ESSAYS ON HOUSEHOLD ECONOMICS

Maximilian Schwefer

München 2018

ESSAYS ON HOUSEHOLD ECONOMICS

Inaugural-Dissertation

zur Erlangung des Grades

Doctor oeconomiae publicae (Dr. oec. publ.)

an der Ludwig–Maximilians–Universität München

2018

vorgelegt von

Maximilian Schwefer

Referent: Prof. Panu Poutvaara, Ph.D.
Korreferent: Prof. Helmut Rainer, Ph.D.
Promotionsabschlussberatung: 30. Januar 2019

Datum der mündlichen Prüfung: 24. Januar 2019
Namen der Berichterstatter: Prof. Panu Poutvaara, Ph.D., Prof. Helmut Rainer, Ph.D., Prof. Helmut Farbmacher, Ph.D.

Acknowledgments

First and foremost, I would like to thank my supervisor Panu Poutvaara. I am very grateful for his diligent feedback and constructive mindset. He was generous with his time and it was a great joy to work with him. His guidance has been an invaluable contribution to this dissertation. I would also like to thank Helmut Rainer for his feedback. Our collaboration focused particularly on the second paper, which immensely benefitted from his comments and expertise. I am grateful for his willingness to serve as a second supervisor, and on my dissertation committee. I owe my thanks to Helmut Farbmacher, who will serve on the dissertation committee as well.

This dissertation would not have been possible without the support of SOS Children's Villages (SOS). I would like to express my gratitude to Astrid Brandl, Stephanie Fähnle, Dario Peter, and Stewart Wilms. They have been great facilitators and mentors who enabled me to understand how this impressive organization works. I particularly would like to thank Dario for his enthusiasm, expertise and supernatural response times. His help with my data requests has been invaluable.

I am grateful to the ifo Institute, the Munich Graduate School of Economics and the Ludwig Maximilian University of Munich for providing me with an excellent research environment. The financial support given by the Foundation of German Business and the German Federal Ministry of Education and Research is gratefully appreciated. Research reported in this dissertation was supported by the Population Studies and Training Center (PSTC) at Brown University through the generosity of the Eunice Kennedy Shriver National Institute of Child Health and Human Development (P2C HD041020). I am indebted to Susan Short for inviting me to Providence. My time at Brown has been a great personal experience and important cornerstone of this work. I am very appreciative of my former colleagues at the PSTC. Besides Susan, I thank Chinyere Agbai, Thomas Alarie, Karen Ball, Quinn Bernier, Jo Fisher, and Susan Silveira for welcoming me to their community.

The first essay is a joint piece of work with Panu Poutvaara, whom I thank for his contributions. We have received helpful comments by Alex Eble, Terry-Ann Craigie, Markus Goldstein and his team at the World Bank. Input from seminar participants at Brown University and the ifo Institute is gratefully acknowledged. The second essay built on the outstanding advice given by Elli Hayati. Her expertise significantly shaped the paper. I am grateful to Yudi Kartiwa and his team at SOS Indonesia, whose detailed knowledge of local conditions helped me considerably. The third essay has benefitted from comments and contributions by Kieron Barclay, Ryan Born, Jan-Philipp Suchta and of workshop participants at the ifo Institute and the German Economic Society's annual meeting.

I am indebted to my office mate Till Nikolka for his continuous support ranging from econometric advice to diverting conversations over coffee. My sincere thanks go to André Schmelzer and Johannes Hemker, who have invested countless hours in coaching and challenging me. I am also grateful for input from Kai Schulze, Clara Welteke, and Julian Zuber. I am glad to call all of them my friends.

Special thanks go to Charlotte for her love and unfailing support. I am very grateful for her continuous moral encouragement and our thought partnership, both of which significantly contributed to this dissertation.

I would like to thank my sister, parents, and grandparents for everything else. Little to nothing would have been possible without the love, support and help from my family.

Meinen Eltern

Contents

List of Tables

List of Figures

Introduction

This dissertation consists of three independent essays in household economics. The essays offer novel empirical perspectives in the domains of decision-making power, domestic violence, and educational achievement of siblings respectively. What all essays have in common is their aim to advance our knowledge of household economics in general and our understanding of intra-household dynamics in specific. They are written with the aspiration to provide insights of relevance to policy-makers and academics alike.

Understanding the economics of the household is important. How household members cope with their limited resources and divergent interests moderates many important life outcomes, for example in the spheres of labor force participation, individual health, and educational achievement. As a subfield of economics, household economics is concerned with the determinants, dynamics, and consequences of household decisions and behavior. The subfield's theoretical underpinning has benefitted from a paradigm shift away from the unitary household model to the cooperative bargaining approach that also informs this dissertation (McElroy and Horney, 1981; Lundberg and Pollak, 1996, 1994). The subfield has also made very valuable empirical contributions. It is thanks to household economics that we can better explain changes in fertility patterns (Ashraf, Field, and Lee, 2014), household consumption (Duflo and Udry, 2004; Duflo, 2000), gender norms (Jensen and Oster, 2009), domestic violence (Aizer, 2011), and divorce rates (Gray, 1998). This dissertation's three essays build on this legacy.
The first essay is joint work with Panu Poutvaara. We explore the methodological and empirical implications of spousal discordance over female decision-making power. There is great academic and public interest in understanding the causes and consequences of female decision-making power in households. However, as *true* power remains unobserved, studies rely on proxies. Surveys are widely considered to be the most unequivocal proxy for female decision-making power, and a growing number of publications relies on them. So far, little attention has been paid to the systematic differences between males' and females' responses to such surveys. Our analyses are based on data from the last three waves of the Indonesian Family Life Survey (IFLS). In our first analysis, we survey whether discordance can be considered *random* or whether it is related to couple attributes. On average, the share of discordant, non-matching responses varies from 33 to 51 percent across decision domains. We find the prevalence

of specific types of discordance to vary with the female income share and other proxies of female power. We confirm that discordance is not *random*, but systematically related to *real* power. In our second analysis, we find that female labor supply and rates of contraceptive use are higher where both partners perceive female power, relative to cases of strong discordance about the female role and relative to concordant reports on the husband as the sole decision-maker. We propose that predictions of these outcomes can be improved by taking both spouses' perceptions of female power into consideration. In our third analysis, we explore changes in relative female economic resources. When female economic resources change, both spouses' shifts in their perception of female decision-making power are in the same order of magnitude. The essay is the first study to use an exogenous income shock to address concerns of potential endogeneity. Summarizing, results suggest that at the population level, one-sided surveys provide a good proxy for shifts in male and female perception of female decision-making power in reaction to a change in female economic resources. However, cross-section estimates suggest that females' outcomes will differ in those families where both spouses' reactions correlate vs. those where only one spouse changes its perception of female power.

The second essay contributes to an emerging strand of literature, that documents the association of disasters and domestic violence, by employing novel panel data from Indonesia. The increasing frequency and severity of natural disasters and their significant impact on affected populations make them subject to extensive research. However, the range of documented outcomes in longitudinal studies has been comparably narrow, and the majority of publications relies on post-disaster data only. This is why many scholars have articulated the need for more (longitudinal) evidence on how disasters affect populations. The second essay is the first study to provide pre- and post-treatment family-level data on the impact of volcano eruptions on domestic violence. Besides, it offers evidence for the channels by which natural disasters might cause domestic violence. I use survey and observational data from 2,024 families in Indonesia, of which a subset has been exposed to two volcano eruptions in late 2013 and early 2014. I estimate the impact of the two eruptions on rates of domestic violence and four alternative outcomes with a difference-in-differences approach. Previous evidence shows that natural disasters cause mental distress in affected populations via various channels. Suffering from mental distress, individuals can develop feelings of aggression and outward, violent interpersonal behavior. I confirm a significant increase in domestic violence in all observed households in the treatment group. Affected communities suffer from lower average household expenditures which are expected to increase distress. Further, increased rates of alcohol/drug abuse and lowered emotional well-being

in affected populations point to increased levels of mental distress. A subsample of families, who have been displaced in a previous eruption, displays substantially higher risk of displaying domestic violence after the volcanoes' eruptions. It is argued that this risk is founded on their previous loss of livelihood, lack of a social network and augmented feelings of a threat of disaster recurrence. Policymakers and emergency response organizations should consider the multi-fold non-economic outcomes of natural disasters when designing interventions.

The third essay is concerned with birth order effects and educational achievement in low- and middle-income countries. A rich body of research recommends that firstborns score better across a range of life outcomes, including educational achievement. While the general existence of birth order effects in high-income countries is widely acknowledged, there is substantial debate over *why* birth order effects exist, and what might explain the heterogeneity in findings in low- and middle-income countries. To address these questions, I use a novel dataset covering 26,898 observations of 4,362 biologically related siblings living in long-term alternative care families in 54 countries. Data are provided by SOS Children's Villages. This study is the first one to provide evidence from a broad set of low- and middle-income countries across multiple continents. Results indicate that birth order effects in low- and middle-income countries are consistent with those in high-income countries. Additionally, this study offers insight into three sources of heterogeneity in birth order effects. First, findings advise that sibships that have suffered extreme economic or emotional hardship (for instance sexual abuse, domestic violence) show attenuated birth order effects compared to other sibships. Secondly, gender-specific effects are identified for Asia, where the firstborn advantage is significantly smaller for girls, compared to boys, suggesting parental gender preferences. Individual hardship *within* a society seems to be as relevant as differences in development *between* societies. This is compatible with previous evidence indicating effect diminishment and reversal for households of low socioeconomic status in high-income countries and reversed birth order effects in low- and middle-income countries. Thirdly, intra-family comparisons of biologically unrelated children of the same biological birth order suggest the existence of tutoring effects between unrelated siblings in their alternative care families. This evidence is only suggestive as large standard errors prevent statements on statistically significant differences between children of the same biological birth order. Holding biological birth order constant, I find superior outcomes for older children ranked *higher* in their alternative care family. I propose more tutoring opportunities as a potential explanation. The finding is consistent with the confluence model – one possible explanation for birth order effects. The conflu-

ence model attributes birth order effects to changing dynamics of social interaction within the family; of which tutoring between children is one element. These findings advance the debates on determinants of educational achievement and the formation of human capital in low- and middle-income countries. They also suggest reasons for how and when intra-family differences in human capital emerge. Larger (alternative care) families could particularly benefit from exploring tutoring as a measure to let children grow personally and intellectually. The results can inform policy making and development interventions by helping to prioritize individuals in highest need.

1. Husbands' and wives' diverging perceptions of female decision-making power: methodological implications and empirical evidence from Indonesia[1]

1. Joint work with Panu Poutvaara

1.1. Introduction

There is great academic and public interest in understanding the causes and consequences of female decision-making power in households. This interest is driven by both concerns about female empowerment and the notion that shifts in bargaining power from men to women are associated with *desirable* changes in household behavior.[2] However, as *true* power remains unobserved, studies rely on proxies. Surveys are widely considered to be the most unequivocal proxy for female decision-making power (Majlesi, 2016; Malhotra, Schuler, and Boender, 2002). Accordingly, a growing number of publications in economics and related fields relies on decision-making surveys.[3] Surprisingly, the vast majority of scholars does not account for the possibility of divergent spousal responses. This is striking as power is rarely exerted in a social vacuum. Studies that consider both spouses' perspectives find high rates of discordance about female decision-making power (We will refer to divergent spousal responses as discordant statements. Matching spousal responses will be referred to as concordant).

This paper explores spousal discordance in statements on female decision-making power. We contribute to existing literature in three ways. First, we confirm that discordance is a phenomenon that is common and *non-random*: specific types of discordance are systematically related to female power proxies such as the female share in the household's income. Additionally, we show that prediction of outcomes can be improved if both spouses' perspectives are taken into account. This study is the first to provide a detailed account of the relationship between specific types of discordance and females' labor market outcomes. Thirdly, we evaluate the elasticity of male and female perception of female decision-making power if relative female economic resources change. We show that spouses' changes in their perceptions are not statistically singificantly different from each other. This is the first study to use an exogenous income shock to address concerns of potential endogeneity.

Our contribution to existing literature is thereby both empirical and methodological. Only a few previous studies consider both spouses' responses. They find discordance to occur frequently and in a non-random manner. Even fewer studies associate specific types of discordant statements with outcomes. So far only a limited number

2. For example, Duflo and Udry (2004) exploit variation in crop yield to show that increases in relative female income cause increased food and educational expenditure.

3. For a review see Donald et al. (2017). For examples see Anderson and Eswaran (2009), Bruins (2017), Jensen and Oster (2009), and Majlesi (2016).

of outcomes has been related to discordant statements (Ambler et al., 2017; Becker, Fonseca-Becker, and Schenck-Yglesias, 2006; Jejeebhoy, 2002).[4] This study benefits from an uncommon breadth and depth of information on households, reaching from household decision-making to labor outcomes and contraceptive use. It is based on data from the Indonesian Family Life Survey (IFLS), a panel survey administered to Indonesian households since 1993. A particular merit of the IFLS is a household decision-making module that provides both partners' perceptions of who makes decisions across 13 different household domains, for instance, on children's health or saving decisions.

In our first analysis, we survey whether discordance can be considered *random* or whether it is related to couple attributes. On average, the share of discordant, non-matching responses varies from 33 to 51 percent across decision domains. We find the prevalence of specific types of discordance to vary with female income share and other proxies for female power. Findings suggest, that discordance is not *random*, but systematically related to *real* power.

Secondly, we relate discordance to two proxies for female power, female labor force participation and contraceptive use. We find that female labor supply and rates of contraceptive use are higher when both partners perceive female power, relative to cases of strong discordance about the female role and relative to concordant reports on the husband as the sole decision maker. We also find that predictions of these outcomes can be improved by taking both spouses' perceptions of female power into consideration. This is based on additional analysis, in which we hold the wife's perception constant and observe outcomes for all possible responses by the husband.

Thirdly, we explore the association of variation in female economic resources and the perception of female decision-making power from both spouses' perspectives. This allows us to predict the potential bias in studies that evaluate the impact of economic interventions on female empowerment and that use one-sided surveys to do so. First, we run a fixed effects model without exogenous variation of income. In the second model, we use a cross-section setting with an exogenous variation of spousal income shares. In both models, we associate changes in relative female economic resources with the male and female perception of female decision-making power. We find that

4. Becker, Fonseca-Becker, and Schenck-Yglesias (2006) and Jejeebhoy (2002) study health outcomes while Ambler et al. (2017) proxy female well-being across multiple dimensions, including labor supply. Compared to this paper, this account of labor supply is less detailed and is limited to a binary variable, indicating whether the wife works more or less than 10.5 hours per day.

both spouses' changes in perception are in the same order of magnitude. This suggests that one-sided surveys will provide a good proxy if one is interested in the average adaptation of perception following an economic shock.

In summary, results suggest that at the population level, one-sided surveys provide a good proxy for shifts in male and female perception of female decision-making power in reaction to a change in female economic resources. However, cross-section estimates suggest that females' outcomes will differ in those families where both spouses' reactions correlate vs. those where only one spouse changes its perception of female power. We conclude that in the context of households, both spouses' perspectives should be taken into account.

The remainder of this paper is structured as follows. In section 1.2, we present an overview of empirical studies that employ one- and two-sided surveys to proxy decision-making power. Our focus will be on studies that explicitly describe discordance in response patterns between husbands and their wives. In section 1.3, we will describe the data. In section 1.4, we will introduce the typology we use to classify spousal statement combinations and prevalence rates of specific types. Then, we will present our main results. In section 1.5, we will survey couple attributes which are associated with different types of con- and discordance. In section 1.6, we will study outcomes of different types of concordant and discordant couples. A particular focus will be on the merit of considering both spouses' perspectives. In section 1.7, we will assess the relationship between variations in female economic and decision-making power. In section 1.8, we will run different robustness checks. We will conclude in section 1.9 by discussing the implications and limitations of our findings.

1.2. Discordance in previous literature

The following account of the previous literature is divided into three subsections. In subsection 1.2.1, we will motivate this study by documenting the wide use of decision-making surveys in economics and related literature. In subsection 1.2.2, we will present previous evidence on discordance in such surveys. In the final subsection 1.2.3, we will discuss potential moderators of discordance.

1.2.1. Measuring female decision-making power

This study is concerned with female decision-making power in the household. Following Kabeer (1999), we understand female decision-making power as one dimension of female agency and female agency as one dimension of female empowerment.[5] Our study thereby relates to the wider empowerment literature (see Duflo (2012) for a review) and the agency literature in specific (see Donald et al. (2017) for a review). There is great academic and public interest in understanding the causes and consequences of female decision-making power in households. This interest is driven by both concerns about female empowerment and the notion that shifts in bargaining power from men to women are associated with *desirable* changes in household behavior (Bruins, 2017; Duflo and Udry, 2004; Jensen, 2012; Lundberg, Pollak, and Wales, 1997; Majlesi, 2016; Bandiera et al., 2018).[6]

Due to the high interest in female empowerment, many scholars are concerned with how to best measure it. Majlesi (2016) posits that *asking* household members about their decision-making power is the most unequivocal way to capture it. Accordingly, many empirical studies rely on surveys to determine decision-making power in the household – in economics (Anderson and Eswaran, 2009; Banerjee et al., 2015; Bruins, 2017; Jensen and Oster, 2009; Majlesi, 2016) as well as in related disciplines such as demography and sociology (Ebot, 2014; Hayes and Boyd, 2017; Kabeer, Mahmud,

5. Kabeer (1999) defines empowerment as command over resources, agency, and achievements. See Malhotra and Schuler (2005) and Malhotra, Schuler, and Boender (2002) for a discussion of female empowerment concepts in international development.

6. For example, Majlesi (2016) documents a positive association between female labor market conditions, female decision-making power, and child health. Duflo and Udry (2004) use gender-specific specialization in crops and weather-induced variation of crop-specific income. Holding total family income constant, higher yields for *female* crops cause higher relative spending on food. Along the same lines, Wang (2014) finds a reduced consumption of *male favored goods* such as cigarettes following a re-allocation of property rights in China. For Bangladesh, Heath (2014) finds that women report being more confident to assert their own decision-making power towards their husbands if they earn a salary.

and Tasneem, 2011; Mboane and Bhatta, 2015; Rahman and Rao, 2004). Malhotra, Schuler, and Boender (2002, p. 26) show that decision-making indicators are some of the "most frequently used indicators" in literature to measure empowerment. Donald et al. (2017) second this view based on their comprehensive account of studies employing household decision-making surveys. The authors confirm the (wide and increasing) use of surveys in literature and discuss their methodological shortcomings. Furthermore, surveys on female decision-making power are now part of the widely used Demographic and Health Surveys (DHS). This will likely encourage frequent use of decision-making surveys in the future (Donald et al., 2017).

1.2.2. Discordance of spousal statements

Despite the popularity of intra-household decision-making questionnaires, studies employing data from both partners challenge the notion that survey data from only one household member is sufficient to describe power dynamics within the household. Studies that do consider both males' and females' perspectives report substantial and systematic differences in power assessments made by men and women respectively (Allendorf, 2007; Ambler et al., 2017; Granbois and Willett, 1970; Story and Burgard, 2012; Twyman, Useche, and Deere, 2015; Uddin, Habibullah, and Sabah, 2016). The subset of studies that focuses on discordance observes variation in discordance between different groups and decision domains (Becker, Fonseca-Becker, and Schenck-Yglesias, 2006; Ghuman, Lee, and Smith, 2006; Jejeebhoy, 2002; Quarm, 2018; Lupri and Brinkerhoff, 1978; Twyman, Useche, and Deere, 2015). However, only a few previous studies link discordant statements to outcomes.

The study with the highest relevance to this one is that of Ambler et al. (2017). Ambler et al. (2017) link discordance to female well-being using data from Bangladesh.[7] Compared to couples in which both partners agree on the husband as the sole owner of the family's assets and primary decision maker, outcomes for women are better if the couple agrees on joint asset ownership and decision-making. Compared to the baseline scenario, female outcomes are also better in discordant couples in which the wife posits female asset ownership and decision-making power with her husband disagreeing.

A small number of previous studies has documented the relationship between discordant reports and health outcomes (Allendorf, 2007; Becker, Fonseca-Becker, and Schenck-Yglesias, 2006; Gasca and Becker, 2017; Jejeebhoy, 2002). Becker, Fonseca-

7. Female well-being is proxied by various measures such as working hours, BMI and use of birth control.

Becker, and Schenck-Yglesias (2006) document female under-reporting of their own decision-making power relative to their husbands as well as a (weak) association of female decision-making power and health behavior in Western Guatemala. Becker, Fonseca-Becker, and Schenck-Yglesias (2006) confirm a positive relationship between female education and concordant statements on joint decision-making. The authors find variation as to whether the male or the female's opinion on female autonomy was the better predictor of preventive health-related behaviors, such as the use of the contraceptive pill. They conclude that studies should elicit both husbands' and wives' perspectives on decision-making power. In Uttar Pradesh, a region characterized by low gender-equity, Jejeebhoy (2002) find men's perspective on female autonomy to be more predictive of outcomes than women's. In contrast, in Tamil Nadu, a region with higher reported female autonomy, women's perspectives are more indicative of contraception-related health outcomes.[8] Allendorf (2007) finds high rates of discordant responses in Nepal. The author finds higher rates of health care utilization in couples that give a concordant report on female decision-making power, compared to couples in which only one partner perceives female decision-making power, while the other does not.

1.2.3. Moderators of discordance

Discordance might arise from multiple causes, a selection of which will be discussed in the following. Moffitt et al. (1997) argue that measurement error will induce statement discordance that might be falsely interpreted as a reflection of disagreement. The authors posit that aggregating responses across questions can reduce this error. Specifically, Safilios-Rothschild (1970) warns that broad and unspecific decision domain might lead to gender differences in understanding of the matter. Ghuman, Lee, and Smith (2006, p. 3) show for five Asian countries that "cognitive and/or semantic meanings" of questions vary between demographic and cultural contexts and between female and male respondents, thereby limiting the generalizability of such comparisons.[9] Anderson, Reynolds, and Gugerty (2017) find higher rates of concordance in couples in Tanzania with higher educated women.

Discordance can also be caused by respondent's intention to paint a socially desirable picture (Allendorf, 2007; Jejeebhoy, 2002). For example, Jejeebhoy (2002) finds that

8. The authors consider the following outcomes: contraceptive use, unfulfilled need for contraception, recent pregnancy, and spousal conversation about contraception.

9. The five countries are India, Pakistan, Philippines, Malaysia, and Thailand.

husbands attribute more power to their wives than wives attribute to themselves in rural India. Accordingly, discordance is particularly high in regions with greater gender inequality. However, if questioned in depth during focus group interviews, men tend to correct previous statements and provide socially less desirable answers. Norms might also provide a focal point to the potentially ambiguous question of power. Allendorf (2007) finds that concordance is higher whenever there are clear gender norms on responsibility.

Surveys on interpersonal violence (IPV) face similar methodological challenges. Moffitt et al. (1997) suggest that social desirability, salience, and self-justification might moderate over- and under-reporting of IPV, among other things. Hayes and Boyd (2017) observe that the husband's presence during the wife's interview led to a reduced stated female acceptance of IPV. Conversely, the presence of another female induced higher stated acceptance of IPV.

1.2.4. Cultural context: female power in Indonesia

Discordance needs to be understood in its cultural context. The following section provides a brief overview of this study's context. The country is characterized by economic dynamism, the predominance of Islam as religion and high ethnic diversity with more than 300 ethnic groups (Blackburn, 2004b). The country is the world's fourth most populous country, tenth largest economy by purchasing power and the largest economy in Southeast Asia (The World Bank Group, 2018). Society is characterized by a high degree of cultural, religious and ethnic heterogeneity. However, as Blackburn (2004b) points out, the fall of the Suharto regime, which emphasized strong role division between men and women, heralded a new era, symbolized by the election of Indonesia's first female president in 2001 (This is not to say, that the cultural echoes of the Suharto period do not prevail in many households and through norms today). Under reference to the United Nations Development Program, Schaner and Das (2016) describe Indonesia as more gender equal than Pakistan and India but less equal than China. The female to male wage ratio increased from 57 percent in 1990 to 84 percent in 2011.[10] Still, male and female labor force participation rates differ significantly and, compared to men, women tend to work more often as unpaid family workers (Schaner and Das, 2016). Schaner and Das (2016) find that younger female cohorts are more likely to enter the formal sector of employment earlier on in their lives, compared to older female cohorts who often started their careers in informal employment. Total

10. Defined as the relative median hourly wage of women compared to men.

female labor force participation has been at a level of around 60 percent in the most recent past.

