huang, W., X. Lei, and Y. Zhao (2016). One-child policy and the rise of man-made twins. *Review of Economics and Statistics 98*(3), 467–476.

Huang, W. and Y. Zhou (2015). One-child policy, marriage distortion, and welfare loss. Technical report.

Institute of Social Science Survey, P. U. (2015). China Family Panel Studies (CFPS). Peking University Open Research Data Platform, version 37. Available at `https://doi.org/10.18170/DVN/45LCSO` (accessed May 10, 2019).

Jack, W., A. Ray, and T. Suri (2013). Transaction networks: Evidence from mobile money in Kenya. *American Economic Review 103*(3), 356–361.

Jack, W. and T. Suri (2014). Risk sharing and transactions costs: Evidence from Kenya's mobile money revolution. *American Economic Review 104*(1), 183–223.

Jia, R. (2014). Weather shocks, sweet potatoes and peasant revolts in historical China. *The Economic Journal 124*(575), 92–118.

Josephson, A., T. Kilic, and J. D. Michler (2021). Socioeconomic impacts of COVID-19 in low-income countries. *Nature Human Behaviour 5*(5), 557–565.

Kaiser, T., A. Lusardi, L. Menkhoff, and C. Urban (2021). Financial education affects financial knowledge and downstream behaviors. *Journal of Financial Economics*.

Kaiser, T. and L. Menkhoff (2017). Does financial education impact financial literacy and financial behavior, and if so, when? *The World Bank Economic Review 31*(3), 611–630.

Kaiser, T. and L. Menkhoff (2018). Active learning improves financial education. *CRC 190 Discussion Paper Series No. 131*.

Kansiime, M. K., J. A. Tambo, I. Mugambi, M. Bundi, A. Kara, and C. Owuor (2020). Covid-19 implications on household income and food security in Kenya and Uganda: Findings from a rapid assessment. *World Development 137*, 105199.

Khalil, U. and S. Mookerjee (2019). Patrilocal residence and women's social status: Evidence from South Asia. *Economic Development and Cultural Change 67*(2), 401–438.

Kwanga, G. M., T. Shabu, and E. M. Adaaku (2017). Natural Disasters and Crime Incidence: A Case of 2012 Flooding in Benue State, Nigeria. *International Journal of Geology, Agriculture and Environmental Sciences 5*(5).

Lackner, S. (2018). Earthquakes and economic growth. Technical report, FIW Working Paper.

Lakemann, T., J. Lay, and T. Tafese (2020). Africa after the Covid-19 lockdowns: Economic impacts and prospects. GIGA Focus Africa (6). https://www.giga-hamburg.de/en/publications/21606562-africa-after-covid-19-lockdowns-economic-impacts-prospects/.

Le Nestour, A., S. Mbaye, J. Sandefur, and L. Moscoviz (2020). Covid19 phone survey Senegal. *Harvard Dataverse*.

Lei, X., Y. Shen, J. P. Smith, and G. Zhou (2017). Sibling gender composition's effect on education: Evidence from China. *Journal of Population Economics 30*(2), 569–590.

Levendosky, A. A., G. A. Bogat, and C. Martinez-Torteya (2013). Ptsd symptoms in young children exposed to intimate partner violence. *Violence Against Women 19*(2), 187–201.

Li, C., E. S. Ford, L. C. McGuire, and A. H. Mokdad (2007). Increasing trends in waist circumference and abdominal obesity among us adults. *Obesity 15*(1), 216–216.

Li, H., J. Yi, and J. Zhang (2011). Estimating the effect of the one-child policy on the sex ratio imbalance in China: Identification based on the difference-in-differences. *Demography 48*(4), 1535–1557.

Li, J. and Z. Liu (2018). Housing stress and mental health of migrant populations in urban China. *Cities 81*, 172–179.

Li, L. and X. Wu (2019). Housing price and intergenerational co-residence in urban China. *Journal of Housing Economics 45*, 101596.

Liang, W., M. Lu, and H. Zhang (2016). Housing prices raise wages: Estimating the unexpected effects of land supply regulation in China. *Journal of Housing Economics 33*, 70–81.

Luca, D. L., E. Owens, and G. Sharma (2015). Can alcohol prohibition reduce violence against women? *American Economic Review 105*(5), 625–29.