Particularly relevant to this study is the work by Frankenberg and Thomas (2001) and their account of measuring power in the context of Indonesian households. Frankenberg and Thomas (2001) acknowledge gifts to family members as one decision domain with particularly high rates of discordance. They find that Indonesian household decisions are subject to the country's diverse norms that vary by region and ethnicity. The authors find that the main decision maker strongly varies between decision domains. They suggest food and routine purchase expenditures as *female domains*, while larger expenditures appear to be a *male domain*.[11] Focus group interviews of Frankenberg and Thomas (2001) reveal that group dynamics influence response behavior. Summarizing, Indonesia's changing and diverse society make the country a relevant and interesting subject for a study of gender relationships.

1.3. Data

We employ data from the Indonesian Family Life Survey (IFLS), a widely used longitudinal household survey dataset. The first and last of five rounds of surveys have been conducted in 1993 and 2014 respectively. We are using the third (Strauss, Witoelar, and Sikoki, 2016), fourth (Strauss et al., 2009), and fifth wave (Strauss et al., 2004). The surveys are administered by the RAND corporation which cooperates with partners and scholars to ensure a continuing high standard of conceptualization and execution. The survey's initial design was set to be representative for 83 percent of the Indonesian population, thereby aiming to cover 7,000 households in 13 provinces (Thomas, Frankenberg, and Smith, 2001). Statistical weights are provided to adjust for changes in population composition. Interviewers did follow up with the original set of households as well as with their split-offs.[12]

The survey is based on multiple books. Each book consists of wide-ranging questions, from education to labor market participation. Some books are administered

11. Frankenberg and Thomas (2001) base their analysis on an earlier wave (IFLS 2) of the same panel and document slightly lower rates of discordance at about 25 percent.

12. This has led to an expansion in the number of households over time. By using sophisticated follow-up designs and tracking, the survey team has been able to achieve very low rates of attrition. For the second survey wave in 1997, the authors were able to reinterview 94 percent of households (Thomas, Frankenberg, and Smith, 2001). Further information on the survey design can be found in Frankenberg and Thomas (2001) and Thomas, Frankenberg, and Smith (2001).

to a subset of household members only.[13] At the core of this study is the household decision-making module in book 3A and the fertility module of book 4. Book 3A is administered to individuals who are currently married or cohabitate and whose spouses live in the same household or lived in the same household in the past six months.[14] Book 4 is administered to ever-married women who are between 15 and 49 years old. The book focuses on marital history, children and fertility. For each book, the interviewer also collects information on who responded and whether other people were present during the interview.

We only use heterosexual couples and refer to husbands, men, male spouses and wife, women, female spouses interchangeably. We omit all individuals who are not married or cohabitate, with missing age data or missing spouse data. We exclude all individuals who are neither the head of household or spouse of the head of household. This omits grandparents or married children still living in their parents' home. We do not include all individuals who neither live with their spouses nor lived with them in the six months prior to the survey. We exclude all individuals with a missing personal identifier and duplicate observations. We also only consider couples that remain *complete* following the application of these restrictions. The application of all restrictions leads to a significant reduction in sample size, the largest share of which is attributable to the exclusion of non-household heads and incomplete couples. For 2000, 2007 and 2014 a total of 21,736, 22,193 and 24,892 observations are available. Application of exclusion restrictions reduces the sample to the number of observations listed in table 1.1.

Table 1.1.: Number of observations after application of exclusion restrictions

	Number of Observations			
	2000	2007	2014	Total
Men	6,532	8,070	8,662	23,264
Women	6,532	8,070	8,662	23,264
Observations	13,064	16,140	17,324	46,528

Notes: For the majority of later analyses, couples are treated as one observation. The number of observations increases over time as the survey tracks and includes spin-off households.

13. Our main analysis is based on the household roster book K, book 3A and book 4. The household roster book K is only administered to the household head or a household member that is knowledgeable about the questions. The household roster book contains questions on all household members' income, their age and their relationship to the household head, inter alia. Book 3A is only administered to respondents who are at least of age 15.

14. See book 3A of IFLS 5 wave, questions PK00a and PK00b.

After application of exclusion restrictions, the sample carries the attributes presented in table 1.2. In most of our analyses, we treat one couple as one observation. We assign the male attributes as additional variables to women. We mark the husbands' values with the prefix "Spouse (Sp.)". Compared to women, men are older, are slightly better educated, earn a larger share of the household income and report almost twice as many worked hours. Very few men are unpaid family workers while 95 percent worked in the past twelve months.[15] One quarter of all women have unpaid family worker status while only 61 percent of them worked in the past twelve months. Household attributes are by definition the same for both sexes since we consider complete couples only.

15. The question asked is "Did [...] work in the last 12 months? ($>$ 5 years)".

Table 1.2.: Descriptive statistics by sex

	Women				Men			
	Mean	SD	Min	Max	Mean	SD	Min	Max
Individual attributes:								
Age	39.16	11.63	15	82	43.59	12.48	18	92
Elementary education	0.36	0.48	0	1	0.35	0.48	0	1
Any secondary education	0.48	0.50	0	1	0.49	0.50	0	1
Any college education	0.12	0.32	0	1	0.13	0.33	0	1
Household attributes:								
N HH adults	2.39	0.82	2	9	2.39	0.82	2	9
N HH children	1.69	1.15	0	9	1.69	1.15	0	9
Migration indicator	0.21	0.41	0	1	0.21	0.41	0	1
Household economy:								
Log HH income	16.68	1.10	9.6	21	16.68	1.10	9.6	21
Any land	0.32	0.47	0	1	0.32	0.47	0	1
Income share	0.19	0.27	0	1	0.81	0.27	0	1
Labor supply:								
Worked in past 12 months	0.61	0.49	0	1	0.95	0.21	0	1
Hours worked per annum	1,132.41	1,401.26	0	11,844	1,973.82	1,331.83	0	14,976
Unpaid family worker status	0.25	0.43	0	1	0.01	0.10	0	1
Observations	8,662				8,662			

Notes: **Data:** IFLS-5 wave (2014), cross-sectional data, one observation is one couple; **Variable definitions:** N HH children: number of children in household; N HH adults: number of adults in household; Migration indicator: household moved since last wave (yes = 1/no = 0); Log HH income: Log of annual income in IDR; Hours worked per annum are stated as reported (later results are robust to exclusion of individuals reporting over 16 hours * 365 days per year; less than .8 percent do report higher values.).

1.4. Typology and prevalence of discordance

1.4.1. Typology of statement combinations

We structure concordance and discordance along a typology of statement combinations that has been developed by Ambler et al. (2017). The focus is on whether the wife has decision-making power or not and on whether the couple agrees with that.[16]

The typology reflects all possible statement combinations that can arise from the response options available to each spouse. The lead question for every domain is "In your household, who makes decisions about: [domain]". The question is asked for 13 domains, eg savings or routine purchases. Appendix table A.1 presents the list of decision domains. In response, individuals can circle letters representing single household members, such as A for the respondent, B for the spouse, and so forth. Thus, each spouse can either respond that he/she makes the decision by themselves, by their spouse or that they engage in joint decision-making. Additional household members or individuals not living in the household that can be named as decision makers are identified as out of scope. If spouses do not that either of the spouses makes the decision, the couple will be assigned to a residual category as described in the following.

The typology of possible statement combinations is exhaustive: We sort all possible combinations in one (and only one) category of the typology. We omit couples in which either spouse's statement is missing. All others are included in this typology. Table 1.3 provides a description of each category. The first four categories (CM, CF, CB and CN) describe concordance whereas the last four (DFM, DMF, DOBOF and DONOM) describe discordance with respect to whether the wife has decision-making power. Couples of category CM give a concordant response of the husband as the sole decision maker. In couples of type CF, both spouses perceive the wife as the sole decision maker. In couples of type CB, both spouses perceive joint decision-making. Finally, couples of category CN give a concordant response that neither of the two spouses makes the decision.

The discordance categories are divided into strong and weak discordance and a residual category. The first two discordance types (DFM, DMF) describe cases in which spouses fundamentally disagree about the wife's role, two cases which we coin *strong*

16. We use this *wife-focused* typology since most related studies focus on the wife's decision-making power.

discordance. One partner perceives female decision-making power while the other does not. The third case DOBOF (weak discordance) depicts the case in which both partners fundamentally perceive female decision-making power but do not agree on whether the husband also has a say. The last, *residual* category DONOM comprises all cases in which either spouse says that the decision is made by a third person while the other spouse perceives the husband as the sole decision maker.

The typology applies to all domains but the contraceptive use decision domain. When asked for decisions on "whether you and your spouse use contraception", respondents were offered an additional "never consider the use of contraception" response option. We will refer to this option as *no use* in the following. We exclude all spousal statement combinations in which either spouse replies *no use*. We do so because the *no use* response can describe an outcome and decision at the same time. This would challenge interpretation, which is why we decide not to consider these couples. We also exclude a small group of individuals who respond that someone else in the household or someone not living in the household makes the decision (coined as "none"). Both exclusion restrictions reduce our sample size by around 34 percent for analyses in the contraception use domain. The full contraception typology including the *no use* response can be found in appendix table A.2.

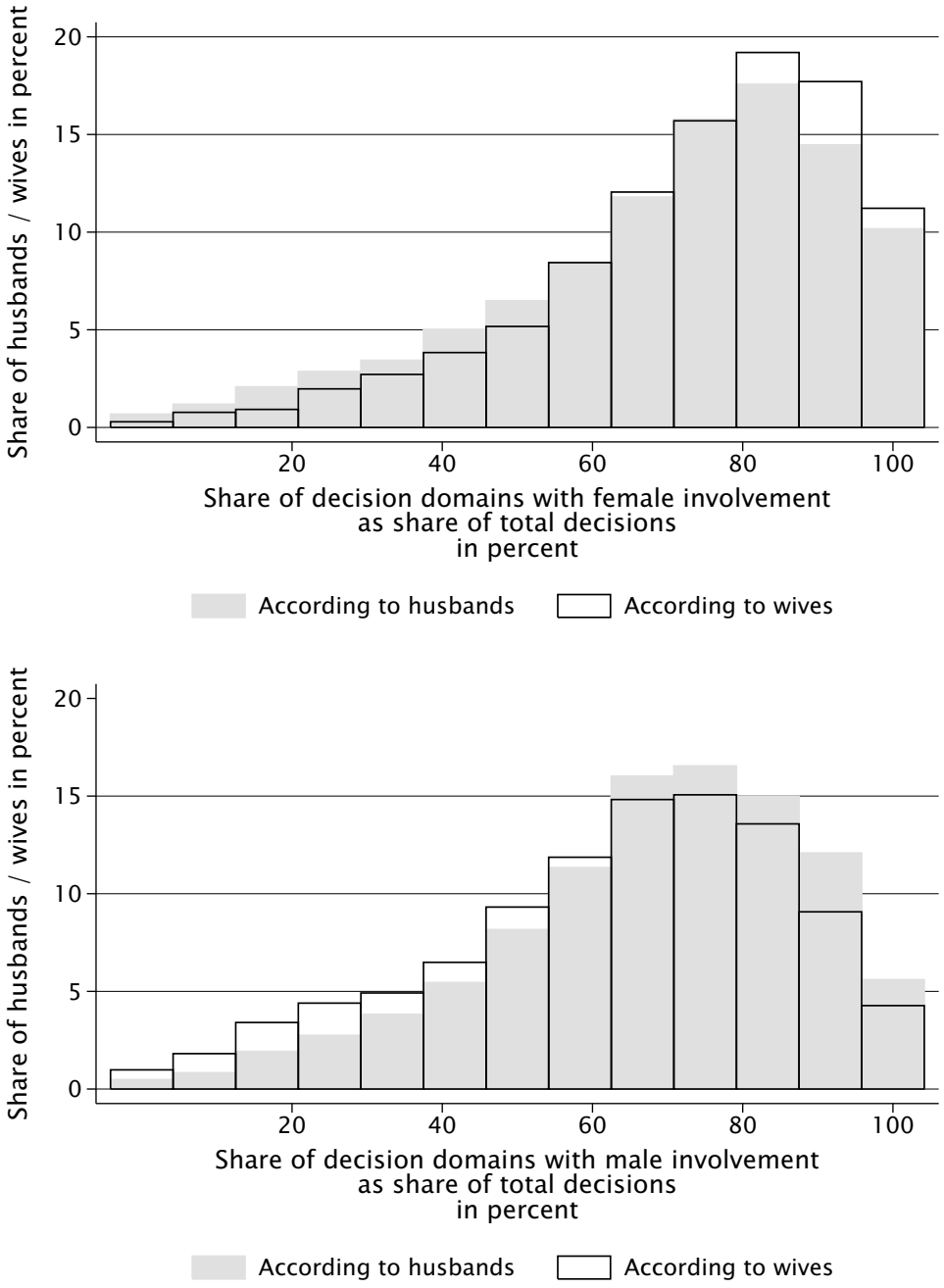
Table 1.3.: Typology of responses of all decision domains except decisions on contraceptive use

Code	**Combination of husband's and wife's statement**	
	Concordance	
CM	Wife and Husb: (m)	Both partners perceive husband as the sole decision maker
CF	Wife and Husb: (f)	Both partners perceive wife as the sole decision maker
CB	Wife and Husb: (mf)	Both partners perceive joint decision-making
CN	Wife and Husb: (none)	Both partners perceive that neither of them makes decision
	Discordance	
DFM	Wife: (f) $\vee$ (mf); Husb: (m) $\vee$ (none)	Strong discordance: Wife perceives female decision-making power (individually or jointly), husband does not
DMF	Husb: (f) $\vee$ (mf); Wife: (m) $\vee$ (none)	Strong discordance: Husband perceives female decision-making power (individually or jointly), wife does not
DOBOF	One Spouse: (mf); Other Spouse: (f)	Weak discordance: Both spouses perceive female decision-making power, one of them perceives wife as sole decision maker, the other joint decision-making
DONOM	One Spouse: (none); Other Spouse: (m)	One of the spouses perceives that neither spouse makes the decision, the other spouse perceives the husband as the sole decision maker

1.4.2. Prevalence

We first analzye the general difference between male and female perception of male and female decision-making power. In figure 1.1, we show the overall decision-making power share of each partner from each perspective. This overall share is coded as 1 (= 100 percent) if an individual is perceived to have a say (individually or jointly) in all 13 decision domains. It is coded as 0 if an individual is perceived to not have a say in any domain. The upper histogram displays the distribution of the male and the female perception of female decision-making power. The lower histogram displays the distribution of the male and the female perception of male decision-making power. Women report a higher average female decision-making share value than their husbands. The same holds vice versa. Women perceive less male decision-making power than husbands perceive themselves. Appendix section A.6.4 offers an overview of the relationship between spousal presence and stated decision power.

Figure 1.1.: Histograms of perceived female (upper graph) and male (lower graph) decision-making power



Note: **Data:** IFLS-5 wave (2014), cross-sectional data, one observation is one couple; Graph: two histograms on male and female perception of male and female decision-making power. Grey area: husband's perception. Black lined bars: wife's perception. **Scale:** upper (lower) graph: decision-making power is captured as number of household decision domains the wife (husband) is involved in over total number of household decision domains; Value 1 on x-axis indicates that wife (husband) has a say in all household decisions.

We present the prevalence of discordance per domain in tables 1.4 and 1.5 (A full list of the actual decision-making domains including the wording can be found in the appendix table A.1). Between 49 and 67 percent of couples agree on who makes the decision in any decision domain (see line *Sum concordance* in the tables). The highest concordance exists with respect to routine purchases (67 percent). The lowest concordance rates are found in the domains of "Time the wife spends socializing" and "Money for monthly savings". We find that across twelve out of 13 domains, the case (DFM) occurs more or at least as often as the opposite case (DMF). This implies that it happens more often that the wife perceives female decision-making power while the husband does not than that the opposite case occurs (the opposite case is that the husband perceives female decision-making power, while the wife does not). This is insofar reasonable as women perceive higher overall female decision-making power than men do. In general, the case of weak discordance (DOBOF) occurs more often than either of the strong discordance cases (DFM, DMF) alone. Decisions on contraceptive use are excluded from the following tables, as they offer additional potential responses. Detailed statistics on contraceptive use responses can be found in appendix table A.21. We find that 50 percent of couples give concordant responses. 13 percent show strong discordance and 23 percent weak discordance. 14 percent give discordant responses that include one partner reporting no use.

Table 1.4.: Concordance and discordance by decision domain (1/2)

	Routine purchases	Your children's education	Your children's health	Large expensive purchases	Giving money to wife's family	Giving money to husband's family
	–	–	–	–	–	–
	in percent of column	in percent of column	in percent of column	in percent of column	in percent of column	in percent of column
Concordance:						
CM: Wife and Husb: (m)	0.02	0.04	0.02	0.07	0.04	0.06
CF: Wife and Husb: (f)	0.60	0.08	0.08	0.06	0.06	0.04
CB: Wife and Husb: (mf)	0.04	0.37	0.41	0.41	0.41	0.40
CN: Wife and Husb: (none)	0.01	0.06	0.05	0.01	0.02	0.02
Sum concordance	0.66	0.55	0.57	0.55	0.53	0.53
Discordance:						
DFM: Wife: (f) $\vee$ (mf); Husb: (m) $\vee$ (none)	0.08	0.15	0.14	0.13	0.14	0.17
DMF: Husb: (f) $\vee$ (mf); Wife: (m) $\vee$ (none)	0.08	0.10	0.07	0.14	0.13	0.14
DOBOF: One spouse: (f); Other spouse: (mf)	0.17	0.19	0.22	0.17	0.19	0.14
DONOM: One spouse: (none); Other spouse: (m)	0.00	0.01	0.01	0.01	0.01	0.02
Sum discordance	0.34	0.45	0.43	0.45	0.47	0.47
Observations	8,662	8,662	8,662	8,662	8,662	8,662

Notes: **Data:** IFLS-5 wave (2014), cross-sectional data, one observation is one couple; **Typology:** CM: concordant report of husband as sole decision maker, CF: concordant report of wife as sole decision maker, CB: concordant report of joint decision making; CN: concordant report of neither partner as decision maker, DFM: wife perceives female decision making power, husband does not, DMF: husband perceives female decision making power, wife does not, DOBOF: both spouses perceive female decision making power, one of them perceives a sole female decision maker, the other joint decision making, DONOM: one of the spouses perceives that neither spouse makes the decision, the other spouse perceives male decision maker.

Table 1.5.: Concordance and discordance by decision domain (2/2)

	Gifts for parties / weddings	Money for monthly arisan (savings lottery)	Money for monthly savings	Time the husband spends socializing	Time the wife spends socializing	Whether you/your spouse works
	–	–	–	–	–	–
	in percent of column	in percent of column	in percent of column	in percent of column	in percent of column	in percent of column
Concordance:						
CM: Wife and Husb: (m)	0.03	0.02	0.03	0.38	0.03	0.21
CF: Wife and Husb: (f)	0.09	0.18	0.10	0.02	0.26	0.01
CB: Wife and Husb: (mf)	0.42	0.13	0.13	0.15	0.20	0.37
CN: Wife and Husb: (none)	0.00	0.21	0.22	0.00	0.00	0.00
Sum concordance	0.54	0.53	0.48	0.54	0.49	0.59
Discordance:						
DFM: Wife: (f) $\vee$ (mf); Husb: (m) $\vee$ (none)	0.12	0.15	0.17	0.24	0.13	0.23
DMF: Husb: (f) $\vee$ (mf); Wife: (m) $\vee$ (none)	0.08	0.13	0.17	0.17	0.09	0.12
DOBOF: One spouse: (f); Other spouse: (mf)	0.24	0.16	0.13	0.05	0.29	0.06
DONOM: One spouse: (none); Other spouse: (m)	0.00	0.03	0.05	0.00	0.00	0.00
Sum discordance	0.46	0.47	0.52	0.46	0.51	0.41
Observations	8,662	8,662	8,662	8,662	8,662	8,662

Notes: **Data:** IFLS-5 wave (2014), cross-sectional data, one observation is one couple; **Typology:** CM: concordant report of husband as sole decision maker, CF: concordant report of wife as sole decision maker, CB: concordant report of joint decision making; CN: concordant report of neither partner as decision maker, DFM: wife perceives female decision making power, husband does not, DMF: husband perceives female decision making power, wife does not, DOBOF: both spouses perceive female decision making power, one of them perceives a sole female decision maker, the other joint decision making, DONOM: one of the spouses perceives that neither spouse makes the decision, the other spouse perceives male decision maker.

1.5. Discordance and couple attributes

In this section we study whether discordance is *random* or whether it is systematically related to couple attributes. We do so in two steps in sections 1.5.1 and 1.5.2. First, we study the prevalence of discordance types with respect to the female income share.[17] Secondly, we relate discordance types to overall couple attributes and further proxies for female power. The observed relationship suggests that concordance is indeed related to *actual* female power.

1.5.1. Prevalence of discordance and female income share

Approach. In figure 1.6, we relate the wife's income share to the prevalence of specific types of con- or discordance in the labor supply decision domain.[18] We create one graph for each of the six types of potential statement combinations. Each graph relates the prevalence of the specific response type as the share of all responses (vertical axis) to female income share (horizontal axis).[19] We use cross-sectional data, based on the IFLS-5 (2014). This is a merely correlative, descriptive way of assessing the relationship. Since the figures display cross-sectional data only, they do not allow to infer any causal statements.

Results. The upper two graphs and the center-left graph in figure 1.6 confirm a positive relationship between female income and decision-making power. The share of couples that agree on the husband as the sole decision maker (type CM) stands in a negative relationship with the female contribution to household income (upper left graph). The opposite holds true for the (small) share of couples in which both partners perceive the wife as the sole decision maker (CF, upper right graph). The center-left graph shows the share of couples in which both partners agree on joint decision-making (CB). Their share among all couples is positively associated with the female contribution to household labor income.

The remaining graphs (center right and bottom row) present prevalence rates of discordant couples. The center-right graph of figure 1.6 shows the association between the

17. Bertrand, Kamenica, and Pan (2015) shows for the United States, a considerable more progressive country than Indonesia with respect to gender roles, that the wife's income share carries an important normative meaning. The wife's contribution to household income not only identifies her own role in society but also her relationship to the husband.

18. The question asked in this domain is "In your household, who makes decisions about: whether you/your spouse works".

19. For a definition of female income share, please refer to section 1.7.1.

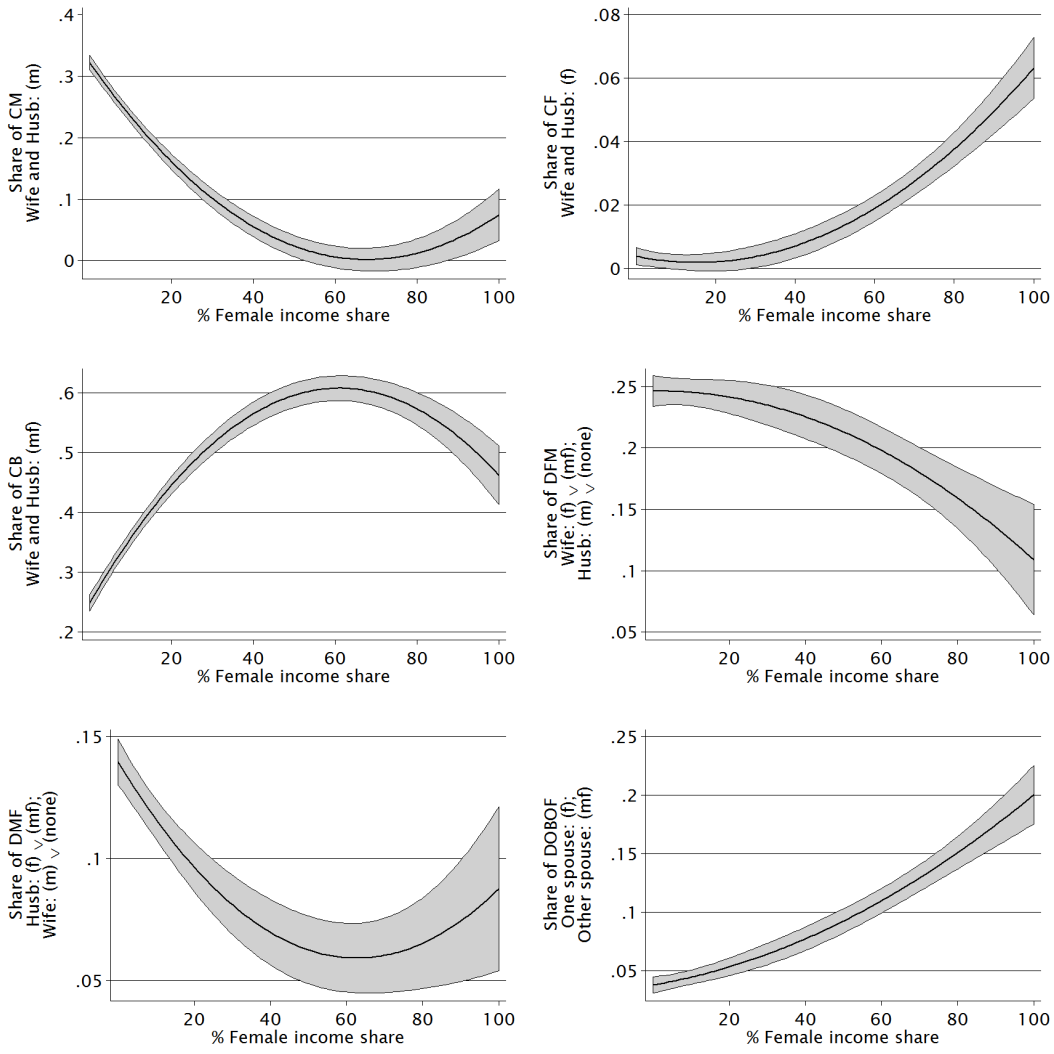
share of couples in which the wife perceives female decision-making power, while the husband does not (DFM) and the female income share. The bottom left graph shows the share of couples in which the wife does not perceive female decision-making power while the husband does (DMF) as a function of the female share in household labor income. They suggest that *strong discordance* is negatively associated with the female share of household income. The opposite holds true for *weak discordance*, displayed in the bottom right graph: The share of weakly discordant couples that fundamentally agree on a female role in decision-making but disagree on the husband's role (DOBOF) is positively associated with female income. One possible explanation is that a more pronounced female income share allows for less ambiguity that could, in turn, result in strong discordance.

1.5.2. Attributes of discordant couples

Approach. We divide the full sample into six subsamples based on the main six possible statement combinations in response to the question *"In your household, who makes decisions about: whether you/your spouse works"*. Couples, whose attributes are presented in columns one through three, have given a concordant statement combination. They agree on the husband as the sole decision maker, the wife as the sole decision maker or joint decision-making respectively. The remaining columns present the attributes of couples that have given a discordant statement combination. Columns four and five present attributes of couples in which either the wife perceives female decision-making power (individually or jointly) while the husband does not (column four) or vice versa (column five). In column six, we present the means for couples that agree on female decision-making power but do not agree on whether the husband also has decision-making power. We use cross-sectional data, based on the IFLS-5 (2014). Each observation is one couple. If a variable is not specified, it shows the wife's value. The husband's attributes are referred to by the prefix spouse (abbreviated by *Sp:*).

Results. The descriptive statistics in table 1.7 suggest that specific types of concordance and discordance are related to *actual* differences in female power. This conclusion is based on the between-subsample variation in variables that are associated with female power. The subsamples differ with respect to the wife's income share, female age, female education, and also the overall female decision-making share. Table 1.7 presents subsample averages. Appendix table A.7 provides T-Tests on the statistical significance of subsample differences.

Table 1.6.: Quadratic linear prediction of prevalence of statement combinations with female income share



Notes: **Data:** IFLS-5 wave (2014), cross-sectional data, one observation is one couple; **Decision making domain:** "Whether you/your spouse works"; Vertical axis: frequency of single statement combination; Horizontal axis: relative contribution of the wife to household labor income (husband and wife only).

One key finding is that female income shares are substantially lower in couples with strong discordance (columns 4 and 5) compared to weak discordance (column 6).[20] The *weakly discordant* (DOBOF) group's females are also older than those in the *strongly discordant* groups (DFM, DMF). Within the concordant groups, the average female income share is lowest in those couples in which both partners agree on the husband as the sole decision maker (column 1) and highest in couples that agree on the wife as the sole decision maker (column 2).

The overall decision share across all domains corresponds to the statement combinations in the labor supply domain. The wife's self-assessed decision share (Wife: own decision share) is higher for couples that respond with DFM and DOBOF than it is for the couples responding DMF. The same holds vice versa for men. The decision share attributed by men to women (Husb: wife's decision share) is higher in the cases DMF and DOBOF than for couples responding DFM. In appendix table A.7, we compare whether the aforementioned differences in income shares and overall female decision shares are statistically significant at the five percent level. We find this to be the case. We conclude that specific types of discordance do not appear to be a random artifact in our data but appear to be systematically associated with *real* female power. In a next step, we use cross-sectional regression to learn which single outcomes have the greatest association with specific types of discordance.

20. The latter category of weak discordance implies that both spouses essentially perceive female decision-making power, but diverge for the husband's role.

Table 1.7.: Comparison of couple attributes by con- and discordance types in labor market decision domain

	(1) CM Wife and Husb: (m)		(2) CF Wife and Husb: (f)		(3) CB Wife and Husb: (mf)		(4) DFM Wife: (f) ∨ (mf) Husb: (m) ∨ (none)		(5) DMF Husb: (f) ∨ (mf) Wife: (m) ∨ (none)		(6) DOBOF One Spouse: (f) Other Spouse: (mf)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Individual attributes:												
Age	35.94	11.01	44.77	11.69	40.32	11.19	38.58	11.60	39.46	12.16	43.61	12.02
Sp: Age	40.68	11.69	51.21	14.11	44.37	12.17	42.96	12.24	44.27	13.01	48.43	13.41
Elementary education	0.36	0.48	0.52	0.50	0.32	0.47	0.38	0.49	0.40	0.49	0.40	0.49
Any secondary education	0.56	0.50	0.38	0.49	0.44	0.50	0.49	0.50	0.47	0.50	0.42	0.49
Any college education	0.05	0.22	0.05	0.22	0.19	0.39	0.08	0.28	0.08	0.27	0.11	0.31
Sp: Elementary education	0.35	0.48	0.46	0.50	0.33	0.47	0.37	0.48	0.36	0.48	0.43	0.50
Sp: Any secondary education	0.55	0.50	0.46	0.50	0.47	0.50	0.50	0.50	0.48	0.50	0.44	0.50
Sp: Any college education	0.09	0.28	0.02	0.16	0.17	0.38	0.10	0.30	0.13	0.33	0.09	0.29
Household attributes:												
N HH adults	2.39	0.85	2.34	0.79	2.37	0.78	2.40	0.85	2.39	0.77	2.47	0.92
N HH children	1.74	1.09	1.50	1.18	1.66	1.17	1.70	1.14	1.75	1.13	1.57	1.22
Migration indicator	0.24	0.43	0.26	0.44	0.21	0.41	0.21	0.40	0.22	0.41	0.16	0.37
Household economy:												
Log HH income	16.59	1.00	16.44	1.12	16.83	1.15	16.65	1.05	16.56	1.09	16.56	1.26
Any land	0.26	0.44	0.17	0.38	0.35	0.48	0.32	0.46	0.34	0.47	0.33	0.47
Female income share	0.05	0.15	0.54	0.39	0.28	0.29	0.16	0.24	0.13	0.24	0.36	0.35
Log h worked pa	6.24	1.68	7.27	0.84	7.22	1.03	6.91	1.28	6.74	1.43	7.17	1.14
Sp: Log h worked pa	7.41	0.86	7.30	1.05	7.39	0.89	7.41	0.90	7.39	0.93	7.26	1.05
Decision share:												
Wife: Own decision share	0.61	0.21	0.81	0.17	0.82	0.16	0.79	0.17	0.61	0.22	0.76	0.18
Husb: Wife's decision share	0.58	0.23	0.76	0.19	0.81	0.17	0.59	0.22	0.78	0.19	0.78	0.18
Wife: Husband's decision share	0.62	0.21	0.30	0.26	0.68	0.21	0.62	0.24	0.62	0.22	0.45	0.26
Husb: Own decision share	0.66	0.19	0.30	0.23	0.72	0.19	0.65	0.21	0.69	0.22	0.58	0.26
Observations	1,849		82		3,186		1,990		998		542	

Notes: **Data:** IFLS-5 wave (2014), cross-sectional data, one observation is one couple; Individual level attributes are wife's attributes by default, Spouse (Sp:) refers to husband's attributes; **Decision making domain:** "Whether you/your spouse works"; **Variable definitions:** N HH children: number of children in household; N HH adults: number of adults in household; Migration indicator: household moved since last wave (yes = 1/no = 0); Log HH income: Log of annual income in IDR; Wife:/Husb.: provides wife's/husband's perspective on decision making power respectively; **Typology:** CM: concordant report of husband as sole decision maker, CF: concordant report of wife as sole decision maker, CB: concordant report of joint decision making; CN: concordant report of neither partner as decision maker, DFM: wife perceives female decision making power, husband does not, DMF: husband perceives female decision making power, wife does not, DOBOF: both spouses perceive female decision making power, one of them perceives a sole female decision maker, the other joint decision making, DONOM: one of the spouses perceives that neither spouse makes the decision, the other spouse perceives male decision maker.

1.6. Discordance and outcomes

Results in section 1.5 suggest that a shared perception of female power is associated with higher *actual* female decision-making power. Accordingly, we also hypothesize that labor and fertility outcomes correspond to this pattern. The endogenous relationship between power and outcomes prevents any form of causal interpretation. However, the strength of the relationship is of interest as the outcomes allow to proxy the degree of female power in a couple, assuming that labor supply and contraceptive use are indicative of female power. A strong relationship between couples' statement combinations and outcomes will inform our idea of female power in couples of a specific statement combination type. Further, outcome prediction is a goal on its own and any insight gained will be informative in this regard.

The first three subsections of this section relate specific outcomes, for example, labor supply, to a couples' perception of female power as expressed in their statement combinations. We group statement combinations along the previous typology to learn about outcome differences between the groups captured in the typology. In the fourth subsection, we slightly change our focus to study the merit of incorporating both spouses' perspectives. We do so by holding the female perception constant and varying the husband's one.

1.6.1. Female labor supply

Female labor supply is both a determinant and consequence of female decision-making power.[21] The following analysis is built on this assumption of a bi-directional relationship between perceived female decision-making power and labor outcomes. Based on previous evidence, we hypothesize that labor supply is higher where female power is higher. Variation between different typology groups will inform our idea of female power within these groups.

Approach. We estimate a standard OLS model. We use cross-sectional data, based on the IFLS-5 (2014). We estimate three labor outcomes in three distinct models. All outcomes are captured in the vector $LABOROUTCOMES$. The three outcomes are employment status, hours worked and status as an unpaid family worker. Employment status is an indicator variable, coded as 1 if a woman has worked in the past twelve

21. Goldin (2006, p. 1) argue that the nature of female decision-making "horizons" are linked to expectations about future female labor participation. Accordingly, Basu (2006) document the bi-directional relationship between decision-making and economic power.

months and zero if not (This includes any work in- and outside of the household, except work classified as unpaid family work). Working hours per annum are coded as the product of hours worked per week and weeks worked per year. Results are only estimated for individuals reporting less than 5.840 working hours per year, a condition which applies to 99.25 percent of the sample.[22] Unpaid family worker status is coded as an indicator variable. It takes on the value one if an individual is considered an unpaid family worker and zero if not.[23]

$$LABOROUTCOMES_i = \beta_0 + \beta_1 SCL_h + \beta_2 INDIVIDUAL_i + \beta_3 HOUSEHOLD_h + \rho_1 PROV_i + \varepsilon_i \quad (1.1)$$

The vector SCL_h contains indicator variables for all possible types of spousal statement combinations to the question: "In your household, who makes decisions about: whether you/your spouse works". We omit the indicator variable for the group which gives a concordant report on the husband as the sole decision maker (CM). Thus, the vector consists of 7 indicator variables for the seven statement combination by the couple as listed in table 1.3, except for case CM.

The level of observation is that of a household. The vector $INDIVIDUAL_i$ contains attributes of the household's individuals while $HOUSEHOLD_h$ contains household attributes. The included individual attributes comprise the individual's educational level[24] and age for both husband and wife. Male values are indicated by "Sp:" for the spouse. The household variables of vector $HOUSEHOLD_i$ capture the number of members of the household (number of all adult members and number of all children), the log of household income and a migration variable indicating whether the household moved since the last time it was surveyed. $PROV_i$ is a vector of province dummies and ε_i is an idiosyncratic error term. Robust standard errors in parentheses are clustered at level of the region.

22. Respondents can provide information on two main jobs in Book 3A of the survey. We use combined statements for the first job on hours (tk22a) and weeks (tk23a) as well as for the second job (tk23b, tk22b respectively) and calculate the sum.

23. There are two opportunities in the survey at which an individual can be classified as an unpaid family worker. We only consider someone to be an unpaid family worker if he or she are classified as such in the family roster (Within survey book K of IFLS wave 5, this is asked as question AR15b: "What were the total earnings of [...] in the last 12 months?" Response options are the salary, "Unpaid Family Worker" and "Don't Know").

24. Each individual is assigned based on the highest educational level that they attended. We cluster all levels into the four groups of a) no education, b) elementary education c) secondary education and 3) tertiary education. Please refer to appendix table A.3 for the distribution and a further specification of the grouping.

Results. Table 1.8 presents the estimates for the relationship between spousal (dis-) agreement and female labor force participation in a cross-sectional OLS estimate using 2014 data. The likelihood of female employment is highest in couples in which both partners perceive female decision-making power, whether con- or discordant regarding the men's role (see coefficients CF, CB, DOBOF in column one).[25] Female labor force participation is lowest in couples in which the husband makes decisions alone (omitted, baseline result) or in which the partners are strongly discordant (see coefficients DFM, DMF). This pattern also holds true for annual worked hours (column 2). Compared to the baseline group, the likelihood of the wife working as an unpaid family worker is lower in all other groups. The point estimates suggest that female unpaid family workers are particularly unlikely to be found in families with both partners agreeing on the wife as the sole decision maker. Point estimates further suggest a (statistically insignificant) difference between the strong and weak discordance categories (DFM, DMF and DOBOF respectively).

It is important to acknowledge that female decision-making power cannot be considered to be an explanatory variable. Looking at the different degrees of correlation with various form of discordant statements is still informative. Our preliminary conclusion is that female decision-making power is higher in couples in which both partners perceive female decision-making power, relative to cases of strong discordance about the female role and relative to concordant reports on the husband as the sole decision-maker.

1.6.2. Use of contraception and covert contraception methods

Previous evidence has shown that women and men tend to differ in their desired number of offspring. In societies with strong traditional norms, husbands tend to report higher overall fertility preferences than their wives (Rasul, 2008b; Mbaye and Wagner, 2017; Westoff, 2010). Accordingly, we hypothesize that contraceptive use is an expression of female power. Furthermore, women choose different strategies to realize their preferences. One of them is covert contraceptive use (Gasca and Becker, 2017). The first model of this section estimates the relationship between couples' statement combinations and contraceptive use. In its second model, we estimate the relationship between couples' statement combinations and covert contraceptive use.

25. At this point we do not focus on those who reply "none" (One of the spouses perceives that neither spouse makes the decision, the other spouse perceives male decision maker) as we consider this to be an exception of little overall relevance. We observe 43 cases in total in all three surveys (with 23,111 total observations).

Table 1.8.: Types of concordance and discordance and their association with labor outcomes

	Dependent variables:		
	(1) Wife working (Yes = 1/ No = 0)	(2) Hours worked (in h worked pa)	(3) Unpaid family worker (Yes = 1/ No = 0)
Concordance:			
CF: Wife and Husb: (f)	0.519***	965.459***	-0.270***
	(0.026)	(151.455)	(0.043)
CB: Wife and Husb: (mf)	0.483***	1046.775***	-0.089***
	(0.021)	(51.719)	(0.027)
CN: Wife and Husb: (none)	0.603***	1513.335***	0.528***
	(0.033)	(85.707)	(0.037)
Discordance:			
DFM: Wife: (f) ∨ (mf); Husb: (m) ∨ (none)	0.278***	540.927***	-0.093***
	(0.014)	(50.530)	(0.021)
DMF: Husb: (f) ∨ (mf); Wife: (m) ∨ (none)	0.199***	375.575***	-0.039*
	(0.025)	(59.270)	(0.023)
DOBOF: One spouse: (f); Other spouse: (mf)	0.448***	1008.506***	-0.184***
	(0.024)	(94.345)	(0.027)
DONOM: One spouse: (none); Other spouse: (m)	0.014	371.200	-0.307***
	(0.109)	(335.884)	(0.104)
Household attributes:			
Log HH income	0.041***	221.281***	-0.053***
	(0.008)	(25.779)	(0.014)
Constant	-0.332**	-3.4e+03***	1.408***
	(0.119)	(392.423)	(0.191)
Control for (spousal) age and education	✓	✓	✓
Control for N HH adults, N HH children, migration	✓	✓	✓
Region dummies	✓	✓	✓
Observations	7,562	7,562	4,802
R^2	0.22	0.16	0.14
Clusters	19	19	19
Robust standard errors clustered at regional level	Yes	Yes	Yes

Notes: Standard OLS estimate; **Dependent variable:** employment status of wife (binary), reported hours worked per year (weekly hours multiplied with weeks worked) and status as unpaid family worker (binary); **Baseline group:** concordant couples, perceiving husband as the decision maker; Individual level attributes are wife's attributes by default, Spouse (Sp:) refers to husband's attributes; **Data:** IFLS-5 wave (2014), cross-sectional data, one observation is one couple; **Decision making domain:** "Whether you/your spouse works"; **Typology:** CM: concordant report of husband as sole decision maker, CF: concordant report of wife as sole decision maker, CB: concordant report of joint decision making; CN: concordant report of neither partner as decision maker, DFM: wife perceives female decision making power, husband does not, DMF: husband perceives female decision making power, wife does not, DOBOF: both spouses perceive female decision making power, one of them perceives a sole female decision maker, the other joint decision making, DONOM: one of the spouses perceives that neither spouse makes the decision, the other spouse perceives male decision maker; Robust standard errors in parentheses are clustered at regional level; ***/**/* indicate significance at the 1%/5%/10% level.

This is built on the hypothesis that women of lower power need to resort to covert methods. We hypothesize that higher female power will manifest itself in higher rates of contraceptive use and lower rates of covert method use.

Use of contraception

Approach. We estimate the relationship between perceived female decision-making power and contraceptive use. We estimate a standard OLS model. We use cross-sectional data, based on the IFLS-5 (2014). The main dependent variable of interest is $ContraceptiveUse_i$, an indicator variable, which takes on the value 1 if the wife reports using contraception, and zero if she reports not using contraception. Contraceptive use is reported by women in a dedicated interview book. Appendix table A.40 shows that rates of reported contraceptive use are not different at a five percent level of significance for women whose partners were present during that dedicated interview compared to those with absent partners.

$$ContraceptiveUse_i = \beta_0 + \beta_1 SCC_h + \beta_2 INDIVIDUAL_i + \beta_3 HOUSEHOLD_h + \rho_1 PROV_i + \varepsilon_i \quad (1.2)$$

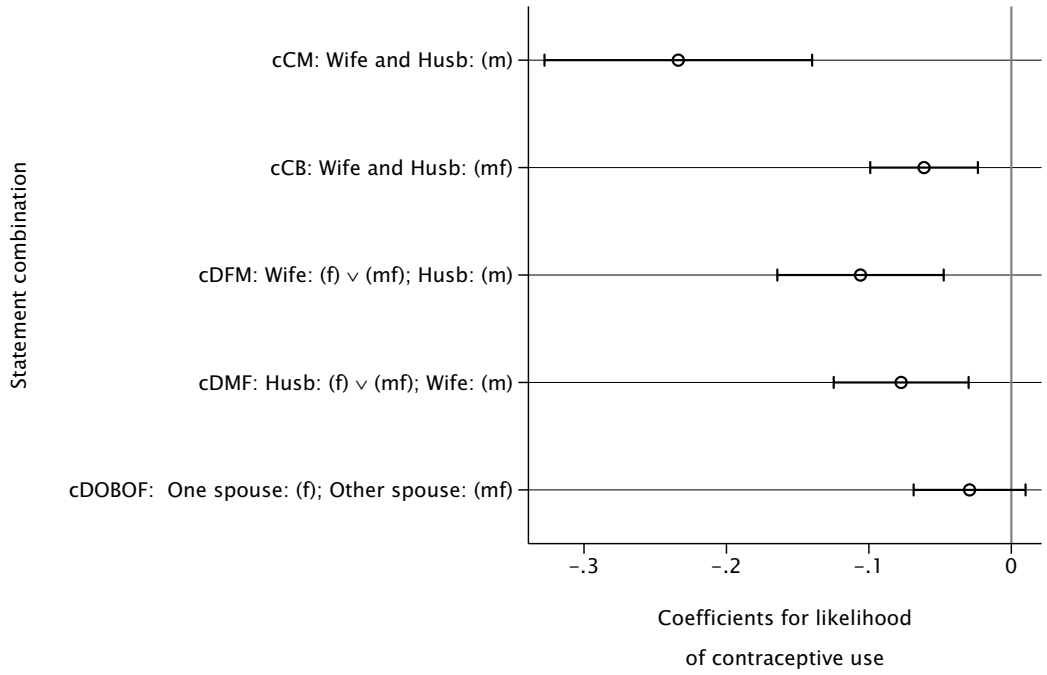
We introduce a vector SCC_h that contains indicator variables for all possible statement combinations to the question on "In your household, who makes decisions about: whether you and your spouse use contraception". The vector consists of one indicator variable for each possible statement combination by the couple as listed in table A.2, except for the statement combination cCF. Diverging from our previous labor supply model, we omit the indicator variable for the group which gives a concordant report on the wife as the sole decision maker (cCF). We use this reference group because our previous reference group (sole male decision maker) is a rare exception in response to this question, with only a small share of couples (N = 119) giving this response. The vectors $INDIVIDUAL_i$, $HOUSEHOLD_h$, $PROV_i$ and ε_i are modeled the same way as in the previous model. Robust standard errors in parentheses are clustered at level of the region.

Results. Results are presented in figure 1.2. The regression table can be found in appendix table A.8, column two.[26] We estimate the highest likelihood of contraceptive use for our baseline category: couples that agree on a sole female decision maker. The

26. In table A.8, we run five specifications. The first specification (column 1) runs the baseline model without controls, the second column presents the baseline model. The other three specifications exclude subgroups of women. Please refer to the robustness section 1.8 for further comments.

lowest likelihood of contraceptive use is estimated for couples that agree on the husband as the sole decision maker (cCM). It is significantly higher for couples that agree on joint decision-making (cCB). Compared to the baseline category, contraceptive use is expected to be lower for the two strongly discordant groups (cDFM, cMDF). The estimated rate of contraceptive use in weakly discordant couples (cDOBOF) is not statistically different from the baseline group. We conclude that outcomes differ based on whether or not couples show discordant opinions. The second analysis in section 1.6.3 permits to learn more about the value of taking both spouses' perspectives into account.

Figure 1.2.: Contraceptive use across statement combinations: coefficient estimates



Note: Standard OLS estimate; **Dependent variable:** contraceptive use (Yes = 1/No = 0); **Baseline group:** concordant couples, perceiving wife as the sole decision maker; **Data:** IFLS-5 wave (2014), cross-sectional data, one observation is one couple; **Decision making domain:** "whether you and your spouse use contraception"; **Typology:** cCM: concordant report of husband as sole decision maker, cCB: concordant report of joint decision making, cDFM: wife perceives female decision making power, husband does not, cDMF: husband perceives female decision making power, wife does not, cDOBOF: both spouses perceive female decision making power, one of them perceives a sole female decision maker, the other joint decision making; Robust standard errors are clustered at regional level; Ticks indicate 95 percent confidence interval.

Use of covert contraception methods

Approach. We estimate the relationship between perceived female decision-making power and covert contraceptive use conditional on contraceptive use. For this, we distinguish between non-covert and covert contraception methods. The main underlying assumption is that women of higher power do not need to conceal their contraceptive use by using covert methods. We define covert methods as those methods that tend to allow women to use contraceptives more discretely.[27] One example of a method that is very difficult to conceal is condom use. Contrastingly, injections are easier to use covertly. Mere use does not imply that they use them covertly.[28] Appendix table A.20 presents a list of all methods and their recorded prevalence. The appendix section A.4 also presents the classification into covert and non-covert. We rely mostly on other literature to identify each of the reported methods as either covert or not covert. See table A.4 for an explanation for the classification. We run a robustness check on an alternative classification.