Macmillan, R. and R. Gartner (1999). When she brings home the bacon: Labor-force participation and the risk of spousal violence against women. *Journal of Marriage and the Family*, 947–958.

Mahmud, M. and E. Riley (2021). Household response to an extreme shock: Evidence on the immediate impact of the Covid-19 lockdown on economic outcomes and well-being in rural Uganda. *World Development 140*, 105318.

Malik, K., M. Meki, J. Morduch, T. Ogden, S. Quinn, and F. Said (2020). COVID-19 and the Future of Microfinance: Evidence and Insights from Pakistan. *Oxford Review of Economic Policy 36*(Supplement_1), S138–S168.

McInerney, M., J. M. Mellor, and L. H. Nicholas (2013). Recession depression: Mental health effects of the 2008 stock market crash. *Journal of Health Economics 32*(6), 1090–1104.

McKenzie, D. and S. Puerto (2021). Growing markets through business training for female entrepreneurs: A market-level randomized experiment in Kenya. *American Economic Journal: Applied Economics 13*(2), 297–332.

McKenzie, D. and C. Woodruff (2017). Business practices in small firms in developing countries. *Management Science 63*(9), 2967–2981.

Meng, X. (2012). Labor market outcomes and reforms in China. *Journal of Economic Perspectives 26*(4), 75–102.

Menon, S. (2020). The effect of marital endowments on domestic violence in India. *Journal of Development Economics 143*, 102389.

Merli, M. G. and A. E. Raftery (2000). Are births underreported in rural China? manipulation of statistical records in response to China's population policies. *Demography 37*(1), 109–126.

Merotto, D. (2019). Uganda: Jobs strategy for inclusive growth. *Jobs Series*. Washington, D.C.: World Bank Group. License: Creative Commons Attribution CC BY 3.0 IGO. Retrieved from `http://documents.worldbank.org/curated/en/693101582561426416/Uganda-Jobs-Strategy-for-Inclusive-Growth`.

Mian, A. and A. Sufi (2011). House prices, home equity-based borrowing, and the US household leverage crisis. *American Economic Review 101*(5), 2132–56.

Miguel, E. (2005). Poverty and witch killing. *The Review of Economic Studies 72*(4), 1153–1172.

Miguel, E., S. Satyanath, and E. Sergenti (2004). Economic shocks and civil conflict: An instrumental variables approach. *Journal of Political Economy 112*(4), 725–753.

Miller, A. R. and C. Segal (2019). Do female officers improve law enforcement quality? Effects on crime reporting and domestic violence. *The Review of Economic Studies 86*(5), 2220–2247.

Miller, M., J. Reichelstein, C. Salas, and B. Zia (2015). Can you help someone become financially capable? A meta-analysis of the literature. *The World Bank Research Observer 30*(2), 220–246.

Munyegera, G. K. and T. Matsumoto (2016). Mobile money, remittances, and household welfare: Panel evidence from rural Uganda. *World Development 79*, 127–137.

Mutter, J. C. (2015). *The disaster profiteers: How natural disasters make the rich richer and the poor even poorer*. Macmillan.

Naved, R. T., M. A. Mamun, K. Parvin, S. Willan, A. Gibbs, M. Yu, and R. Jewkes (2018). Magnitude and correlates of intimate partner violence against female garment workers from selected factories in Bangladesh. *PLoS One 13*(11), e0204725.

Niles, A. N. and A. O'Donovan (2019). Comparing anxiety and depression to obesity and smoking as predictors of major medical illnesses and somatic symptoms. *Health Psychology 38*(2), 172–181.

Ñopo, H. (2009). The gender wage gap in Peru 1986-2000: Evidence from a matching comparisons approach.

OECD (2020a). Recommendation of the Council on Financial Literacy, OECD/LEGAL/0461.

OECD (2020b). *Violence against women (Attitudes toward violence)*. doi: 10.1787/f1eb4876-en (Accessed on 30 December 2020).

Ohrnberger, J., E. Fichera, and M. Sutton (2017). The relationship between physical and mental health: A mediation analysis. *Social Science & Medicine 195*, 42–49.

Ozili, P. K. and T. Arun (2020). Spillover of Covid-19: Impact on the global economy. *Available at SSRN 3562570*.

PDC (2015). *National disaster preparedness baseline assessment: Peru*. Pacific Disaster Center.

Plichta, S. B. (2004). Intimate partner violence and physical health consequences: Policy and practice implications. *Journal of Interpersonal Violence 19*(11), 1296–1323.

Prentice, C., D. McKillop, and D. French (2017). How financial strain affects health: Evidence from the Dutch National Bank Household Survey. *Social Science & Medicine 178*, 127–135.

Radloff, L. S. (1977). The ces-d scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement 1*(3), 385–401.

Raghupathi, W. and V. Raghupathi (2018). An empirical study of chronic diseases in the United States: A visual analytics approach to public health. *International Journal of Environmental Research and Public Health 15*(3), 431.

Rahman, H. Z. and I. Matin (2020). Livelihoods, coping, and support during Covid-19 crisis. *Dhaka, BRAC Institute of Governance and Development*.

Rao, V. (1993). The rising price of husbands: A hedonic analysis of dowry increases in rural india. *Journal of Political Economy 101*(4), 666–677.

Rees, D. I. and J. J. Sabia (2013). Forced intercourse, mental health, and human capital. *Southern Economic Journal 80*(2), 324–344.

Riley, E. (2018). Mobile money and risk sharing against village shocks. *Journal of Development Economics 135*, 43–58.

Ritchie, H. and M. Roser (2019). Gender ratio. *Our World in Data*. https://ourworldindata.org/gender-ratio.

Rong, Z., W. Wang, and Q. Gong (2016). Housing price appreciation, investment opportunity, and firm innovation: Evidence from China. *Journal of Housing Economics 33*, 34–58.

Roser, M. and E. Ortiz-Ospina (2013). Global extreme poverty. *Our World in Data*. https://ourworldindata.org/extreme-poverty.

Roy, S., M. Hidrobo, J. Hoddinott, and A. Ahmed (2019). Transfers, behavior change communication, and intimate partner violence: Postprogram evidence from rural Bangladesh. *Review of Economics and Statistics 101*(5), 865–877.

Sá, F. (2015). Immigration and house prices in the UK. *Economic Journal 125*(587), 1393–1424.

Sabia, J. J., A. K. Dills, and J. DeSimone (2013). Sexual violence against women and labor market outcomes. *American Economic Review 103*(3), 274–78.

Saiz, A. (2010). The geographic determinants of housing supply. *The Quarterly Journal of Economics 125*(3), 1253–1296.

Sekhri, S. and A. Storeygard (2014). Dowry deaths: Response to weather variability in India. *Journal of Development Economics 111*, 212–223.

Sen, A. (1992). Missing women. *BMJ: British Medical Journal 304*(6827), 587.

Stevenson, B. and J. Wolfers (2006). Bargaining in the shadow of the law: Divorce laws and family distress. *The Quarterly Journal of Economics 121*(1), 267–288.

Stroebel, J. and J. Vavra (2019). House prices, local demand, and retail prices. *Journal of Political Economy 127*(3), 1391–1436.

Suárez Farfán, A. V., F. d. M. Monzón Rodríguez, J. Valenzuela Jiménez, J. A. Calderón Marmolejo, K. Pérez Díaz, L. Huaylinos Oré, M. T. Gutiérrez Pajares, M. Gamarra Valencia, M. Vargas Cuno, and S. Zapata Gonzales (2016). Violencia basada en género: Marco conceptual para las políticas públicas y la acción del estado.

Sun, A. and Q. Zhang (2020). Who marries whom in a surging housing market? *Journal of Development Economics 146*, 102492.

Suri, T. (2017). Mobile money. *Annual Review of Economics 9*, 497–520.

Sweet, E., A. Nandi, E. K. Adam, and T. W. McDade (2013). The high price of debt: Household financial debt and its impact on mental and physical health. *Social Science Medicine 91*, 94–100.

Tan, Z., S.-J. Wei, and X. Zhang (2021). Deadly discrimination: Implications of "missing girls" for workplace safety. *Journal of Development Economics 152*, 102678.

Teye, A. L. and D. F. Ahelegbey (2017). Detecting spatial and temporal house price diffusion in the netherlands: A Bayesian network approach. *Regional Science and Urban Economics 65*, 56–64.

Tibshirani, R. (1996). Regression shrinkage and selection via the LASSO. *Journal of the Royal Statistical Society: Series B (Methodological) 58*(1), 267–288.