We use cross-sectional data, based on the IFLS-5 (2014). We follow the specific typology described in the previous section. We only use couples in which *no* partner reports *no use*. The dependent variable is $CovertMethod_h$. All methods are coded as either covert (value of $CovertMethod_h = 1$) or non-covert (value of $CovertMethod_h = 0$). In the appendix, we provide alternative specifications for our classification of covert methods as a robustness check.[29]

$$CovertMethod_h = \beta_0 + \beta_1 SCC_h + \beta_2 INDIVIDUAL_i + \beta_3 HOUSEHOLD_h + \rho_1 PROV_i + \varepsilon_i \quad (1.3)$$

Results. Results are presented in figure 1.3 and the regression table is presented in appendix table A.9. We find that all discordant couples have estimated higher rates of covert contraceptive use compared to baseline couples that agree on the wife as the sole decision maker, although only two out of three are statistically different from zero. We find slightly higher rates of covert method use in couples that agree on joint decision-making compared to the omitted group of couples that agree on the wife as the sole

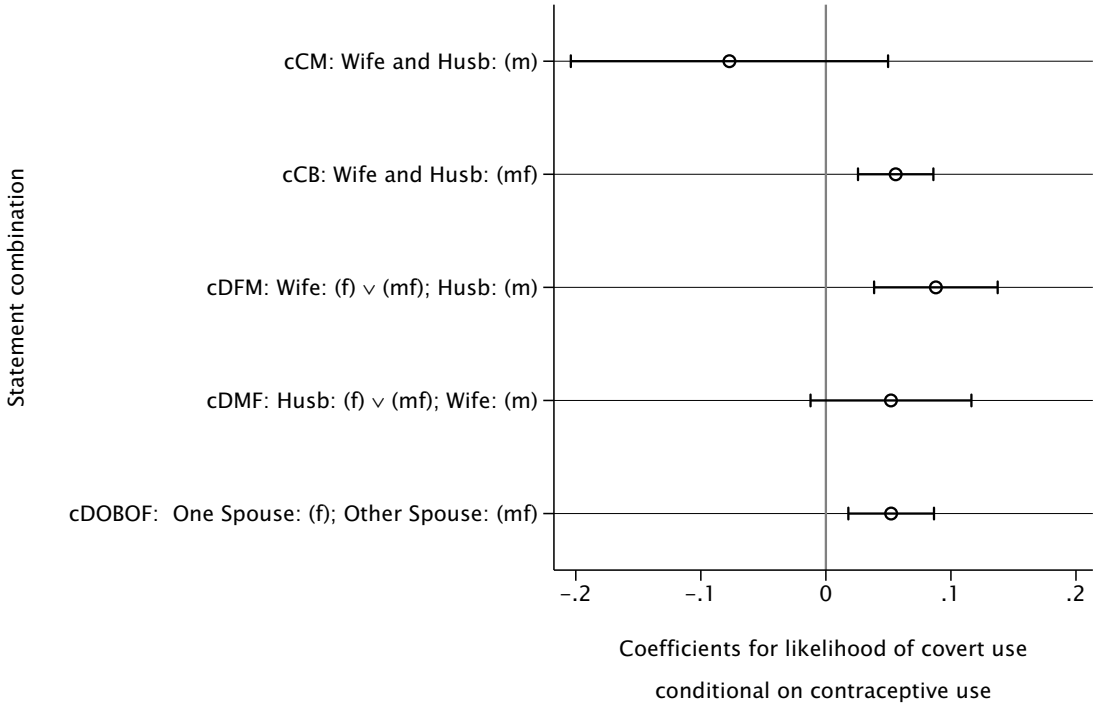
27. Gasca and Becker (2017, p. 4) define contraceptive methods as "The contraceptive methods most likely used covertly are those that can be easily hidden from a spouse, have few side-effects and can be easily utilized."

28. Biddlecom and Fapohunda (1998) show that women use methods both covertly and openly that are easier and harder to conceal.

29. Table A.26 presents results for the same estimation if users of the pill are excluded from the analysis. Table A.27 presents results for the same estimation if the pill is coded as a covert method.

decision maker. As will be discussed later, this result is not robust to classifying *the pill* as a covert method, as is observable in appendix tables A.26 and A.27.

Figure 1.3.: Covert method use across statement combinations conditional on contraceptive use: coefficient estimates



Note: Standard OLS estimate; **Dependent variable:** covert contraceptive use (Yes = 1/No = 0) conditional on contraceptive use; **Baseline group:** concordant couples, perceiving wife as the sole decision maker; **Data:** IFLS-5 wave (2014), cross-sectional data, one observation is one couple; **Decision making domain:** "whether you and your spouse use contraception"; **Typology:** cCM: concordant report of husband as sole decision maker, cCB: concordant report of joint decision making, cDFM: wife perceives female decision making power, husband does not, cDMF: husband perceives female decision making power, wife does not, cDOBOF: both spouses perceive female decision making power, one of them perceives a sole female decision maker, the other joint decision making; Robust standard errors are clustered at regional level; Ticks indicate 95 percent confidence interval.

1.6.3. Prediction of outcomes with both spouses' perspectives

By default, most previous studies use the wife's perspective only. The previous part of this section established that outcomes vary by type of con- and discordance. In this section, we explicitly test whether the prediction of outcomes can be improved if the husband's perspective is taken into account.

Approach. We estimate all previous baseline models for labor and contraceptive use outcomes again with a new, alternative typology. Before, we used a typology that *sorts* couples by their perception of female decision-making power.

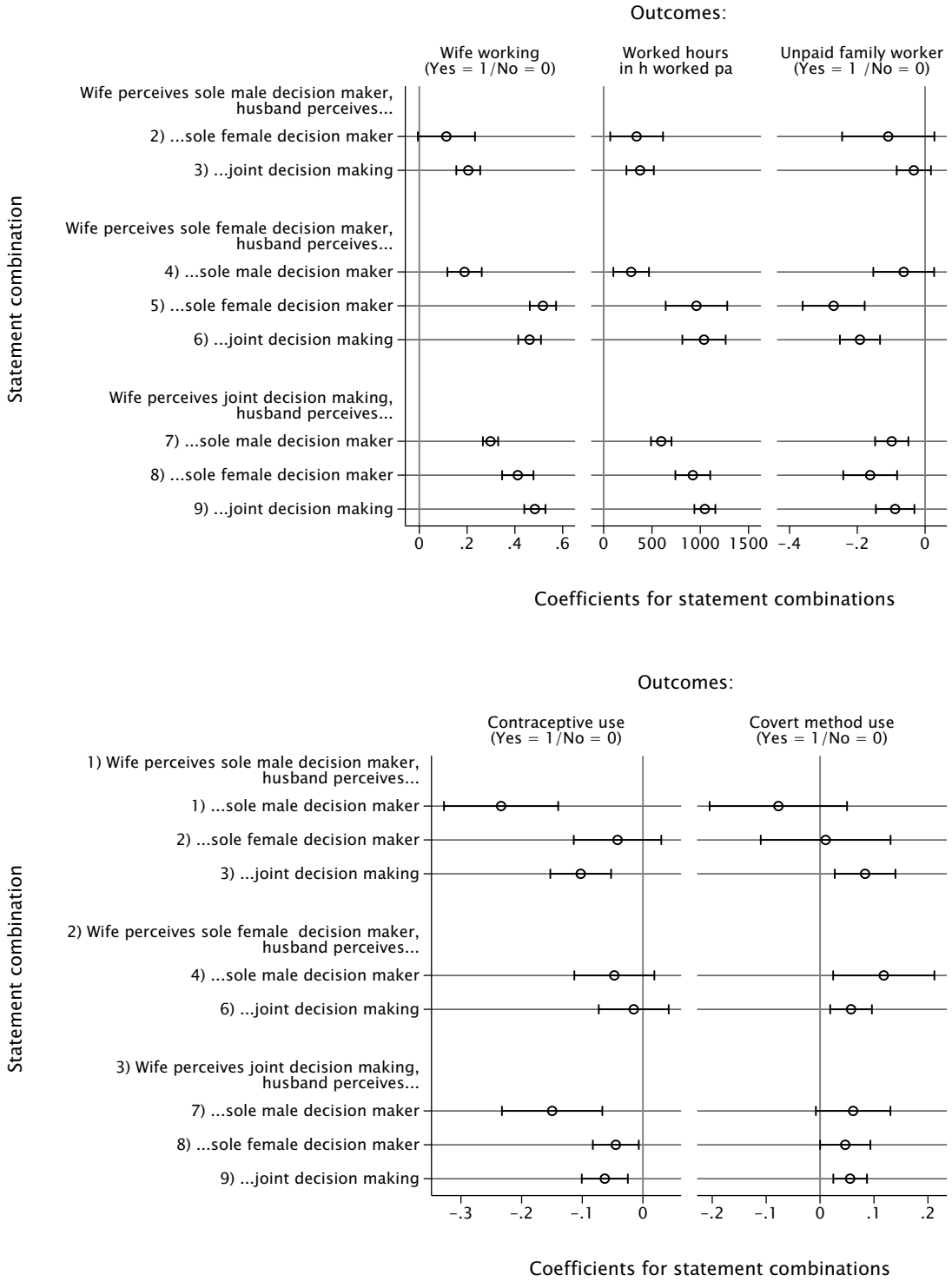
Now, we first sort couples by the wife's response. We limit our analysis to three potential responses: perception of a sole male decision maker, a sole female decision maker and joint decision-making. We omit couples in which the wife responded that neither partner makes the decision.[30] Within the female response groups, we create three new-subgroups for each possible response by the husband. Again, we limit our analysis to the three responses stated previously. The combination of three female responses and three male responses yields nine possible statement combinations.

We introduce one indicator variable for each possible combination. We use this indicator vector instead of the previous vectors (SCL and SCC) in the baseline model equations 1.1 and 1.2. All other baseline specifications remain the same. We omit the same groups as before. For the estimation of labor outcomes, we omit the indicator for the couples in which both spouses perceive a sole male decision maker. For contraceptive and covert contraceptive use outcomes, we omit the indicator for concordant couples that perceive a sole female decision maker.

Results. In figure 1.4, we find that including the husband's perspective can improve the prediction of outcomes. Outcomes often vary by type of male response, holding the female response constant. For labor outcomes, the baseline group is that of concordant couples, in which both spouses perceive a sole male decision maker. Within all three subgroups (2,3 and 4,5,6 and 7,8,9), female labor force participation is higher if the husband states that decisions are made jointly as opposed to by himself as the sole decision maker. The male perspective hence allows predicting outcomes more precisely. Results on working hours confirm this. These results are substantial. The estimated difference between groups 4 and 6 is around 700 annual working hours. The only within-group difference in the unpaid family worker column is found for groups 4 and 5. For contraceptive use, we also confirm that in some cases, the male perspective will improve our predictions. If the wife perceives a sole male decision maker and if her husband concurs, the likelihood of her using contraceptives is significantly lower than if her husband reports that she is the sole decision maker. With respect to covert use, we observe that rates of covert use are higher in the cases 4 and 6 compared to the omitted group number 5.

30. For the response to the question "whether you/your spouse works", we thereby exclude less than 1 percent of the total sample. For the contraception analysis, we exclude around 34 percent of all observations, because we also exclude all couples that report *no use*. We do so as inclusion would challenge interpretation. The no use response describes an outcome as well as a decision.

Figure 1.4.: Prediction of outcomes with both spouses' perspectives: coefficient estimates



Note: Standard OLS estimate; Dependent variables listed on top of subgraphs; **Upper three graphs: Baseline group:** concordant couples, perceiving husband as the decision maker; **Decision making domain:** "Whether you/your spouse works"; **Lower two graphs: Baseline group:** concordant couples, perceiving wife as the sole decision maker; **Decision making domain:** "whether you and your spouse use contraception"; **Data:** IFLS-5 wave (2014), cross-sectional data, one observation is one couple; Robust standard errors are clustered at regional level; Ticks indicate 95 percent confidence interval.

1.7. Changes in female economic resources and decision-making power

We established that female outcomes vary between groups that differ in their perception of female power. In this section, we study how male and female perception diverge *if* female economic power changes. The goal is to learn about (potential) bias that could arise in studies that are based on one-sided surveys. In order to achieve this, we pursue two approaches: one based on a fixed effects model and one based on an instrumental variable model.

1.7.1. Fixed effects model

We study the relationship between changes in female income share and changes in female decision-making power from both spouses' perspectives. We are interested in whether the association between either spouse's perception is more responsive to changes in female income contribution. We employ a fixed effects model to account for unobserved heterogeneity and to study changes in female economic resources and decision power over time. The fixed effects approach can control for unobserved heterogeneity between couples. However, it cannot account for the *dynamic*, bi-directional interaction of income and power over time. That is, we cannot determine in which way economic resources and decision-making power influence each other. We argue, however, that the general *strength* of relationship is of informational value.

Approach. We run two fixed effects models per decision domain, one to reflect the husband's perception and one to reflect the wife's perception. The main dependent variable of interest is whether the wife (husband) perceives the wife to have decision-making power in a given domain. Our dependent variable takes on the value 1 if the wife (husband) perceives female decision-making power in a domain, and zero if not. Accordingly, we run 26 models: two for each of the 13 domains. We use panel data, based on IFLS waves 3, 4 and 5. The vector of all 26 outcome variables is denoted $FemDecShare_{w,d,i,t}$. PFDMP expresses the perceived female decision-making power (PFDMP) of wife w in decision domain d as perceived by wife or husband i in period t.

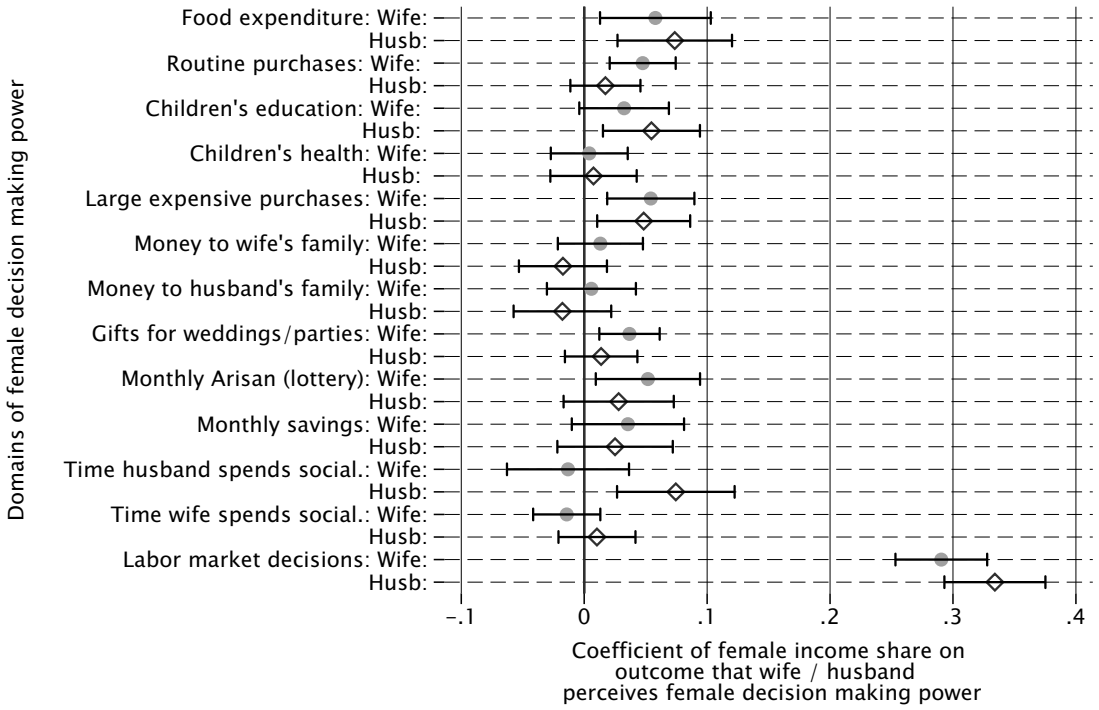
The predicting variable of interest is the female's income share $FemIncShare_{h,t}$. The income share is the wife's contribution to household labor income divided by the sum of wife's and husband's income.[31] We use individual income as stated by the household head. We set income to zero if a subject has "not worked in past 12 months". We do not exclude unpaid family workers. We do so because economic inactivity is relevant to define a baseline of power. The income share is then calculated as the female income over the sum of male and female income. The data offer estimates of income based on *unfolding brackets* for individuals who did not report their income. We disregard those, as including them would require making many assumptions. The remaining variables are analogous to previous models. Time-invariant variables, such as educational level, are excluded. The model is specified as follows.

$$FemDecShare_{w,d,i,t} = \beta_1 FemIncShare_{h,t} + \beta_2 INDIVIDUAL_{i,t} + \beta_3 HOUSEHOLD_{h,t} + \rho_1 PROV_{i,t} + \gamma_t + \alpha_i + \varepsilon_{i,t} \quad (1.4)$$

Results. We can confirm the association between female economic and decision-making power. The results are presented in figure 1.9. The figure presents the coefficients of female income share across all 26 models. Detailed estimates are to be found in the appendix section A.3.3. A positive increase in female income share over time is associated with a perceived female decision-making power increase. This holds true for both perspectives. From the husband's perspective, female income share is positively associated with female decision power in the domains of children's education and the husband's time spent socializing. Women tend to perceive their own decision-making power in the domains of routine purchases, monthly Arisan (lottery), and gifts for weddings/parties when their income share increases. Both women and men tend to perceive more female power over food expenditure, large, expensive purchases and labor market decisions when the female income share increases. The within-domain estimates are not statistically significantly different from each other between men and women. One-sided surveys might draw different conclusions within specific domains, dependent on whether they take the husband's or the wife's perspective into consideration. That is, because some effects are statistically significantly different from zero for women (men), while they are not for men (women). However, there is no statistically significant within-domain difference at the five percent level in the elasticity of the male and female perception of female power with respect to changes in female income.

31. By design, a gain in relative female income share implies a loss in male income share as the function is $1 - FemaleIncomeShare = MaleIncomeShare$.

Table 1.9.: Association of female income and decision-making power across domains: summary of coefficients



Note: Fixed effects estimate; **Dependent variable:** wife (husband) states that wife has decision-making power in specific domain; **Data:** IFLS waves 3, 4 and 5, panel data 2000 to 2014, one observation is one couple; Ticks indicate 95 percent confidence interval; Please see appendix section A.3.3 for numeric estimation results.

1.7.2. Instrumental variable model

Approach. For this model, we exploit natural disasters and crop loss as a shock to household income shares to study the relationship between female economic and decision-making power. For the first stage, we regress the female share in labor income on the instrument vector $SHOCK_i$. In order to code $SHOCK_i$, we use information in IFLS-3 wave (2000) data on exposure to two possible events (natural disasters and crop loss) and six possible household labor supply reactions. For example, one possible reaction is that a household member started working for pay. All recorded events are listed in appendix table A.5 and the possible household reactions are listed in appendix table A.6

There is one indicator variable for each of the six reactions to each of the two events. Thus, the vector comprises twelve indicator variables in total. The default coding of each indicator variable is 0. If a household has been exposed to an event **and** reacted to the event, the respective indicator variable is coded as 1. For example, the indicator variable for event e_1 and reaction r_1 is coded 1 if and only if the household responds that it was affected by disaster e_1 **and** reacted in way r_1. If it did neither show reactions $r_{2,3,4,5,6}$ to event e_1 nor experienced e_2, all other eleven variables will also take on the value zero.

The household does not report whether it was the wife or husband or another household member who adjusted their labor supply. However, first stage estimates in appendix table A.19 suggest that crop losses reduce the female income share, on average. This is in line with previous findings.[32] Following natural disasters, the picture is more nuanced: If a household member changed or quit their job, the female income share tends to increase. If a family member takes on an additional job, the female income share tends to decrease (This appears sensible given that the *average* wife does not work. The average additional job is thus expected to be taken on by the husband.) As crop loss and the severity of exposure to natural disasters are associated with ownership of land, we estimate two models. One for the subsample of landowners and one for the full sample, controlling for land ownership. The vectors $INDIVIDUAL_i$, $HOUSEHOLD_h$, $PROV_i$ and ε_i are modeled in the same way as in the previous section. The resulting first stage equation is as follows.

32. Kochar (1999) find for India that men tend to adjust their labor supply in order to account for crop loss. Cameron and Worswick (2003) use IFLS 1990 data and find that the average household response to crop loss is a shift from *unproductive* farm work to more productive work outside of the household.

$$FemIncShare_h = \beta_0 + \beta_1 SHOCK_i + \beta_2 INDIVIDUAL_i + \beta_3 HOUSEHOLD_h + \rho_1 PROV_i + \varepsilon_i \quad (1.5)$$

The reduced form equation estimates the effect of instrumented changes in female income share on perceived female decision-making power. Female decision-making share is measured as decision share over all domains ($FemDecShare_i$). We estimate it for husband and wife separately. For example, if a couple reports on 13 domains and the wife perceives female decision-making power in twelve out of these domains, the variable would take on the value .92. The husband's assessment might differ, taking on any value between 0 and 1. A perfectly concordant egalitarian couple would be assigned the value one for both spouses.

$$FemDecShare_i = \beta_0 + \beta_1 FemIncShare_h + \beta_2 INDIVIDUAL_i + \beta_3 HOUSEHOLD_h + \rho_1 PROV_i + \varepsilon_i \quad (1.6)$$

Results. If the female income share changes, women adjust their perception of their own power slightly more than men adjust their perception of female power. However, this difference is not statistically significant. Table 1.10 presents the reduced form results. Appendix table A.19 provides the first stage estimates. One-sided surveys will yield similar conclusions, irrespective of whether they use the husband's or the wife's perspective. In appendix table A.17, we use a fixed effects approach to estimate the relationship between the wife's income share and her overall decision share. The analysis yields results similar to this estimation, supporting the robustness of the IV approach.[33]

33. This robustness check is itself robust to exclusion of the labor supply decision domain as a part of the overall decision share. The female and male perspective is still similar but the overall effect size is smaller. The estimate is reported in appendix table A.18.

Table 1.10.: IV Model reduced form results: instrumented female income share's effect on wife's and husband's perception of female power

	Dependent variable: overall female decision-making share from the perspective of the...			
	(1) ...wife (Land-owners only)	(2) ...husband (Land-owners only)	(3) ...wife (Full sample)	(4) ...husband (Full sample)
Female income share	0.561*** (0.210)	0.547*** (0.200)	0.532** (0.214)	0.489** (0.198)
Any land			0.034*** (0.012)	0.036*** (0.011)
Constant	0.561*** (0.105)	0.652*** (0.104)	0.466*** (0.091)	0.524*** (0.088)
Control for (spousal) age and education	✓	✓	✓	✓
Control for HH income, N HH adults, N HH children, migration	✓	✓	✓	✓
Observations	2,276	2,276	6,053	6,053
Underid. F-Statistic	20.8	20.8	18.5	18.5
Underid. P-val.	0.0228	0.0228	0.0468	0.0468
Weak id. F-Statistic	13.49	13.49	14.45	14.45
Overid. F-Statistic	7.625	13.45	10.15	20.50
Region dummies	Yes	Yes	Yes	Yes
Robust standard errors	Yes	Yes	Yes	Yes

Notes: Reduced form IV results; **Dependent variable:** female share in household decision-making across all decision domains; Individual level attributes are wife's attributes by default, Spouse (Sp:) refers to husband's attributes; **Data:** IFLS-3 wave (2000), cross-sectional data, one observation is one couple; ***/**/* indicate significance at the 1%/5%/10% level.

1.8. Robustness checks

In order to verify the robustness of our results, we have conducted six checks. First, we run the cross-sectional models on alternative data. We use alternative data from the IFLS-4 Survey (2007).[34] Results are presented in appendix section A.5.1. Except for one bar, the histogram results are of similar shape. Additionally, the labor outcomes analysis supports previous conclusions. The contraception results are statistically insignificant at the five percent level, however. This *deviant* finding might be driven by two factors. Due to survey design, new and young couples enter the survey from 2007 through 2014. Relatedly, we observe a change in contraception method mix over time (see appendix table A.20). A longitudinal survey design with a focus on contraception might shed light on the question how and why contraceptive use changed over time.

Secondly, we use an alternative covert method specification in appendix section A.5.2. We first exclude all pill users and then classify *the pill* as covert.[35] The findings are not robust to this specification. Effects are reversed. However, we still observe differences between discordant groups. Further, we test whether our results are robust to the exclusion of specific groups of women, who might not intend to prevent pregnancy. This assessment can be found in appendix table A.8 for the contraception analysis and in appendix table A.9 for the covert method use analysis. Results are robust to their exclusion. We find that our assessment in its general conclusions is robust to the exclusion of women that a) do state that they wish to have another child[36] and b) state that they are not fertile at the moment, for instance, due to menopause. However, exclusion restriction a) (in column 3 of appendix table A.8) reduces the statistical significance of the estimates. Same holds true for a third specification that excludes both groups, a and b.

Thirdly, one might suspect that discordance is a phenomenon that is mainly driven by a subsample that is consistently discordant while other couples are consistently concordant. To check this, we study the within-household covariance of discordance across domains (results not reported in this paper). We cannot confirm that it is a sin-

34. Our cross-section estimates on the association between spousal statements and labor outcomes and contraceptive use outcomes are based on cross-section data from IFLS-5 wave (2014).

35. The pill can be concealed and taken without the husband's knowledge. However, this might prove more difficult than using other methods covertly, such as injections. Chikovore et al. (2002, p. 329) discuss for a sample in Zimbabwe how contraceptive pills cause a "hide-and-seek game" at home.

36. The question does not specify the time at which women would like to receive another child.

gle group of individuals displaying *consistent inconsistency*. The average discordance rate is 37 percent, and only 10 percent of all couples show now discordance at all. Appendix table A.36 cross-tabulates statement combinations in the labor supply and contraceptive decision domains. While there is some correlation between statements, it is not possible to predict one with the other. Remarkably though, 49.3 percent of couples that agree on joint decision-making on labor supply also do so on contraception.

Fourthly, we check whether our model is robust to alternative specifications of household income shares. In our main specifications, we assign zero income to individuals whose spouses report income and who have not worked in the past twelve months. We do so, as we do consider this to be a better reflection of true average female economic power than dropping all observations of economically inactive women. For example, unpaid family workers are disproportionately often women and usually do not report income. The fact that they do not have command over disposable income is an important empirical insight. However, one might be afraid of this biasing our results, in particular, if non-reporting is associated with power. We provide an alternative specification in appendix section A.5.3. We code non-reported income as missing (income is for example not reported if an individual did not work in the past 12 months) and omit the full couple for the analysis. Our findings are robust to this alternative specification, but standard errors grow, as one would expect. Results for the fixed effects analysis also do not differ, if one limits the analysis to couples in which the wife's share decreases (increases).

Fifthly, we check the validity of our instrumental variable approach to address the concern that a subset of regions was particularly exposed to shocks. If these regions also differed in other unobservable ways from the other regions, this might induce bias. Appendix section A.4.2 shows that all but one regions have been subject to at least one of the shocks we instrument for. We also provide an alternative coding of our instrument in section A.5.4. This is based on the inclusion of additional instruments. This model does not pass identification tests. It yields estimates of a similar order of magnitude as our main results.

Sixthly, we check whether our main OLS results are robust to probit estimation. We use OLS to ease interpretation and because we are interested in sample average estimates. Some dependent variables (eg labor market participation, contraceptive use) are binary, suggesting probit estimation. The marginal probit estimates for labor and contraceptive outcomes offer the same conclusions as OLS estimates do.

1.9. Discussion

This study's overall discordance levels are in line with previous publications. In a review Allendorf (2007) document average discordance rates from 50 to 75 percent across various previous studies and their own. Further, we document higher self-perceived power by females compared to males' perceptions. Ambler et al. (2017) review the previous literature and find a mixed picture in this regard.

Results are in general agreement with the literature on discordance and health and well-being outcomes (Allendorf, 2007; Ambler et al., 2017). However, analyses of Ambler et al. (2017) propose differences between groups that we classify as *strongly discordant* – groups in which the husband or the wife does not perceive female decision-making power, while the other spouse does. We cannot confirm fundamental differences between these groups (Results are similar for women, who perceive their own power, without their spouses concurring and for those women who do not perceive any female decision-making power while their spouses do). In general, our observations suggest that future research will benefit from taking into account both spouses' perceptions when assessing female decision-making power. The covert contraceptive use analysis provides new insight on decision-making power and types of contraceptive use. These findings are also supported by previous literature that documented different fertility preferences by men and women (Rasul, 2008a). However, the results are sensitive to the chosen classification of methods into covert and non-covert. In-depth focus group interviews could be one way to understand contraceptive use and gendered perception of power better. Frankenberg and Thomas (2001) suggest that focus groups are helpful in order to learn about decision-making power in the household, as husbands, in particular, provide more truthful answers.

The findings on changing perceptions of female economic resources and decision-making power are also closely related to literature that evaluates changes in the perception of female power in general. For example, Beaman et al. (2009) document that exposure to powerful women can shift the perceptions of gender roles of men and women. Future work should elaborate on interventions that shift the perception of either or both sexes.

So far, there is no comprehensive framework for the reasons why partners can be discordant. Supportingly, Ambler et al. (2017) posit that there is no household model

that permits the inclusion of diverging spousal preferences. Future work should aspire to develop a theory in order to conceptualize these and related empirical findings. In this regard, the appendix offers some additional discussion of potential moderators of discordance and power. Section A.6.1 provides descriptive information on the association of relative female and male education and female power. Holding male education constant, higher female education is associated with higher female power. Section A.6.2 estimates the relationship between potential sources of female power and the husband's and wife's perception of female power. We find that from the wife's perspective, female business ownership is positively associated with female decision-making power, while it is not from the husband's perspective. From both perspectives, female education is positively associated with female decision-making power as is the overall household income. In appendix section A.6.4 we present outcomes for the four possible combinations of spousal presence at interviews. We find that the average rates of total discordance across all dimensions are highest where neither spouse was present during the other spouse's interview.

1.10. Conclusion

We observe high rates of discordance and find them to be associated with female income shares as well as labor and contraceptive outcomes. We conclude that taking the husband's perspective into account can improve prediction accuracy. We use two different approaches to test the sensitivity of both husbands' and wives' perceptions of female decision-making power to changes in female income shares. Summarizing, the results suggest that at the population level, one-sided surveys provide a good proxy for the reaction of either gender to economic shocks. However, cross-section estimates suggest that females' outcomes will vary in those families where both spouses' reactions correlate vs. those where only one spouse changes its perception of female power.

Female empowerment is one of the central topics in international development, which is reflected in its relevance in policy-making and academia. In international development, many interventions aim at increasing female decision-making power in order to improve the outcomes of women and children. This study suggests that female power is relevant for female outcomes but it also shows that female power is not an objective measure that is independent of subjective evaluation. Rather, it is the interplay of spouses' perceptions that predicts female well-being.

This study motivates a paradigm shift in development economics, moving away from a strong focus on women and taking the husband into the picture to promote female empowerment with the help of both partners.

2. Sitting on a volcano: domestic violence in Indonesia following two volcano eruptions

2.1. Introduction

Natural disasters threaten the well-being of affected individuals and communities.[1] Global warming is expected to augment this threat by increasing the frequency and severity of extreme weather conditions (Van Aalst, 2006; Watson and Albritton, 2001; Cavallo and Noy, 2010). The increasing prevalence of natural disasters and their significant impact make them subject to extensive research. So far, the range of documented outcomes in longitudinal studies has been comparably narrow, and the majority of publications rely on post-disaster data. This is why many scholars have articulated the need for more (longitudinal) evidence on how disasters affect populations (Rezaeian, 2013; Sety, James, and Breckenridge, 2014; Cavallo and Noy, 2010; Parkinson and Zara, 2013; World Health Organization (WHO), 2002; Rosborough, Chan, and Parmar, 2009; Schumacher et al., 2010).

This study contributes to this emerging strand of the literature that documents the association of disasters and violence with novel panel data from Indonesia. It is the first one to provide pre- and post-treatment family-level data on the impact of volcano eruptions on domestic violence. In addition, it offers evidence for the channels by which natural disasters might cause domestic violence.

The causal chain from the occurrence of natural disasters to domestic violence is detailed hereafter. Previous evidence shows that natural disasters cause mental distress in affected populations (Neria, Nandi, and Galea, 2008). The causes of disaster-related distress are multi-fold. They can be psychological (eg existential fear), social (eg loss of social network) or economic (eg loss of livelihood) (see Rezaeian (2013) and Overstreet et al. (2011) for frameworks). Victims' distress symptoms can range from feelings of helplessness to posttraumatic stress disorder (PTSD).[2] The anticipation or experience of recurrent disasters, such as volcano eruptions, can amplify distress (Overstreet et al., 2011). Suffering from mental distress, individuals can develop feelings of aggression and outward, interpersonal violent behavior (Berkowitz, 1993; Curtis, Miller, and Berry, 2000; Denlay and Shrader, 2000). The likelihood of violent behavior is increased

1. Geophysical disasters caused economic damage of USD 763 billion between 1995 and 2015. Asia experienced 37 percent of the global economic loss between 1994 and 2015. Volcanic activity accounted for eight percent of all natural disasters globally during this period (Wahlstrom and Guha-Sapir, 2015).

2. For examples of studies documenting increases in symptoms of mental distress, eg PTSD, after natural disasters, see Fernandez et al. (2017), Goenjian et al. (2000), Neria, Nandi, and Galea (2008), Paxson et al. (2012), Rezaeian (2013), Rhodes et al. (2010), Rubonis and Bickman (1991), and Warsini et al. (2014).

by multiple factors, for example, worsened living conditions or a lack of social networks and social control (Curtis, Miller, and Berry, 2000; Rezaeian, 2013). Rezaeian (2013) provides a framework and reviews previous literature that links disaster experience to (interpersonal) violence.

In this study, the impact of two eruptions on rates of domestic violence and four alternative outcomes is estimated with a difference-in-differences approach. On a descriptive level, I confirm an increase in domestic violence in the treatment group. The affected communities suffer from lower average household expenditure which is expected to increase distress. Further, increased rates of alcohol/drug abuse and lowered emotional well-being in affected populations point to increased levels of mental distress (alcohol abuse is also associated with IPV).[3] A synthetic control approach allows the verification of these results. A subsample of families with *internally displaced people* (IDP) status displays substantially higher levels of domestic violence after the volcanoes' eruptions. This subsample has previously suffered from natural disasters and proves to be particularly vulnerable to repeated disaster exposure. These families suffer from a loss of livelihood, lack of a social network and augmented feelings of a threat of disaster recurrence. The treatment coefficient estimate suggests an increase of nine percentage points.

Data are provided by an NGO that serves rural and urban communities across Indonesia. I use survey and observational data from 2,024 families. The families receive support based on their vulnerability to family breakdown. In late 2013 and early 2014, some of the families have suffered from the eruption of the two volcanoes Mount Kelud and Mount Merapi. The volcanoes' eruptions caused ashfall, evacuations, and deaths across multiple hundreds of kilometers on Java, Indonesia's main island.

This study's results suggest an increase in domestic violence after natural disasters. It further points at the roles of economic conditions, IDP status and social networks of affected communities. Policymakers and emergency response organizations should consider the causes of multi-fold non-economic outcomes of natural disasters when designing interventions.

The remainder of this study is structured as follows. First, I will review related studies

3. McFarlane (1998) reviews studies on the association of alcohol abuse and PTSD. Sonne et al. (2009) present evidence on the sequence of the onset of PTSD and alcohol abuse respectively. Bueno and Henderson (2017) explore the association of IPV with alcohol abuse. Bech et al. (2003) discuss the relationship between mental distress and well-being.

that survey the impact of natural disasters and domestic violence separately. I will then discuss how both strains of literature are linked. I will provide an overview of the Indonesian cultural context. Secondly, I will present my empirical strategy and discuss alternative specifications. Afterwards, I will present the results that are accompanied by robustness checks. I will conclude by discussing the implications of the findings.

2.2. The relationship between natural disasters and domestic violence

In the following section, I will first present previous work that surveys the consequences of natural disasters on affected populations. I will then present evidence on the determinants and consequences of domestic violence and its local context in Indonesia. I will conclude by discussing literature linking natural disasters with domestic violence.

2.2.1. Natural disasters

Natural disasters such as hurricanes, volcano eruptions, earthquakes, and cyclones threaten the well-being of affected communities. Global warming is expected to augment this threat via an increase in the frequency and severity of those disasters that are weather-linked (Cavallo and Noy, 2010; Van Aalst, 2006; Watson and Albritton, 2001).

The strain of studies in economics dedicated to evaluating the impact of natural disasters is comparably novel. Evidence regarding the overall impact on long-term economic growth remains mixed (Loayza et al., 2012). Gignoux and Menéndez (2016) identified public investment in response to natural disasters as a critical moderator that can turn short-run losses in long-term gains. In their study on the welfare impact of earthquakes in rural Indonesia, they found that public investment can lead to local welfare gains within six to twelve years following an earthquake. Arouri, Nguyen, and Youssef (2015) document adverse income effects in rural Vietnam following storms, floods, and droughts while Deryugina, Kawano, and Levitt (2018) find income increases for some communities affected by Katrina.[4] For short-run outcomes, Noy (2009) found that disasters can lead to slowdowns in production and that these are expected to be worse in developing countries. Accordingly, Strobl (2012) estimate an average output drop of .83 percentage points following hurricane strikes in Central America and Caribbean regions.

Apart from macroeconomic evaluations, micro-level assessments have been conducted with respect to human capital accumulation (Caruso, 2017; Baez, Fuente, and Santos, 2010; Janvry et al., 2006; Ferreira and Schady, 2009; Gitter and Barham, 2007), income and expenditure (Arouri, Nguyen, and Youssef, 2015; Deryugina, Kawano, and Levitt, 2018), as well as health and physical growth (Caruso, 2017; Hoddinott and Kinsey,

4. In the aftermath of hurricane Katrina, Deryugina, Kawano, and Levitt (2018) find that the incomes of affected individuals outgrow those of unaffected individuals in control communities.

2001; Maccini and Yang, 2009). Caruso (2017) find that the health of young children is particularly susceptible to natural disasters, while Ferreira and Schady (2009) observe lower investment in children in developing countries after droughts.

Despite this wealth of previous publications, few studies address natural disasters as a cause of domestic violence. To my knowledge, there are no longitudinal studies using family level data to observe changes in domestic violence following a natural disaster.[5]

2.2.2. Domestic violence

Domestic violence as a peril to public health

IPV poses a major peril to public health. In line with previous literature, IPV is defined in the following as violence between intimate partners. Domestic violence is defined as general violence in the household, including IPV as well as violence against children and other household members.[6] IPV can result in stress, fear and physical as well as psychological trauma, and incur the sentiment of loss of control (García-Moreno et al., 2013). Moreover, domestic violence is found to negatively affect children born to mothers exposed to violence during pregnancy (Aizer, 2011). Domestic violence towards children can affect their development and might result in trauma and lower ability, among other things (Reading, 2008). According to estimates by the WHO, 30 percent of ever-partnered women experience intimate partner violence (IPV) in their lifetime. In South East Asia, prevalence rates of IPV are even higher than the global average, at an average rate of 37.7 percent (García-Moreno et al., 2013).

Multiple individual and structural risk factors have been associated with domestic violence in general and IPV in specific. A perpetrator's record of violence in the recent past, drug and alcohol abuse, threatening behavior, previous psychological issues, are some traits that are found disproportionately often in offenders (Dutton and Kropp, 2000).[7] Circumstantial moderators of domestic violence include but are not limited to gender-specific labor market conditions (Aizer, 2010; Anderberg et al., 2015), social welfare transfers (Bobonis, González-Brenes, and Castro, 2013; Hidrobo, Peterman, and Heise, 2016) and adverse emotional cues (Card and Dahl, 2011). Two competing

5. An exception is research using rainfall shocks. While one can consider rainfall shocks a natural disaster, their consequences are commonly different from volcano eruptions or earthquakes. They usually pose a non-existential threat, particularly if not accompanied by flooding.

6. For a general discussion of definitions of IPV and domestic violence, please refer to Reading (2008).

7. Both Dutton and Kropp (2000) and Jewkes (2002) provide excellent overviews of this literature and discuss appropriate assessment methods of domestic violence risk.

theories exist on how changes in relative female economic power might change the prevalence of domestic violence. Women might expect to leverage higher economic power to *negotiate* better outcomes in household bargaining. By contrast, one might expect (violent) male backlash as an adverse reaction to a decrease in relative male status (Bueno and Henderson, 2017). In a Sub-Saharan context, Cools and Kotsadam (2017) propose that economic inequality both at the level of the household and at the community is associated with higher IPV rates. With new data from Africa, Alesina, Brioschi, and La Ferrara (2016) argue that ancient cultural norms and current economic conditions interact in a non-trivial way. Empirically, Bueno and Henderson (2017) find that household bargaining based approaches are more predictive of general IPV whereas male backlash theories are so for sexual IPV. The authors also point to the relationship between excessive alcohol consumption and IPV.

Domestic violence in the Indonesian context

There is no systematic, continuous tracking of IPV prevalence on a national level in Indonesia.[8] Both the World Health Organization and the United Nations base their estimates on a SUSENAS (National Census Survey) survey from 2006. In this, 3.07 percent of women reported any lifetime experience of either IPV, non-partner violence or both (UN Women (United Nations Entity for Gender Equality and the Empowerment of Women), 2011; World Health Organization (WHO), 2008). The WHO documents that 66 percent of all reported violence cases are classified as psychological (World Health Organization (WHO), 2008). In contrast, 22 percent of Javanese women of reproductive age that have been part of a pregnancy preparation program reported a "lifetime exposure to sexual and physical violence" (Hayati et al., 2011, p. 1). All numbers deserve cautious interpretation since local norms might lead to biased and possibly understated reporting of IPV.

IPV needs to be understood in the context of norms. For example, in an African context, Alesina, Brioschi, and La Ferrara (2016) show that interaction of ancient norms and female economic power can explain current rates of domestic violence. In Indonesia, IPV has traditionally been considered a personal, private issue that deserves private intervention. This notion prevails until today and might be reinforced by the norm of *harmony* (njaga praja) (Hayati et al., 2013). The norm stresses the protection of the husband's honor towards non-family members. This is likely to result in fewer

8. It was not until the mid-1990s that international organizations started raising awareness for the topic, and initial public reactions ranged from surprise to denial (Blackburn, 2004a). Reported figures vary greatly depending on time, region and source.

reports of IPV incidences. Fewer and potentially biased reports create an unclear empirical picture and a selective public focus on *out-of-home* violence such as non-marital rape and trafficking. Moreover, victims refrain from reporting as they are typically not convinced that things will change in their favor if they come forward (Blackburn, 2004a). In their study, Nilan et al. (2014) found Indonesian men to be reluctant to talk about violence against women. Men tended to engage in victim indictment and overall refutation of the phenomenon of IPV.

Acceptance of IPV is *high* in Indonesia despite political progress.[9] Opposed to global trends, rates of rejection of IPV amongst Indonesian women stayed relatively stable over the recent past (Pierotti, 2013). In the most recent IDHS study, 27 percent of Indonesian women considered wife-beating to be a justified reaction to maternal child neglect (Statistics Indonesia (Badan Pusat Statistik - BPS) National Population and Family Planning Board (BKKBN) Kementerian Kesehatan (Kemenkes—MOH) and ICF International, 2012). Similarly, 24 percent of women considered male violence justified in cases in which women left home without giving notice to their husbands. Both numbers are higher for younger age groups. Compared to women, men report lower levels of acceptance of wife-beating (Statistics Indonesia (Badan Pusat Statistik - BPS) National Population and Family Planning Board (BKKBN) Kementerian Kesehatan (Kemenkes—MOH) and ICF International, 2012) – a finding potentially driven by social desirability.

In the Indonesian context, Hayati et al. (2013) found that female economic independence and conservative values are associated with higher rates of IPV. Hayati et al. (2011) show for this study's treatment area, that female main breadwinners were at particular risk of violence experience yet were particularly unwilling to accept help. Nilan et al. (2014) use interviews and find that one self-reported cause of violence in Indonesia is the (perceived) male inability to satisfy female expectations. In particular, financial difficulties seemed to be predictive of intimate partner violence in the context of Indonesia.[10] In the study, 48 interviewed men, all of whom were considered *community leaders*, expressed a discordance between idealistic, partially religiously informed expectations towards men and a more egalitarian reality resulting in men feeling challenged about their identity (Hayati, 2013).

9. In the recent past, the country progressed, for example, by introducing a National Commission on Violence Against Women in 1998 and passing of the Domestic Violence Act in 2004 (Hayati et al., 2011).

10. A more detailed account of the concepts of masculinity in Indonesia and female coping strategies in response to IPV can be found in Hayati (2013).

In summary, previous evidence confirms that domestic violence poses a significant threat to public health in Indonesia. Precise measurement of prevalence is complicated by social norms. Household economics and female economic dependence and independence have both been presented as potential risk factors.

2.2.3. Stress and domestic violence in consequence of natural disasters

This study's findings propose that natural disasters lead to an increase in domestic violence. It thereby relies on previous studies that propose channels through which this happens. Rezaeian (2013) suggests that the psychological, social and economic consequences of natural disasters cause mental distress which in turn causes domestic violence. The following section presents previous evidence on this hypothesis.

The first strand of literature shows how natural disasters cause stress in established populations. A comprehensive account of disaster psychiatry can be found in Usano et al. (2017). Rubonis and Bickman (1991) review 52 empirical studies and find a *small but consistently positive* association of disasters and a subsequent increase in indicators of psychopathology. Goenjian et al. (2000) document higher levels of posttraumatic stress disorder (PTSD) in populations affected by severe earthquakes compared to those exposed to milder trauma. Effects persist without major improvement within 1.5 and 4.5 years following a disaster and that decreases in living conditions, and livelihood can amplify stress. Pre-event conditions such as mental illness predict the later severity of PTSD levels.[11] The recurrent threat of natural disasters, whether anticipated or actual, is expected to augment mental distress further (Overstreet et al., 2011). Two studies have surveyed the impact of Mount Merapi's 2010 eruption on mental health and confirm previous findings for this study's context. Warsini et al. (2014) report higher rates of distress in survivor communities located close to the peak of Mount Merapi for early 2013. Victims attributed distress to *volcanic dust* on roads and *mining and construction* following the events. The same authors also record higher levels of PTSD in affected areas, in particular among women, individuals of young and middle age and homeowners (Warsini et al., 2014). Further, general ev-

11. Fernandez et al. (2017) show for a sample in Chile that pre-disaster attributes in patients (such as panic disorders) predict the likelihood of post-disaster PTSD. Rhodes et al. (2010) find a doubling of mental illness prevalence amongst disadvantaged communities post the Katrina Hurricane and estimate PTSD rates at close to fifty percent. In the same context, Paxson et al. (2012) confirm the long-term consequences of natural disasters by finding that rates of post-traumatic stress symptoms did not return to pre-Katrina levels even 43 to 54 months after the event.

idence suggests that stress is one driver of domestic violence (see for example Runyan et al. (2002) and Bardi and Borgognini-Tarli (2001) for reviews). In this realm, stress can be moderated by both, primary and secondary stressors. Primary stressors are immediate consequences of the disaster such as a felt or actual threat to life. Secondary stressors are stressors that have been caused by the disaster, such as the loss of one's job (Overstreet et al., 2011).

The second strand of literature links stress to domestic violence in the context of natural disasters. Rezaeian (2013) offers a systematic review of the literature that documents the link between mental distress and domestic violence. The author finds that most studies document an increase in interpersonal violence after natural disasters. Indeed, outward, interpersonal violence and aggression are often found reactions of humans exposed to mental distress and feelings of helplessness (Berkowitz, 1993; Curtis, Miller, and Berry, 2000; Denlay and Shrader, 2000). For example, Denlay and Shrader (2000) link violence and aggression as one stress-coping mechanism *chosen* by men, based on a study conducted in the aftermath of Hurricane Mitch in Honduras and Nicaragua. Rezaeian (2013) proposes a model for the relationship between natural disasters and violence that includes the multi-fold psychological, social and economic ways by which disasters affect individuals.[12] Worsened living conditions and lack of social networks and social control have been put forward as important moderators of domestic violence (Curtis, Miller, and Berry, 2000; Rezaeian, 2013). Curtis, Miller, and Berry (2000) argue that the social control of antisocial behavior is reduced in the aftermath of disasters permitting increased rates of violence as an otherwise sanctioned behavior. This is supported by evidence on individuals who have lost their social network. Anastario, Shebab, and Lawry (2009) document high rates of gender-based violence in populations who have been internally displaced following Hurricane Katrina.

The third strand of literature links natural disasters to domestic violence. Previous empirical research on the relationship between natural disasters and domestic violence has mostly relied on cross-section data and post-event surveys or police/administrative data. Adams and Adams (1984) were first to empirically link disaster experience on the one and domestic violence on the other hand. The authors argue that stress resulting from disaster manifests in physiological and psychological responses, among them

12. Rezaeian (2013, p. 1104) proposes that natural disasters lead to "personal threats to life, loss of loved ones, propertyloss [sic]", "interruption and failure of social systems & services", "collapse of social cohesion & harmony" and "massive destruction, Population [sic] displacement".

domestic violence. Based on police reports, they document an increase in domestic violence in the aftermath of the eruption of Mount Saint Helens, a volcano located in the United States of America. Sety, James, and Breckenridge (2014) suggest significant increases in domestic violence after disasters in high-income countries. Based on a post-tsunami survey in Sri Lanka, Fisher (2010) propose that disasters amplify pre-event violence patterns. This is confirmed by research on Hurricane Katrina. In a survey during the aftermath of Katrina, Picardo, Burton, and Naponick (2010) found an increase in existing and new abuse of displaced women. Schumacher et al. (2010) confirm this for individuals who lived in affected areas at the time of Katrina's impact. The authors argue that their study is the first one that assesses pre- and post-disaster IPV prevalence but also remark that this information stems from post-event surveys making it subject to potential recall and reporting biases. Curtis, Miller, and Berry (2000) use public reports on child abuse for periods before and after three natural disasters to survey the link between abuse and disasters. They report increased child abuse rates in two of the three observed samples. Parkinson (2017) interviewed women after bushfires and documents an increase in domestic violence.