Trako, I., M. M. Sviatschi, and G. Kavanaugh (2018). Access to justice, gender violence and children: Evidence from women's justice centers in Peru. Technical report, Princeton University, Woodrow Wilson School of Public and International Research Program in Development Studies.

Tseng, T.-S. and H.-Y. Lin (2008). Gender and age disparity in health-related behaviors and behavioral patterns based on a national survey of Taiwan. *International Journal of Behavioral Medicine 15*(1), 14–20.

Tur-Prats, A. (2019). Family types and intimate partner violence: A historical perspective. *Review of Economics and Statistics 101*(5), 878–891.

UBOS (2021). The population projection of Kabarole district from 2015-2020. Uganda Bureau of Statistics. Retrieved from http://catalog.data.ug/dataset/uganda-s-district-population-projections-2015-2020/resource/904a83cf-3735-4d98-ac02-8fd6188b25a4.

Valenti, M., F. Masedu, M. Mazza, S. Tiberti, C. Di Giovanni, A. Calvarese, R. Pirro, and V. Sconci (2013). A longitudinal study of quality of life of earthquake survivors in L'Aquila, Italy. *BMC Public Health 13*(1), 1143.

Wang, F., C. Milner, and J. Scheffel (2021). Labour market reform and firm-level employment adjustment: Evidence from the hukou reform in China. *Journal of Development Economics 149*, 102584.

Waxman, A., Y. Liang, S. Li, P. J. Barwick, and M. Zhao (2020). Tightening belts to buy a home: Consumption responses to rising housing prices in urban China. *Journal of Urban Economics 115*(C).

WDI (2020). World Development Indicators. DataBank. Washington, D.C.: The World Bank. Retrieved from https://databank.worldbank.org/reports.aspx?source=2&country=UGA.

Wei, S.-J. and X. Zhang (2011a). The competitive saving motive: Evidence from rising sex ratios and savings rates in China. *Journal of Political Economy 119*(3), 511–564.

Wei, S.-J. and X. Zhang (2011b). Sex ratios, entrepreneurship, and economic growth in the People's Republic of China. Technical report, National Bureau of Economic Research.

Wei, S.-J., X. Zhang, and Y. Liu (2017). Home ownership as status competition: Some theory and evidence. *Journal of Development Economics 127*, 169–186.

Weissman, M. M., D. Sholomskas, M. Pottenger, B. A. Prusoff, and B. Z. Locke (1977). Assessing depressive symptoms in five psychiatric populations: A validation study. *American Journal of Epidemiology 106*(3), 203–214.

Weitzman, A. and J. A. Behrman (2016). Disaster, disruption to family life, and intimate partner violence: The case of the 2010 earthquake in Haiti. *Sociological Science 3*, 167–189.

WHO (2013). *Global and regional estimates of violence against women: prevalence and health effects of intimate partner violence and non-partner sexual violence*. World Health Organization.

Wong, I. S., D. Dawson, and H. P. Van Dongen (2019). International consensus statements on non-standard working time arrangements and occupational health and safety. *Industrial Health 57*(2), 135–138.

World Bank (2020). Uganda Economic Update, 16th Edition, December 2020: Investing in Uganda's Youth.

WorldBank (2020). *World Development Indicators (Population, total).* URL: https://data.worldbank.org/indicator/SP.POP.TOTL?locations=PE (Accessed on 30 December 2020).

Wrenn, D. H., J. Yi, and B. Zhang (2019). House prices and marriage entry in China. *Regional Science and Urban Economics 74*, 118–130.

Yang, G., J. Hu, K. Q. Rao, J. Ma, C. Rao, and A. D. Lopez (2005). Mortality registration and surveillance in China: History, current situation and challenges. *Population Health Metrics 3*(1), 3.

Yaribeygi, H., Y. Panahi, H. Sahraei, T. P. Johnston, and A. Sahebkar (2017, July). The impact of stress on body function: A review. *EXCLI Journal 16*, 1057–1072.

Zhang, J. and W. Chan (1999). Dowry and wife's welfare: A theotrical and empirical analysis. *Journal of Political Economy 107*(4), 786–808.

Zhao, Y., J. Strauss, and G. Yang (2015). China Health and Retirement Longitudinal Study (2013 wave2). Peking University Open Research Data Platform, V1. Available at `https://doi.org/10.18170/DVN/AQIU6C` (accessed September 20, 2019).