The only micro-level panel evidence about the impact of natural disasters on domestic violence is on rainfall shocks.[13] While one can consider rainfall shocks a natural disaster, their consequences are commonly different from volcano eruptions or earthquakes. They usually pose a non-existential threat, particularly if not accompanied by flooding.

So far, the range of documented outcomes in longitudinal studies has been comparably narrow, and most publications rely on post-disaster data. Buttell and Carney (2009) point at multiple challenges in measuring changes in gender-based violence. One of them is the lack of baseline information. This is why many scholars have articulated the need for more (longitudinal) evidence on how disasters affect populations (Cavallo and Noy, 2010; Parkinson and Zara, 2013; Rezaeian, 2013; Rosborough, Chan, and Parmar, 2009; Schumacher et al., 2010; Sety, James, and Breckenridge, 2014; World Health Organization (WHO), 2002). The unique panel dataset on individuals employed in this study allows learning about domestic violence and potential risk factors following a disaster.

13. Sekhri and Storeygard (2014) find a relationship between rainfall shocks and dowry deaths in India. Chin (2011) associate changing power dynamics following rainfall shocks and spousal violence in the same context. Miguel (2005) identifies extreme rainfall as a reason for economically motivated *witch murders* in Tanzania in which relatives murder elderly women during times of economic scarcity.

2.3. Context

2.3.1. Volcano eruptions on Java in 2013 and 2014

Volcano eruptions on Java. With 130 active volcanoes and its location on the *pacific ring of fire*, Indonesia is a country with one of the world's highest rates of seismic activity. In late 2013 and early 2014, it has witnessed two major volcanic eruptions: Mount Kelud and Mount Merapi erupted on its main island Java. The increase in domestic violence in affected areas and its potential causes are subject of this study.

Concerning total evacuations and death toll, Mount Kelud's eruption has been far more severe. However, Mount Merapi's eruption is likely to have been a significant source of stress to surrounding populations. Its eruption in November 2013 caused reminiscence of its last, catastrophic eruption in 2010, which took the lives of 353 individuals. 2013 also marked the end of a three year period in which the volcano displayed a very low level of observable activity.[14] It showed first activity on November 18, 2013, resulting in a 2 km high plume of ash (Wunderman, 2014b). This eruption has been the first significant one, after its 2010 outburst. Following this first incident, Mount Merapi erupted again on multiple occasions in March 2014; creating a 9.8 km high plume on March 27th and affecting neighboring regions and communities in Kemalang and Klaten regency (Wunderman, 2014b). Besides and after one and a half months of increased seismic activity, Mount Kelud erupted on the 13th of February 2014, causing the evacuation of 100,000 people, killing 7 and destroying around 11 thousand buildings in the surrounding communities (Wunderman, 2014a). I observe effects in communities that have not been evacuated but still struck by ashfall.

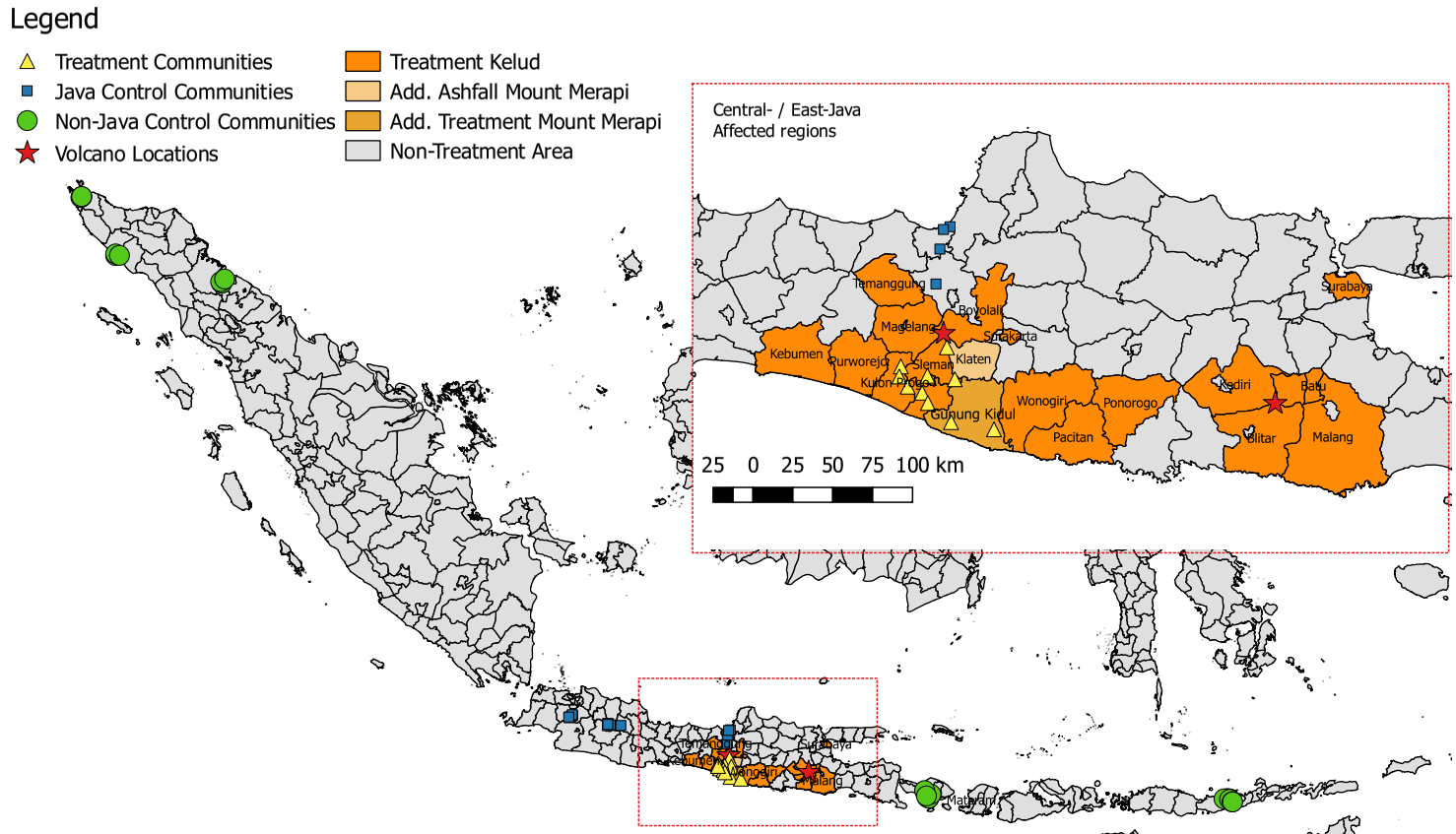
Assignment to treatment and control groups. Figure 2.1 maps the location and assignment of treatment and control communities. Treatment communities are marked with a yellow triangle. Communities on Java that are outside the treatment area are selected as the control group and marked with a blue square. Communities outside Java, marked with a green circle, will serve as a control group for a synthetic control robustness check but are not considered for the main analyses. The primary treatment group is identified based on the latest assessment of the International Federation of Red Cross and Red Crescent Societies (see appendix figure B.2). This group of districts is named Treatment Kelud on the map's legend. I add two additional treatment areas. The first area, named Additional Ashfall Mount Merapi on the map's legend, is added

14. Mount Merapi is an active stratovolcano with a continuous level of activity over time.

based on Wunderman (2014b) and has been exposed to ashfall by Mount Merapi in 2014. This area does not contain any communities in this sample and hence does not affect estimates. A second sub-district named Kabupaten Gunungkidul[15] is included based on the combination of two factors. First, it has been exposed to light ashfall right after Mount Kelud's eruption (The United Nations Office for the Coordination of Humanitarian Affairs (OCHA), 2014). However, it has been excluded from the final assessment of affected areas that have been used for the overall treatment identification (see International Federation of Red Cross and Red Crescent Societies (2014) and map in appendix figure B.2). Secondly, and more importantly, the area Daerah Istimewa Yogyakarta (marked as Additional Treatment Mount Merapi) is closely linked to the primary treatment area and is thereby expected to be exposed to the consequences of the treatment shock. Kabupaten Gunungkidul is part of Daerah Istimewa Yogyakarta which is home to all treatment communities. The used regency level economic data are the same for treatment area and this area. Moreover, a social center located in the treatment area administers its support. The light exposure combined with the economic and social dependency on the treatment area recommend its inclusion as opposed to assigning it to a control area or dropping it. Appendix table B.4.3 runs a robustness check on how exclusion informed results. More detailed information on the eruption is offered by the Red Cross (International Federation of Red Cross and Red Crescent Societies, 2014).

15. Also referred to by Gunung Kidul Regency

Figure 2.1.: Overview map of locations of treatment and control groups

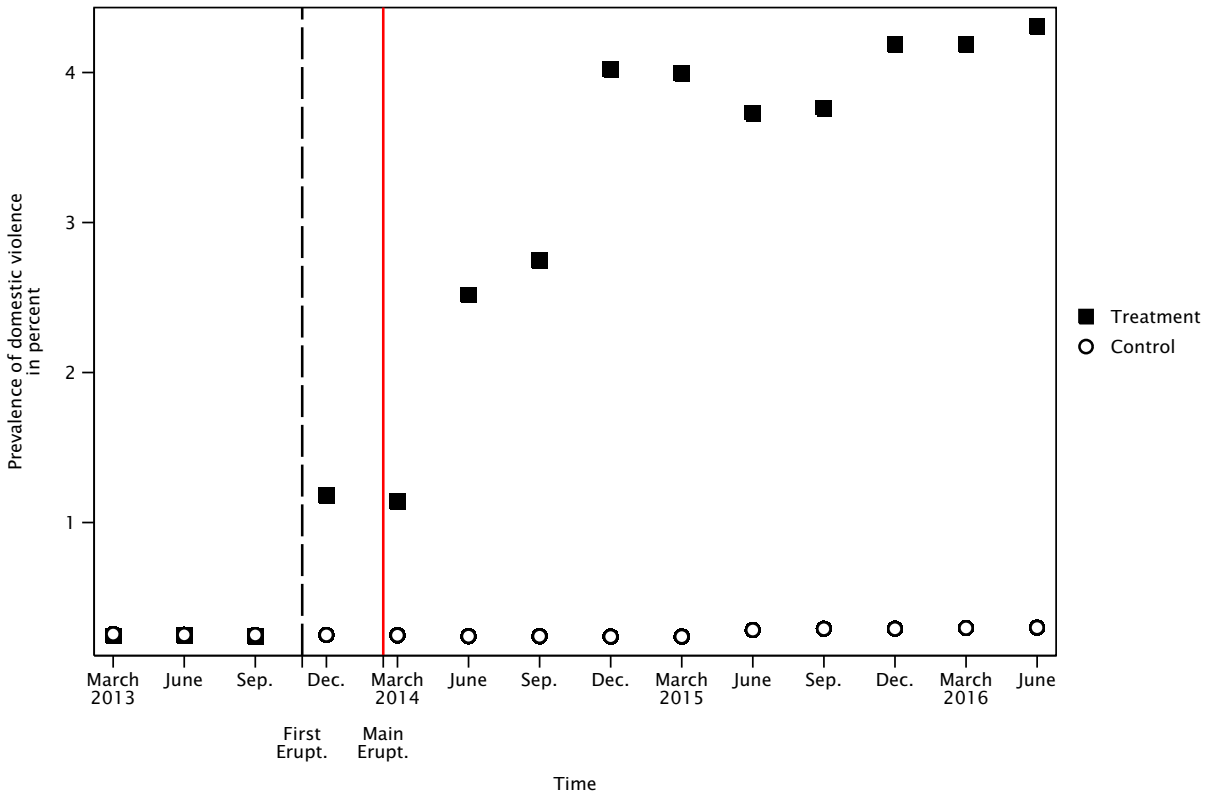


Comment: Identification of treatment based on hazard map provided by the International Federation of Red Cross and Red Crescent Societies. Additional ashfall area based on Wunderman 2014b. Additional treatment area based on second map by UN including area and (administrative, economic and support) associaion of eastern part of DI Yogyakarta with rest of DI Yogyakarta (see Data chapter for further description).

2.3.2. Increase in domestic violence following the eruptions

Table 2.2 displays the increase in domestic violence that is at the center of this study. Beginning with the first eruption of Mount Merapi in November 2013, a trend of increasing rates of domestic violence is observable. The main eruption of the two volcanoes occurred in February and March of 2014. The trend of increasing domestic violence continues until late 2014 – a point at which it reached a plateau of around 4 percent of all households in affected regions. Mount Merapi is geographically closer to the sample communities, but its next outbreak in 2014 has not been as impactful as the 2014 outbreak of Mount Kelud. However, it might still have induced significant stress in communities that have been displaced in its last major outbreak in 2010. It is expected that both volcanoes contribute to the increase in domestic violence.

Figure 2.2.: Domestic violence over time by treatment status



Black squares: treatment areas; White circles: no treatment; First eruption: the first eruption of Mount Merapi, Main eruptions: eruptions of Mount Merapi and Mount Kelud.

2.3.3. Macro-level decrease of average per capita household expenditure following the eruptions

Nationally, Indonesia's economy grew significantly over the past decades. Despite a recession in the late 1990s, its total GDP has more than tripled between 1995 and 2014. In 2014, it reached a level of 888 billion US Dollar[16] More recently, the country's growth rates have been decreasing: from 8.2 percent in 1995 to 5.0 percent in 2014 (OECD, 2015). 2014, the year of interest to this study, does not mark an exceptional year as such but fits the larger trend of decline in economic growth rates.

Regionally, Indonesia is divided into 34 provinces, which are divided into regencies (*Kabupaten*) and cities (*Kota*).[17] Regencies and cities are divided into subdistricts. In figure 2.3 the average household expenditure of treatment and control groups are compared. The data are provided by the World Bank Group. Importantly, the data are sometimes provided at regency-level (for example for West Java) and sometimes at city-level (for example for the city of Yogyakarta). The regency-/city-level data are then assigned to the individual. The average values depicted in the figure 2.3 are thus weighted by the number of observations in each regency and city respectively.

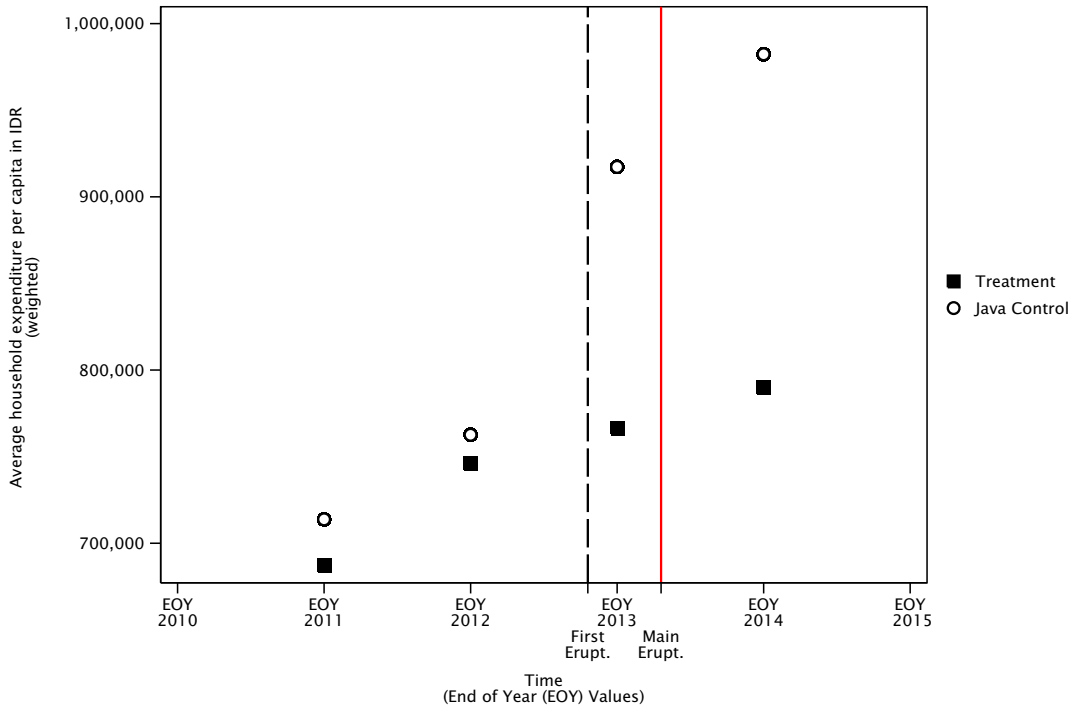
The treatment region exhibits moderate household expenditure growth previous to the eruptions and then slows down thereafter. Exact growth figures are found in appendix table B.1. The net effect of positive growth in urban areas and negative growth in rural areas results in mildly positive but below average net growth figures for the treatment group.

2.3.4. Micro-level development of living conditions and interventions following the eruptions

Figure 2.4 reports the development of living conditions at micro-level. These data are collected by the NGO and are based on the sample used in this study. The observed pattern differs from the macro-level observation of figure 2.3. Two reasons might drive this. First, universal health care (UHC/JKN) has been formally introduced on first of January of 2014 which might have influenced statements on living conditions and explain the slight uptick observed between 2013 and 2014. Secondly, interventions might have moderated the living conditions of the sample.

16. Measured in then-current US Dollar terms.
17. Some regency-sized areas are listed at the level of provinces, eg the Special Region of Yogyakarta or the Special Capital Region of Jakarta.

Figure 2.3.: Weighted regency-/city-level data: development of household expenditure per capita over time (in IDR)

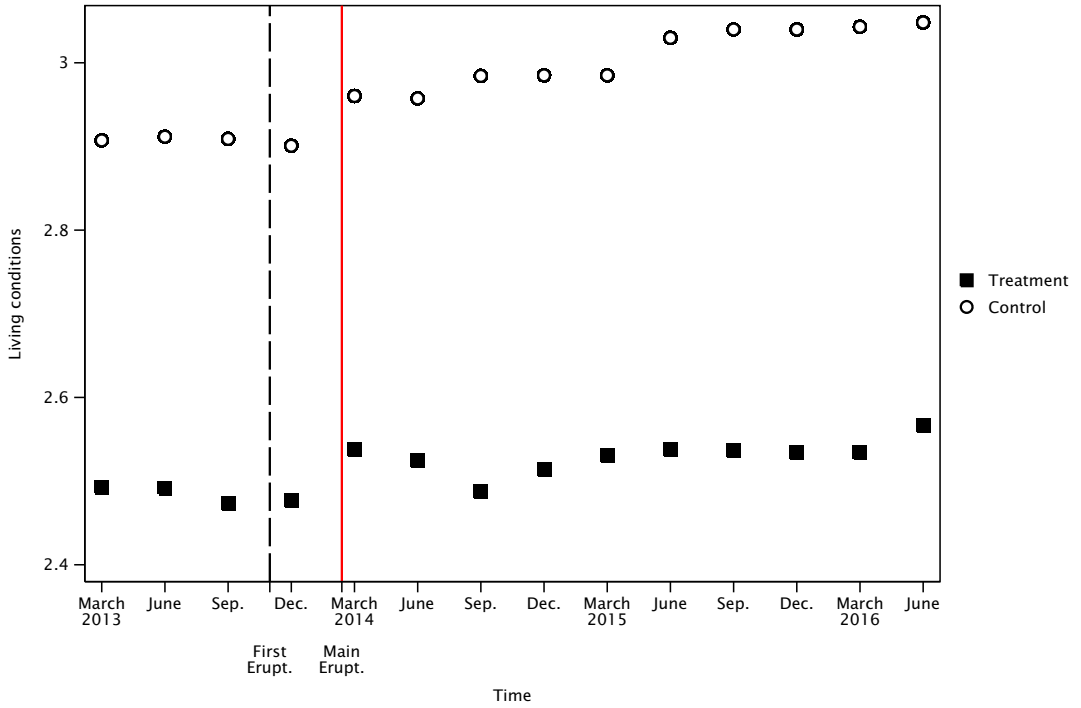


Notes: **Data:** Indonesia Database for Policy and Economic Research (INDO-DAPOER), World Bank Group; Total household expenditure per capita over time in IDR. EOY = End of year. Data for 2015 is not yet available. Assigning of regencies/cities to Java and Non-Java and treatment and control according to table B.1, based on geographical location; First eruption: the first eruption of Mount Merapi, Main eruptions: eruptions of Mount Merapi and Mount Kelud.

To mitigate the adverse effects of natural disasters, Indonesia has created the National Agency for Disaster Management (BNPB). It orchestrates response activities and all relevant stakeholders in the case of a natural disaster-induced emergency (Jati, 2015). Following Mount Kelud's eruption, the BNPB supported the District's Disaster Management Agency (BPBD) in coordinating all relief efforts (International Federation of Red Cross and Red Crescent Societies, 2014). While all construction-related activities were in the hands of the military, economic and non-economic interventions were carried out by a cluster network of governmental and non-governmental agents such as the Red Cross, the World Food Program, Plan International, World Vision, Catholic Relief Services and the Yakkum Emergency Unit. Inter alia, interventions addressed the psycho-social, economic and nutritional well-being of affected communities (International Federation of Red Cross and Red Crescent Societies, 2014). The sample's treatment and control groups received interventions by the NGO, which also provided the data for this study. Unfortunately, there is no exact record of disaster-specific intervention. Going forward, it is hypothesized that these interventions attenuated the severity of the eruptions' ramifications. All negative changes in livelihood outcomes should hence be understood as an upper bound estimation of outcomes in the counterfactual case of a non-intervention and vice versa for positive changes. How this might affect this study's external validity is discussed in the section on identification concerns, section 2.6.[18]

18. A second concern that will be addressed later on is the multidirectional impact that disaster aid might have on the likelihood of domestic violence. It will also be discussed in the section on identification concerns, section 2.6.

Figure 2.4.: Development of living conditions over time



Notes: Black squares: treatment; White circles: no treatment. Y-axis-scale: 4 on the scale equals "Family lives in conditions that are adequate, as per local standards (defined on the level of the community in consultation with key stakeholders)", 1 on the scale equals "Family lives in conditions that are below local standards, and are compromising the personal well-being of individual (and/or family)"; First eruption: the first eruption of Mount Merapi, Main eruptions: eruptions of Mount Merapi and Mount Kelud.

2.4. Empirical strategy

2.4.1. Data

Data are provided by a Non-Governmental Organization (NGO), SOS Children's Villages (SOS). The data have been collected on a quarterly basis by social workers via a standardized questionnaire. The social workers are employed by SOS. The unit of analysis is that of a family. I use an unbalanced quarterly panel of 2,029 families in Indonesia from the first quarter of 2013 to the second quarter of 2016 (The number of observations can be found in appendix table B.4. The lowest number of observations is given for the last quarter, with 6.1 of all observations, the highest number of observations is given in the third quarter with 2029 observations.) The maximum number of observations is 14. For the average family, I use 13.3 quarters of data.

The families are part of the on-going support program by SOS. They are selected on the basis of their likelihood of family breakdown. The support program aims at preventing this family breakdown. Interventions comprehend economic, psycho-social, health and legal support. I only consider families that have been admitted to the support program before 2013 and for which reports are given until at least 2015. All observations of families that do not fulfill these prerequisites are dropped. The data are collected by social workers who are interviewing the families on a monthly basis with a structured questionnaire.

Data on domestic violence are collected by a child protection team. The child protection team collects data from two main sources. The first source is community leaders. These are heads of different local administrative levels (Head of Dukuh, Head of Rukun Warga/Rukun Tetangga) or religious leaders. There is a general awareness that incidences of domestic violence should be reported. The second source is the extended family or the nearest party of the victim. Sometimes, there are self-reports of female victims.

For the difference-in-differences approach, I use a subsample of the full Indonesian sample. I limit the sample for the difference-in-differences analysis to communities located on Java. This is done in order as the Java communities (control and treatment) show similar trends prior to the eruptions and hence comply with the common trend assumption.

The main descriptive statistics are summarized in table 2.1 for the pre- and post-treatment period. The treatment group is younger and better educated than the control group. Reported household expenditure and living conditions are higher in the control group than in the treatment group during the pre-treatment period. The gap between the two groups widens over time. Pre-treatment, domestic violence and reported alcohol abuse rates are low and below one percent in both control and treatment. For the pre-treatment period, the treatment group's average well-being score is 2.6 whereas the control group's score is 3.1 on a scale where four is good, and one is bad (Please refer to the appendix for a description of scales).

The share of biological children is comparably high in both the control and treatment group at rates of 100 percent and 98 percent respectively. There is a shrinking but statistically significant difference between the two groups in this respect. The average family has been part of the program for around six years. The treatment group receives more interventions and has been part of the program for a shorter duration of time than the control group. However, both groups receive a similar level of psychosocial and childcare support. Levels of received support do not increase significantly over time.

The descriptive statistics in table 2.1 show that pre- and post-treatment levels of key variables are different. This does not violate the assumptions of a difference-in-differences estimation per se. However, one might question whether these groups, that are different from each other, really would have developed in the same way if it was not for the treatment. To check the robustness, we conduct a synthetic control analysis. We will also use communities outside of Java for this analysis as described in section 2.6.

Table 2.1.: Descriptive statistics by treatment status

	Pre-treatment			Post-treatment		
	Control	Treatment	Diff.	Control	Treatment	Diff.
Outcome variables:						
Share of fam. with domestic vio.	0.002	0.002	0.000	0.002	0.029	-0.026***
Log household expenditure per capita	13.610	13.555	0.054***	13.699	13.580	0.119***
Living conditions (1 to 4=good)	2.909	2.486	0.424***	2.976	2.509	0.467***
Emotional well-being (1 to 4=good)	3.097	2.611	0.486***	3.181	2.657	0.524***
Share families with alcohol/drug abuse	0.007	0.004	0.003	0.007	0.009	-0.002
Cargegiver characteristics:						
Caregiver education (1 to 4)	2.547	2.670	-0.123***	2.544	2.701	-0.157***
Age female caregivers	47.355	44.881	2.474*	47.858	41.563	6.295***
Age male caregivers	45.527	43.751	1.776***	47.020	44.181	2.839***
Share biological children	0.998	0.979	0.018***	0.998	0.986	0.012**
Support variables:						
Time since program admission in years	6.491	5.391	1.099***	6.555	5.484	1.070***
Support: Food	0.425	0.653	-0.228***	0.423	0.653	-0.230***
Support: Healthcare	0.547	0.798	-0.252***	0.544	0.798	-0.254***
Support: Material	0.195	0.304	-0.109***	0.194	0.304	-0.110***
Support: Economic	0.288	0.650	-0.362***	0.288	0.650	-0.362***
Support: Living conditions	0.440	0.578	-0.138***	0.437	0.573	-0.136***
Support: Psychosocial	0.783	0.788	-0.005	0.780	0.782	-0.002
Support: Childcare	0.757	0.763	-0.006	0.754	0.764	-0.010
Support: Legal	0.337	0.461	-0.124***	0.336	0.453	-0.117***

Notes: **Data:** Pre-treatment statistics based on pooled data from first three quarters of 2013; Post-treatment statistics based on pooled data from three last quarters of 2014; Age and education data might vary within-family over time based on responding primary caregiver; **Definition variables:** Education caregiver on scale 1 to 4 indicates 1) no formal education, 2) primary education, 3) lower secondary education, and 4) higher achievement than lower secondary education respectively; Support domains indicate support given by SOS in respective domains; See appendix section B.1 for exact other scales and definitions of variables.

2.4.2. Difference-in-differences identification strategy

I employ a difference-in-differences estimation. The method is frequently employed to study the ramifications of natural disasters (Caruso, 2017; Gignoux and Menéndez, 2016; Jensen, 2000; Shah and Steinberg, 2017). Specifically, and following Gignoux and Menéndez (2016), I am using a fixed effects model to account for unobserved between-family variation. I estimate the model for five outcomes. The main outcome variable is domestic violence. The four alternative outcomes are average household expenditure (based on macro-level data), household living conditions (based on micro-level data), emotional well-being and alcohol/drug abuse.[19]

Domestic violence

I estimate the following equation to identify the effect of two volcano eruptions on domestic violence prevalence.

$$y_{j,t} = \alpha_j + \beta_t + \gamma T_{j,t} + +\delta' X_{j,t} + \epsilon_{j,t} \quad (2.1)$$

, where $y_{j,t}$ describes a binary outcome variable for household j at time t that is one if domestic violence is reported and zero otherwise. α_j is a fixed effect accounting for time-invariant household attributes (A non-reported Hausman test supports the fixed effects approach). β_t is a time dummy-vector with dummies for all quarters to capture general time trends. $T_{j,t}$ is a dummy that is 1 for all affected areas post-treatment and 0 otherwise. γ is the coefficient of interest and measures the increase in domestic violence attributable to the event.

$X_{j,t}$ is a vector of family attributes that are determined in the pre-treatment period but time-variant. It includes the following variables: age-group of primary caregiver[20], number of children in the household as well as time since program admission and a binary indicator for whether a family received support. ϵ is an idiosyncratic, time-varying error term. Standard errors are clustered at a regional level (Results are robust to clustering at family level).

To describe the treatment effect over time, I report a series of dummies by interacting a binary *living in a treatment region* indicator with time fixed effects. The coefficients

19. Alcohol/drug abuse is reported at low rates, which might be either due to underreporting or low rates or both. Irrespective of this, the models will provide suggestive evidence towards the plausibility of the proposed channel – stress as a cause of domestic violence.

20. Age groups are created based on the decade of parental birth: 2000-10, 1990-99, 1980-1989, etc.

of this indicator vector capture the increased likelihood of development of domestic violence for individuals living in treated regions compared to individuals living in control regions. This also allows testing the common trend assumption. It shows that there is no pretreatment difference between control and treatment groups with regard to the dependent variable.

Alternative outcomes

For the alternative outcomes, I estimate the previously specified difference-in-differences equation 2.1 and replace outcome $y_{j,t}$ by the respective alternative outcome variable. I will estimate the baseline equation of the previous specification. Opposed to the previous specification, I employ a random-effects model as suggested by Hausman test results.

I estimate four outcomes. The first two are economic outcomes. A threat to livelihood would be expected to increase mental distress and act as a secondary stressor (Overstreet et al., 2011). The latter two outcomes are emotional well-being and alcohol/drug abuse. Emotional well-being is considered a proxy for the state of mental well-being of the sample. Previous research has documented the comorbidity of alcohol abuse and PTSD.[21]

The first outcome is household expenditure per capita in Indonesian Rupiah (IDR) as reported by the Indonesia Database for Policy and Economic Research.[22] This value does not vary by household but by region. The data per region can be found in appendix tables B.1 and B.2.

The second outcome is living conditions. Living conditions of a household are measured on a scale from one to four, with four indicating a positive outcome. The survey defines value four as "Family lives in conditions that are adequate, as per local standards (defined on community level in consultation with key stakeholders)", while the definition for value one is "Family lives in conditions that are below local standards, and are compromising the personal well-being of individual (and/or family)". The full scale is to be found in appendix section B.1.

21. McFarlane (1998) reviews studies on the association of alcohol abuse and PTSD. Sonne et al. (2009) present evidence on the sequence of the onset of PTSD and alcohol abuse respectively. Bueno and Henderson (2017) explore the association of IPV with alcohol abuse.

22. See Indonesia Database for Policy and Economic Research (INDO-DAPOER) by the World Bank Group.

The third outcome is emotional well-being. This variable takes on values from one to four, with four being *good*. The survey defines value four as "Care-giver is pro-active in addressing the situation of her/his family, and is emotionally stable, with a generally positive outlook". The definition for value one is "Care-giver is passive (not taking any action to address the situation of her/his family) and/or is emotionally unstable (showing signs of anger, irritability, aggression or depression)". Appendix section B.1 provides the full scale. One potential issue with this outcome is that the reporting care-giver sometimes changes over time. The threat to identification will be discussed later.

The fourth outcome is alcohol/drug abuse. Social workers report whether alcohol and/or drug use affect the family in a negative way. If either one of the two caregivers abuses alcohol or drugs, it is coded as 1, if none of the two abuse alcohol or drugs, it is coded as 0.

2.4.3. Heterogeneity analysis

Based on conversations with experts of the local situation and previous literature (Anastario, Shebab, and Lawry, 2009), I hypothesize that households with IDP status (internally displaced people) are particularly likely to develop domestic violence after the eruption. Households with IDP status had to resettle in the past due to natural disasters. These households were forced to migrate and often suffer on multiple dimensions in their new environment. In the setting of the sample, many families from the Huntap community had to move due to the 2010 major outbreak of Mount Merapi. A family is classified with IDP status if the family has IDP status at any given point in time during the observational period.

It is hypothesized that they will develop higher rates of domestic violence for three reasons. First, they are likely to suffer from a reduced livelihood even before treatment. This is because their previous sources of income have either been destroyed or the displacement forced them to create a new livelihood (While some studies on Katrina show that this might be beneficial for some, local experts suggest that overall living conditions suffer from displacement). These families are thereby likely to suffer from reduced income, reduced home size and a loss of their previous environment. Secondly, individuals with IDP status are expected to suffer from a loss of their social network and thereby social control. Thirdly, the volcano eruptions could act as traumatic reminders. Although I am not able to show this, I hypothesize that a *re-eruption* of Mount Merapi will cause significant trauma to this subpopulation.

Overstreet et al. (2011) summarizes literature that shows that even anticipation of a recurrent disaster threat can induce distress. I, therefore, estimate the baseline specification for both individuals with IDP and without IDP status in two separate models.

2.4.4. Proposed channel of causality

As outlined in the introduction and literature review sections, previous studies found that the psychological, social and economic consequences of natural disasters cause mental distress in affected populations. Outward, interpersonal violence and aggression are one type of reaction of humans exposed to distress. Worsened living conditions and lack of social networks and thereby social control act as important moderators.

This study follows this proposed causal chain. To test these predictions, I first estimate the impact of treatment on domestic violence. I then survey four alternative outcomes. I estimate the treatment effect on two economic outcomes. This is to test whether the affected population also suffers from economic loss. Secondly, I estimate the treatment effect on alcohol/drug abuse and emotional well-being. I argue that both are associated with mental distress (Alcohol abuse has also been associated with IPV).[23] A heterogeneity analysis with a subsample of individuals with IDP status estimates the treatment effect of people that lack a social network.

2.5. Results

2.5.1. Change in domestic violence prevalence

I estimate the change in the prevalence of domestic violence with a difference-in-differences model using fixed effects estimates (A Hausman test rejects the equivalence of random effects). Column 1 of table 2.2 presents the results of a fixed effects model without controls. Column two shows fixed effects estimates with a minimum set of controls, including quarter and village as well as caregiver age group dummies. Column three shows estimates from a fixed effects model with full controls, adding an indication of received types of NGO support, time since admission to SOS and number of children living in the household. Treatment is defined as living in an exposed community during and post the first volcano eruption. Results suggest an increase of approximately 2.2

23. Bech et al. (2003) discuss the relationship between mental distress and well-being. Bueno and Henderson (2017) explore the association of IPV with alcohol abuse.

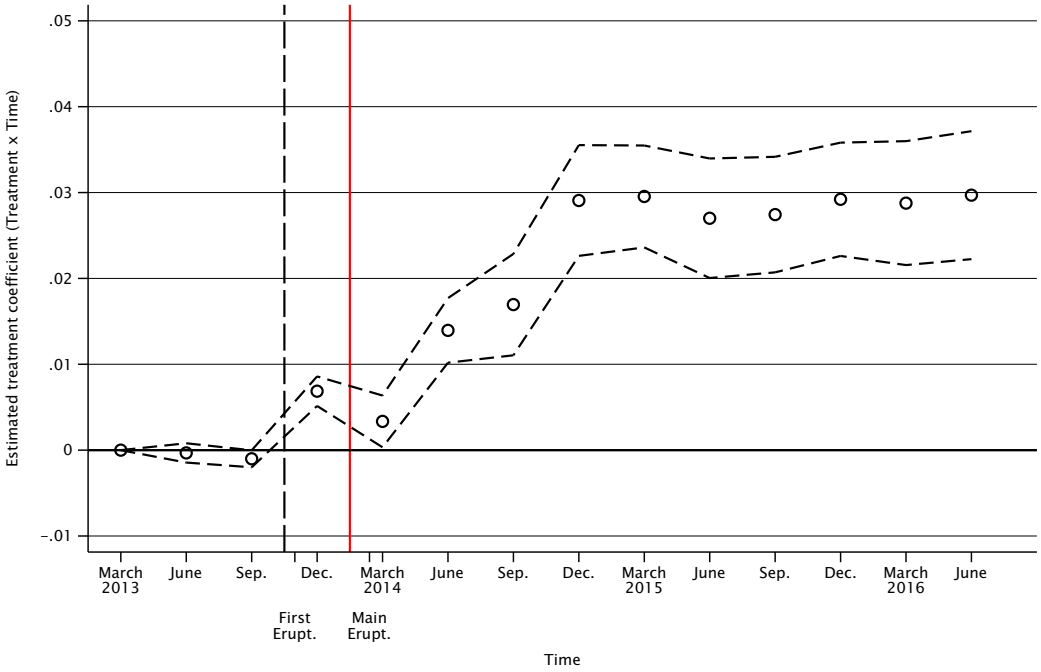
percentage points in domestic violence after the event of an earthquake. Compared to the low baseline level, this is a very meaningful increase. The effect size does vary slightly between estimation approaches and is found in random effects as well as fixed effects models. Figure 2.5 shows treatment coefficients over time.

Table 2.2.: Baseline model: effect of volcano eruptions experience on domestic violence

	Dependent variable: domestic violence		
	(1)	(2)	(3)
	No controls	Minimum controls	Full controls (Baseline model)
Treatment	0.030***	0.022***	0.022***
	(0.000)	(0.002)	(0.002)
Constant	0.002***	-0.077**	-0.217***
	(0.000)	(0.017)	(0.026)
Time dummies		✓	✓
Programme dummies		✓	✓
Age group dummies		✓	✓
Support dummies			✓
Time since admission and no. of children			✓
Observations	12,169	12,128	12,128
$Adj.R^2$	0.016	0.058	0.067
Clusters	4	4	4
Family level fixed effects	Yes	Yes	Yes
Standard erros clustered at regional level	Yes	Yes	Yes

Notes: Difference-in-differences estimate with family fixed-effects; **Dependent variable:** domestic violence (yes = 1/no = 0); Robust standard errors are clustered at regional level; ***/**/* indicate significance at the 1%/5%/10% level.

Figure 2.5.: Baseline model:
effect of volcano eruptions experience on domestic violence over time



Difference-in-differences estimate with family fixed-effects; ***Dependent variable:*** *domestic violence in treated vs. non-treated regions. Estimated coefficient: interaction of treatment and time. Controls: caregiver age group, regional control, number of children, support, time since admission; Robust standard errors in parentheses are clustered at regional level; First Eruption: marking the first eruption of Mount Merapi; Main Eruption: main eruptions of Mount Merapi and Mount Kelud; Lines indicate 95 percent confidence interval.*

2.5.2. Change in alternative outcomes

Changes in the four alternative outcomes support the hypothesis that the volcano eruptions lead to a decrease in material and emotional well-being. I find an adverse effect of treatment on average household expenditure while the individual living conditions do not change. The latter result might also be influenced by the fact that the sample receives economic interventions (See discussion in section 2.6). Importantly, the first outcome variable is measured on macro-level, while the latter is measured on household level. The treatment coefficient is negative for emotional well-being. It is positive and significant for alcohol/drug abuse negatively affecting families. The small effect of the latter outcome has to be interpreted in light of the overall low rates of alcohol/drug abuse in the overall sample.

Table 2.3.: Alternative outcomes: effect of volcano eruptions experience on alternative outcomes

	Alternative outcomes:			
	(1) Household expenditure	(2) Household living conditions	(3) Emotional well-being	(4) Alcohol/ drug abuse
Treatment	-0.054***	-0.059	-0.067***	0.005***
	(0.019)	(0.053)	(0.017)	(0.000)
Time s. Prog. Adm.	-0.002	0.010	0.004	0.001
	(0.001)	(0.013)	(0.019)	(0.001)
Constant	13.553***	0.000	0.000	-0.006
	(0.020)	(.)	(.)	(0.011)
Time dummies	✓	✓	✓	✓
Programme dummies	✓	✓	✓	✓
Age group dummies	✓	✓	✓	✓
No. of children	✓	✓	✓	✓
Observations	7,207	11,134	11,583	12,128
$Adj.R^2$				
Clusters	4	4	4	4
Family level random effects	Yes	Yes	Yes	Yes
Standard erros clustered at regional Level	Yes	Yes	Yes	Yes

Notes: Difference-in-differences estimate with family random-effects; **Dependent variable:** column 1: log of household expenditure in IDR (measured at regency-/city-level), column 2: living conditions (measured at micro level, scale of 1 to 4, where 4 is good), column 3: emotional well-being (scale of 1 to 4, where 4 is good), column 4: alcohol/drug abuse affects family (yes = 1/no = 0) respectively, see appendix section B.1 for exact scales and definitions of variables; Robust standard errors in parentheses are clustered at regional level; ***/**/* indicate significance at the 1%/5%/10% level.

2.5.3. Heterogeneity analysis: domestic violence prevalence among internally displaced people (IDP)

I find that households that are classified as internally displaced people (IDP) have a significantly larger chance of displaying domestic violence. The estimated effect size is roughly four to five times that of non-IDP households. This suggests that IDP families are at more substantial risk of developing domestic violence. Unobserved pretreatment heterogeneity of IDP vs. Non-IDP households is accounted for by fixed effects.

2.6. Identification concerns and robustness checks

Multiple concerns potentially threaten the validity of correct identification of the treatment effect. In the following section, I will discuss them and several robustness checks to address the concerns where possible.

Sample selection and attrition. The sample is particularly vulnerable to family breakdown compared to the general population. Its selection into a programme by SOS Children's Villages happens based on its assessed risk of family breakdown. While this renders the sample not representative of the Indonesian population, it is of particular relevance for policymakers. Due to its vulnerability, the sample is often the primary target group for emergency programs. Reactions of this group hence remain relevant for policy design. We only use individuals that have been present for the core of the observational period. Selective attrition from the program can upward bias the estimates. However, only 5.3 percent of all observations are omitted because they left the program before the end of the observational period. This indicates that selective attrition is not responsible for the observed effects. The remaining number of observations is comparably stable over time (see appendix table B.4). Throughout the years 2013 and 2014 the number of families remains at around 900 and gradually decreases thereafter.

Interviewer behavior. By definition, a natural disaster is a visible event. This event might affect interviewer behavior as well. The thoroughness by which interviews were conducted and the attention of social workers to signs of violence might have been influenced. This would upward bias reporting of domestic violence as discussed in Sekhri and Storeygard (2014). I cannot entirely reject this hypothesis. However, two arguments increase the likelihood that the results are not entirely driven by ob-

Table 2.4.: Heterogeneity analysis: effect of volcano eruptions experience on domestic violence prevalence by internally displaced people (IDP) status

	Dependent variable: domestic violence			
	(1) No IDP status no controls	(2) IDP status no controls	(3) No IDP status full controls	(4) IDP status full controls
Treatment	0.020***	0.091***	0.014***	0.092**
	(0.000)	(0.000)	(0.001)	(0.018)
Constant	0.002***	-0.003***	-0.150***	-0.520**
	(0.000)	(0.000)	(0.023)	(0.119)
Quarter dummies			✓	✓
Programme dummies			✓	✓
Age group dummies			✓	✓
Support dummies			✓	✓
Time since admission and no. of children			✓	✓
Observations	10,763	1,084	10,740	1,066
$Adj.R^2$	0.0091	0.062	0.042	0.23
Clusters	4	3	4	3
Family fixed effects	Yes	Yes	Yes	Yes
Standard erros clustered at regional level	Yes	Yes	Yes	Yes

Notes: Difference-in-differences estimate with family fixed-effects; **Dependent variable:** domestic violence (yes/no); sample split by previous IDP experience; Robust standard errors in parentheses are clustered at regional level; ***/**/* indicate significance at the 1%/5%/10% level.

servational sensitivity. First, the identified results in domestic violence fit the overall development of related variables. In addition to domestic violence, I also find increases in alcohol/drug abuse prevalence rates and reductions in emotional well-being. Furthermore, the full sample is restricted to families who have been admitted previously to the event. This implies that the same questionnaire has been filled for on average 5 to 6 years before the eruption happened. I expect that this long-term exposure to the organization and interviewers will increase the likelihood of accurate, trust-based reporting and detection of domestic violence.

Related to this is a concern associated with the main outcome variable. Domestic violence is indicated by a simple binary variable reflecting the overall presence of domestic violence within a family. This issue limits the overall depth of insight that can be gained from this study, apart from potential backlash identification issues. However, as Reading (2008) point out, there is a high correlation between different forms of domestic violence. From a humanitarian point of view, it is also relevant to detect and prevent violence in general, irrespective of the actor and victim.

Caregiver attributes. Most variables of interest are collected at the household level (for instance domestic violence, alcohol/drug abuse, living conditions). However, variables like caregiver age and emotional well-being are collected at the level of the caregiver. In some cases, the *primary caregiver* changes. It is hypothesized that this occurs (mostly) not due to death or divorce but dependent on which parent has been interviewed. To test the relevance of *switching caregivers*, the baseline model is only estimated for households that report a male primary caregiver (This does neither imply, that excluded households are led by female singles, nor that the primary male caregivers are single). I find a positive treatment effect suggesting an increase in domestic violence. This effect is smaller compared to the one found when using the full sample. See appendix table B.8 for the model.

Simultaneous treatment (eruption and support). A related concern is associated with the disaster and regular support programs that the sample has received. Interventions might bias outcomes. While other opinions exist, Cavallo et al. (2013) posit that post-disaster aid commonly covers only a minor share of real damages in affected communities. Nonetheless, the data on living conditions might be upward biased. That is, compared to other vulnerable groups, due to sample selection, the sample might benefit from more (disaster) support than otherwise equally vulnerable groups that do not receive support. It is not observed how families behaved that re-

ceived no support at all. It is also not possible to distinguish between disaster and regular support. However, even these outcomes are informative. It is of interest what happens in light of intervention, as scenarios with intervention are more common in most countries than non-intervention. Yet, this study cannot comment on whether this specific set of interventions increased domestic violence due to male backlash. Observing overall economic and family-level data suggests that domestic violence is associated with worsening economic conditions. However, the data do not provide information on relative income shares within families and the association between relative economic power and domestic violence.

Assignment to treatment group. Assignment to treatment is identified via maps provided by disaster aid organizations and additional reasoning. I do not have GPS location information on single households but identify via the location of the community the household is assigned to. The communities are then located as single points on the map and matched to disaster data. Certain villagers might live further away outside of the treatment area and thereby experience *less* treatment. This could downward bias the results. To check the robustness of the results to the exclusion of mildly affected treatment areas, the communities located in Kabupaten Gunungkidul are excluded. Results are robust to limiting the treatment sample to the resulting smaller sample size (see appendix table B.10).

Pre-treatment differences of control and treatment group. As argued, compared to the control group, the treatment group is on average less educated and suffers from lower living conditions, inter alia (see table 2.1). While overall household expenditure per capita is at a similar level (see figure 2.3), the variation in observables might cast doubt on the adequacy of the comparison of treatment and control group. As pointed out earlier, pre-treatment differences would only threaten difference-in-differences identification if they lead to a violation of the common trend assumption.

I run a synthetic control approach to test whether the post-treatment trends are rooted in the worse socioeconomic position of the treatment group compared to the control group. The approach has been pioneered by Abadie and Gardeazabal (2003) and Abadie, Diamond, and Hainmueller (2010) to evaluate the effects of policy interventions and civil conflict. Cavallo et al. (2013) were first to employ it for the identification of natural disaster effects. It creates a weighted average of control units to create a new synthetic control group, which then parallels the treatment group in its pretreatment features.

The results suggest that the negative post-treatment development of the treatment region is not founded on pretreatment differences of control and treatment groups. The synthesized control groups that have been created based on the pre-treatment trends of the treatment group develop more positively than the treatment group in all cases. Results are reported in the in appendix figure B.1.

2.7. Discussion

This study does not allow to comment on whether male backlash theories or bargaining models provide a better explanation for violence in the aftermath. However, a net increase in violence is observed. The overall rates of observed violence are in line with the previous literature. Domestic violence does not level off over time. This is in accordance with previous reviews that have shown a mixed picture with respect to the persistence of PTSD in the aftermath of disasters (Neria, Nandi, and Galea, 2008).

The overall results tie in with previous research that suggests an increase in PTSD following disasters and propose violence as one channel of how individuals react to existential stress and a challenged livelihood (Rezaeian, 2013). The four alternative outcomes support this theory. Previous hypotheses by Nilan et al. (2014) who point out the association between violence and the (self-perceived) male inability to satisfy female (economic) expectations are supported as well.

Families who have been displaced in the past are facing a particular risk of developing domestic violence. This longitudinal observation confirms previous post-disaster cross-sectional observations in a hurricane-related IDP context (Anastario, Shebab, and Lawry, 2009). It also alludes to the findings by Berkowitz (1993) who propose that a lack of social control will lead to higher rates of domestic violence (Curtis, Miller, and Berry (2000) discuss this in the specific context of natural disasters). Individuals with IDP status live outside of their previous social environment and are hence subject to less social control. The findings also confirm with Warsini et al. (2014), who documented higher rates of distress in survivor communities around Mount Merapi, which are the communities that individuals with IDP status come from. The high rates of domestic violence suggest that individuals with IDP status should be tracked closely following their displacement.

Appendix table B.6 supports the previous findings. The table offers a comparison of families in the treatment region by their domestic violence status. Families that show domestic violence in the aftermath of the disaster also report below average emotional well-being. We also find significantly higher rates of documented alcohol/drug abuse at a rate of 15 percent. These differences within the treatment group hint at the validity of the previously proposed channels.

2.8. Conclusion

This study's results suggest an association between natural disasters and domestic violence. In the aftermath of two volcano eruptions, I find an increase in domestic violence, a reduction in emotional well-being and a strong relationship between domestic violence and IDP status. A synthetic control approach supports the robustness of findings. The unique data thereby offer a longitudinal perspective on a particularly vulnerable group. As such, it is the first family level panel dataset from a developing country. Causal identification might suffer from multiple shortcomings in the data. While there is no final causal certainty that domestic violence has been caused by the stress induced by volcano eruptions, I argue that one has good reason to investigate the relationship further and act *preemptively* from a policy perspective. This could also result in providing (further) special assistance to families with IDP status. Findings should encourage the collection of more evidence and a potential increase in sensitivity of disaster aid workers for a prevalence of domestic violence in post-disaster areas.

3. Birth order effects and educational achievement in low- and middle-income countries

3.1. Introduction

The importance of education for individual prosperity (Hartog and Oosterbeek, 1998; Rosenzweig, 1995) and economic growth (Hanushek and Woessmann, 2015; Mankiw, Romer, and Weil, 1992) is undisputed. However, what determines educational achievement itself? There has been ongoing debate on the influence of individual determinants of educational achievement, such as individual aptitude (Rowe, Vesterdal, and Rodgers, 1998), parental attributes (Desforges and Abouchaar, 2003; Harris, 2008), teacher quality (Darling-Hammond, 2000), and their collective dynamics (Belley and Lochner, 2007; Cameron and Heckman, 2001; Heckman, Lochner, and Todd, 2006). Part of this debate centers around the impact of birth order effects. A rich body of research documents that firstborns in high-income countries tend to fare better across a range of life outcomes, like IQ, height, and all-cause mortality. Studies show that the same holds true for educational achievement. For example, Black, Devereux, and Salvanes (2005) find that secondborns receive on average four months less of education than their first-born siblings.

While the general existence of birth order effects in high-income countries is widely acknowledged, there is substantial debate over *why* birth order effects exist (Barclay and Myrskylä, 2014; Kristensen and Bjerkedal, 2007), and what might explain heterogeneity in findings in low- and middle-income countries (De Haan, Plug, and Rosero, 2014). Some studies with data from low- and middle-income countries demonstrate reversed birth order effects, that is better outcomes for later-born children. For example, evidence from South America (De Haan, Plug, and Rosero, 2014; Emerson and Souza, 2008; Lafortune and Lee, 2014) and Asia (Ejrnæs and Portner, 2004) proposes positive birth order effects, implying better educational outcomes for later-born children. As De Haan, Plug, and Rosero (2014) point out, there is a need for more evidence from low- and middle-income countries.

This study addresses these gaps in the literature by employing novel and unique data. It is the first one to provide evidence from a broad set of low- and middle-income countries across multiple continents. Overall results indicate that birth order effects in low- and middle-income countries are consistent with those in high-income countries. Higher educational achievement for firstborns is identified when estimating the relationship for the full sample. A second analysis of heterogeneity suggests reasons for why previous studies might have found contradictory results. Three sources of het-

erogeneity are surveyed. Extreme hardship, parental gender preferences, and tutoring between siblings are identified as moderators of birth order effects.

I use a novel dataset covering 26,898 observations of 4,362 biologically related siblings living in long-term alternative care families in 54 countries. The data are provided by a global childcare NGO which places children and young adults in alternative care families. In these families, children live together with their biological siblings, one non-biological mother and up to eight non-biological siblings.[1] Individuals living in these families have parents who passed away or who are no longer able to take care of them. This dataset is particularly valuable as it spans across several continents and permits the observation of family structures where biological and social birth order coexist.

Findings suggest that sibships that have suffered extreme economic or emotional hardship (for instance sexual abuse, domestic violence) show attenuated birth order effects compared to other sibships. This is compatible with previous evidence indicating effect diminishment and reversal for households of low socioeconomic status in high-income countries and reversed birth order effects in low- and middle-income countries. Individual hardship *within* a society seems to be as relevant as differences in development *between* societies. Gender-specific effects are identified for Asia, where the firstborn advantage is significantly smaller for girls, compared to boys, suggesting parental gender preferences. These effects are mainly driven by data from India, a country with a widespread preference for male offspring.

Intra-family comparisons of biologically unrelated children of the same biological birth order provide suggestive evidence that supports the existence of tutoring effects between unrelated siblings. Holding biological birth order constant, I find superior outcomes for older children ranked *higher* in their alternative care family. I propose more tutoring opportunities as a potential explanation. This evidence is only suggestive as large standard errors prevent statements on statistically significant differences between children of the same biological birth order. The finding is consistent with the confluence model – one possible explanation for birth order effects. The confluence model attributes birth order effects to changing dynamics of social interaction within the family; of which tutoring between children is one element.

1. The NGO takes full custody of the children admitted into villages on a long-term basis. The family-like care approach is one type of alternative care. It can be thought of as a hybrid of a foster model and adoption. It emulates a family environment. It will be referred to by the general term of alternative care.

These findings advance the debates on determinants of educational achievement and the formation of human capital in low- and middle-income countries. They also suggest reasons for how and when intra-family differences emerge. Some evidence points to the possible existence of tutoring effects. Larger (alternative care) families could particularly benefit from exploring tutoring as a way to let children grow personally and intellectually. The results can inform policy making and development interventions by helping to prioritize individuals in highest need.

The structure of the remainder of this paper is as follows. First, I will provide an overview of the related literature. I focus on existing theories to explain birth order effects and summarize previous empirical evidence, with a particular focus on low- and middle-income countries. I will then describe the data and the estimation strategy. Subsequently, I will present the results on *classic* birth order effects between biological siblings for the full sample. I will then split the sample by experience of hardship and estimate this relationship again. This will be followed by an analysis on the interaction of gender and firstborn status with a regional split. Subsequently, I will show the results of an analysis comparing biologically unrelated siblings of the same biological birth order. I will conclude by relating results to previous literature, discussing their external as well as internal validity and deriving policy implications.

3.2. Theory and empirical evidence on birth order effects

This section presents theories to explain birth order effects (section 3.2.1), empirical evidence in general (section 3.2.2) and findings from low- and middle-income countries in specific (section 3.2.2). An emphasis will be on studies that used samples similar to this one. Critics have challenged the existence of birth order effects and the methods employed to analyze them. A selection of their objections will be covered as well.

I will conclude that first approaches to elucidate the mixed evidence in low- and middle-income countries exist. A lack of multi-country studies and the ambiguous empirical evidence call for studies that allow to further compare outcomes across countries and elaborate on the direction of birth order effects in low- and middle-income countries (De Haan, Plug, and Rosero, 2014).

3.2.1. Theories to explain birth order effects

This section presents three main theories that explain why birth order effects surface: (i) resource dilution theory, (ii) confluence theory, and (iii) immunoreactive theory. (i) Resource dilution theory attributes superior outcomes for earlier-born children to the gradual dilution of parental resources with every additional child being born into the family. (ii) Confluence theory posits that within-family social dynamics are an important factor to explain birth order effects. Throughout their lives, firstborns are exposed to an environment of higher average intellectual maturity, compared to later-borns. (iii) Immunoreactive theory attributes these effects to biological causes, namely mothers' biological reactions to the male fetus. It is important to note that these three theories could hold true in parallel: the verification of one will not falsify the other.

(i) Resource dilution theory. Resource dilution theory is based on the assumption that parents' resources, such as attention and financial means, are divided among children living in a household. Hence, they dilute with every additional child.[2] The firstborn will benefit from access to the highest average amount of resources (Blake, 1981; Downey, 2001). This effect is amplified by the fact that investments during early childhood are expected to be more productive than investments later on in life (Cunha and Heckman, 2009). Based on American Time Use Survey data, Price (2008) finds that first-born children experience 20 to 30 minutes of average additional quality time per day compared to secondborns at the same age in similar families due to an equal split of parental time amongst siblings. Also, active discrimination by parents can augment birth order effects (Findings will be presented in the empirical literature review, section 3.2.2). Downey (2001) propose that a strong argument in favor of resource dilution theory is the low likelihood of the null hypothesis being true. It appears unlikely that neither parental resources nor the number of siblings in a household carries any impact on a child's development.

(ii) Confluence theory. The confluence model is based on the conceptualization of the family as an intellectual environment that follows complex dynamics. It has been first described by Zajonc and Markus (1975). The model is based on the assumption that the child's intellectual development is partly driven by the dynamics of its social environment's average intellectual maturity. The authors argue that firstborns bene-

2. This alludes to Becker and Lewis (1973), who described the quality vs. quantity trade-off that parents are facing regarding their offspring.

fit, all other things equal, from a household age which is on average higher compared to that of younger siblings. Siblings are understood as important peers, and child development is analyzed on the grounds of this peer structure and changes thereof. Moreover, older siblings are expected to benefit from a tutoring effect, namely by reinforcing their skills via teaching them to younger siblings. An important aspect is that this firstborn advantage is dynamic and not a linear function of birth order. As such, it depends on the individual's and their siblings' age. The model suggests a positive association between birth order and achievement from a crossover age of 11 +/-2 onwards (Zajonc and Mullally, 1997; Zajonc and Sulloway, 2007). The crossover age describes the age at which earlier-born siblings start to particularly benefit from tutoring their later-born siblings, resulting in more pronounced negative birth order effects. The authors propose that this also explains conflicting findings in other studies dependent on the respective sample's average age. An additional dimension associated with variation in effect sizes is that of age spacing between siblings (Zajonc, 1976). Confluence theory has received substantial criticism. Galbraith (1982) point out that the model is unable to explain observed birth order effects in French data. For confluence theory to explain birth order effects, spacing would be required to be substantially longer than it is in reality. Furthermore, Galbraith (1982) posit that the model is not able to explain the finding that in France, positive birth order effects (improved outcomes for later siblings) occur in conjunction with negative family size effects (worse outcomes for larger families). Retherford and Sewell (1991) show that the confluence model provides a poor fit if within-family data is used (as opposed to between-family estimation) – a finding seconded by others (Rodgers, 1984; Wichman, Rodgers, and Maccallum, 2006).[3]

(iii) Immunoreactive theory. The final and third type of explanation for birth order effects is based on immunoreactive theory (IMRT) (Gualtieri and Hicks, 1985). Immunoreactive theory hypothesizes that maternal antibody reactions grow stronger with increasing birth order. The authors argue that the male fetus with its particular, *male* genetic attributes causes maternal antibody attacks on the fetal brain. These, in turn, affect male fitness negatively. The reactions are expected to grow stronger for later-born male children with the female body *learning* from the first male fetus and developing stronger anti-body attacks over time. Findings of Kristensen and Bjerkedal (2007) and Barclay (2015) cast doubt on whether biological explanations in general and IMRT-based approaches in specific help to explain birth order effects. In their

3. Retherford and Sewell (1991) are also not able to replicate the findings with a between-family estimation and a representative sample for the United States of America.

landmark study, Kristensen and Bjerkedal (2007) show that second-born men whose biologically older brothers passed away before reaching the age of one resemble first-borns in their achievements. The argument is that the passing of their older brothers renders them biological secondborns but social firstborns.

Apart from theory-specific criticism, birth order research has been subject to general objections that question the validity of employed research methods to study birth order effects. *Spurious association theories* attribute IQ related birth order effects to the analysis of between-family data (Kanazawa, 2012). However, as shown by Sulloway (2007), these theories are unable to provide explanations for various phenomena, such as the observable distinction between biological and functional causes of birth order effects (Kristensen and Bjerkedal, 2007) or the role of age spacing in explaining birth order differences (Buckles and Munnich, 2012). Another objection is based on the observation that the decision to receive a second child is not independent of the outcome of the first pregnancy. This endogenous relationship could serve as an alternative explanation for the firstborn advantage (Ejrnæs and Portner, 2004). However, Bagger et al. (2013) and Black, Devereux, and Salvanes (2010) use instruments to account for endogeneity in fertility decision-making and show that birth order effects are robust to this specification.

In light of the findings by Kristensen and Bjerkedal (2007), I consider two theories as possible explanations for birth order effects in the context of this sample: resource dilution theory and confluence theory.

3.2.2. Empirical evidence on birth order effects

In the following, I provide a selection of studies that have particular relevance to this one, focusing on studies associating differences in educational achievement with birth order effects. Furthermore, I will summarize findings of studies that employ data that share peculiarities comparable to this sample, namely a non-biological family setting, above average sibship size, and a low- and middle-income country environment. I do not discuss criticism of empirical birth order effects research in-depth. Schooler (1972), Galbraith (1982), and Kanazawa (2012) offer some of the main arguments that have been brought forward by critics.

Ability and educational achievement in biological families. Belmont and Marolla (1973), Black, Devereux, and Salvanes (2011), Calimeris and Peters (2017),

and Kristensen and Bjerkedal (2007) explore the relationship between birth order, ability and educational achievement. Black, Devereux, and Salvanes (2011) estimate the IQs of first-born children to be three percent higher than those of second-born children. Confirming evidence has been recognized in regard to negative educational outcomes for higher birth order ranks (Barclay, 2015; Black, Devereux, and Salvanes, 2005; Haan, 2010; Kantarevic and Mechoulan, 2006). This has been established for biological siblings (Black, Devereux, and Salvanes, 2005; Härkönen, 2014) and adoptive sibship groups alike (Barclay, 2015).

Ability and educational achievement in alternative care families. In their study on children growing up with adoptive parents, Björklund, Lindahl, and Plug (2006) argue that both adoptive and biological families carry features that are important for outcomes but that the influence of biological and adoptive parents varies by type of the observed determinant (for instance parental education, gender, income). Beckett et al. (2006) and Lindblad, Hjern, and Vinnerljung (2003) associate positive cognitive development with an earlier child age at adoption. While Hjern, Lindblad, and Vinnerljung (2002) identify substantial differences in psychosocial life outcomes of adopted children compared to non-adopted children, Barclay (2015) can replicate birth order effects in alternative care families living in Sweden. Finally, there is an overlap between adoption studies and studies on peer effects because sibling interactions in large sibship groups are comparable to peer interactions.[4] Scholars have been able to show that the behavior, societal background and educational performance of peers can change individual attainment in both directions (Ammermueller and Pischke, 2009; Sacerdote, 2014). Sacerdote (2011) estimates the influence of peers on par with other important determinants, such as class size (Biddle and Berliner, 2002) or teacher quality (Darling-Hammond, 2000). While most peer research focuses on classmates, similar effects have also been estimated for cohorts living in close-knit settings comparable to SOS Children's Villages (Carrell, Fullerton, and West, 2009).

Sibship size and spacing. Multiple studies find sibship size to be negatively associated with educational attainment (Black, Devereux, and Salvanes, 2010; Blake, 1981). Rodgers et al. (2000) were the first to show that large families do not imply lower ability per se. Black, Devereux, and Salvanes (2005) found that educational attainment gaps between larger and smaller families disappeared once birth order ef-

4. In SOS Children's Villages, children are exposed to biologically unrelated siblings inside their new family's house and other children living in the same village but in separate houses. Both groups form a peer environment.

fects were taken into account. According to the authors, it is fertility decisions of low socioeconomic status parents that cause mediocre outcomes for larger families, ceteris paribus. However, the same authors found mixed evidence in a later publication (Black, Devereux, and Salvanes, 2010). Other scholars find a reversal (Davis, Cahan, and Bashi, 1977) or total disappearance of birth order effects (Kanazawa, 2012) if family size is considered. Building on confluence theory, Zajonc and Mullally (1997) predict a reversal of effects between first and secondborns at a turnover age of 11 +/- 2 years, and for final life outcomes, a persistent advantage for firstborns and a persistent disadvantage for lastborns, irrespective of sibship size.

Controlling for sibship size, the spacing of siblings continues to provide a potential force influencing birth order effects. Zajonc (1976) predicts varying birth order effects dependent on age spacing of siblings – with a generally positive association of individual outcomes and spacing. Buckles and Munnich (2012) find a positive effect of spacing on older siblings only. The evidence remains mixed as other studies propose a null effect of spacing (Belmont, Stein, and Zybert, 1978; Black, Devereux, and Salvanes, 2010).

Birth order effects in low- and middle-income countries and low-income families in high-income countries

Evidence in low- and middle-income countries is heterogeneous (Mechoulan and Wolff, 2015). Studies from South America (De Haan, Plug, and Rosero, 2014; Emerson and Souza, 2008; Lafortune and Lee, 2014) and Asia (Ejrnæs and Portner, 2004) find positive birth order effects, suggesting better educational outcomes for later-born children. Conversely, Moshoeshoe (2016), Hammitt, Liu, and Tsou (2012) and Calimeris and Peters (2017) find negative birth order effects in Lesotho, Taiwan and Indonesia respectively.

Amongst others, three potential explanations exist for these ambiguous findings: variation in family resources, child labor and parental discrimination/selective investment or a combination thereof. All three factors differ regarding their total and relative levels in low- and middle-income countries compared to high-income countries. Moshoeshoe (2016) proposes family wealth as a determinant of within-country variation of birth order effects in low- and middle-income countries. Lafortune and Lee (2014) find that positive birth order effects are attenuated by increasing family assets in Mexico. In low- and middle-income countries, wealth can be thought of as a proxy for the neces-

sity to send earlier-born children off to work early, potentially leading to their sub-par educational achievement. Where present, child labor effects confirm this hypothesis. Emerson and Souza (2008) attribute positive birth order effects in low- and middle-income countries to higher rates of child labor among earlier-born children. According to Emerson and Souza (2008) and Edmonds (2006), earlier-born children tend to engage more in child labor compared to their younger siblings and receive less education in consequence.

Besides child labor and wealth, variation in parental investments will moderate birth order effects, leading to birth order variation by country and culture. Jayachandran and Kuziemko (2011) show that mothers in India engage in shorter breastfeeding spells if the first-born child is female. The authors offer cultural gender preferences as an explanation. This has been confirmed by Fors and Lindskog (2017), who find negative birth order effects in India and inferior outcomes for first-born girls compared to first-born boys. So do Lafortune and Lee (2014) for South Korea. Ejrnæs and Portner (2004) treat fertility decisions as endogenous and find positive birth order effects on the Philippines. Multiple authors document other forms of discrimination. Mechoulan and Wolff (2015) observe parental discrimination with respect to the allocation of financial resources and gifts. Hotz and Pantano (2015) find weaker sanctioning towards later-born siblings, who do not meet the expectations of their parents. De Haan, Plug, and Rosero (2014) document less parental quality time for earlier-born children in Ecuador.

The first study that has used multi-country data from low- and middle-income countries is that of Tenikue and Verheyden (2010).[5] The authors explore wealth and child labor as potential explanations for birth order patterns. However, their data do not indicate whether any young adults have left the household already. Arguably, this threatens the identification of birth order effects. Moshoeshoe (2016) shows how this factor is likely to bias results.

In high-income countries, multiple studies confirmed that low-income families tend to show the most pronounced positive birth order effects, implying better outcomes for later-born siblings (De Haan, Plug, and Rosero, 2014; Lafortune and Lee, 2014). For families with low income in high-income countries, Bonesrønning and Massih (2011) and Lafortune and Lee (2014) find smaller and reversed birth order effects. In South Korea, the US and Mexico, Lafortune and Lee (2014) document more years of educa-

5. Lafortune and Lee (2014) use a multi-country dataset from one middle-income country (Mexico) and two high-income countries (South Korea and the United States of America).

tion for firstborns in families with higher educated fathers, while the opposite holds true for families with fathers without formal education. Bonesrønning and Massih (2011) find some evidence for more pronounced birth order effects in families with highly educated mothers.

3.3. Empirical strategy

3.3.1. Data

The data are provided by SOS Children's Villages, an international NGO. The NGO operates as a federation with national organizations in 133 territories and headquarters in Innsbruck, Austria. Its budget exceeded one billion US Dollars in 2015. Individual donors, corporations as well as governments and other institutions are the main funders. The NGO runs two main childcare programs in addition to a multitude of services along the themes of education, health care, and emergency response.

The two care programs combined serve a total of 553,600 beneficiaries worldwide (SOS Children's Villages International, 2016). The larger program by the number of beneficiaries is the *family strengthening program* with 467,400 beneficiaries. It provides interventions to improve childcare and prevent the breakdown of biological families. The second program, called *family based care* provides care to 86,200 beneficiaries who have lost parental care or whose parents are no longer able to take care of them. Children live in so-called villages. Their educational achievement is the subject of this study.

Before admission, the organization determines a child's need for alternative care by means of a standardized process. Children are admitted on the grounds of four potential reasons for admission: loss of either one or both parents, the inability of caregivers to take care, the referral from another care placement or child abandonment. Table C.12 in the appendix displays the share of each reason for admission by region. After a positive admission decision, SOS admits children to a village, where they will become part of a family.

Families typically consist of one caregiver and up to ten children. The vast majority of caregivers is female and working full time for SOS Children's Villages in the functional equivalent of a biological mother. Assignment to a specific family depends on child-mother interaction during an initial trial period. Importantly, biological siblings are never split up between two families. About ten families form one village. Villages provide additional infrastructures such as schools, sports facilities, and the village's head office. Whether children have access to a school on site varies by location. Once children are grown up and self-sustaining, they will move out, and new children enter the family.

Data for this study are obtained from a central program database which provides data for the vast majority of all children living in SOS Children's Villages in low- and middle-income countries. Data are collected at village level monthly. A committee of social workers, head of the village and other representatives of the respective national organization is responsible for the collection of data, and for tracking educational achievement. Commonly, data collection is paper-based with subsequent data entry into a computer. Alternative care mothers can provide input, but a committee makes the final decision on performance.[6]

The data provide the following information on all individuals: A codified surname and first name, an individual's gender and date of birth, their reason and date of admission. Besides, detailed reason for admission is available for some *reasons for admission*. The database offers information on the house, village, and home country of an individual. Educational performance is tracked on a scale from one to four, where four is a good outcome. The variable will be described in more detail in the next section. Appendix section C.1 provides the exact wording. According to related literature, it will be referred to by *educational achievement*. For caregivers, full names are available. The study is based on panel data obtained from collection cycles from September 2014 through September 2016. For this timespan, data is retrieved at the end of each quarter. Applying all exclusion restrictions results in 26,898 observations, from 4,362 individuals, living in 54 countries. The motivation to use multiple observations per person is that the primary independent variable of interest (biological birth order) is time-invariant and secondly, that multiple outcome observations increase statistical power.

All figures presented in the following refer to the sample after application of exclusion restrictions. Appendix table C.10 provides a detailed overview of the number of observations per country, and grading distributions within countries after application of exclusion restrictions. The average age of children in the sample is 12 with a standard deviation of 3.5 years. The final sample contains a slightly higher percentage of female children (52 percent). On average, children have been admitted at the age of 6.4 with a standard deviation of 3.1 years. The highest number of observations included in the sample is Asian (17,577), followed by children from Latin America (5,121) and Africa (4,200).

6. Not all villages can offer a full committee. In these cases, a social worker will carry out data collection and data entry.

The average number of children analyzed per biological sibship is 2.4 (SD = 0.7). The equivalent figure for alternative care sibships is 9.5 (SD = 2.2). On average, children have spent 46 percent of their life in SOS care.[7] Further descriptive statistics for first-, second-, and thirdborns with a split by region can be found in appendix tables C.4, C.5, C.6 respectively. Unfortunately, the data do not provide further parental background information, that goes beyond the child's reason for admission. Parental information such as employment status or age would be desirable.

Table 3.1.: Descriptive statistics for children by birth order rank

	Firstborns		Secondborns		Third- or higherborns		Full sample	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Outcome variable:								
Educational achievement	3.02	0.73	3.02	0.72	3.07	0.71	3.03	0.73
Individual characteristics:								
Age	13.65	3.17	11.16	3.12	9.75	2.76	12.02	3.46
Age at entry	7.77	3.06	5.54	2.75	4.54	2.46	6.35	3.13
Gender = female	0.56	0.50	0.48	0.50	0.51	0.50	0.52	0.50
No. bio. siblings	2.30	0.57	2.30	0.57	3.22	0.42	2.44	0.66
No. all siblings	9.50	2.17	9.50	2.17	9.61	2.18	9.52	2.17
Lifeshare spent in SOS care	0.42	0.22	0.49	0.24	0.52	0.25	0.46	0.23
Reason for admission:								
Abandonment	0.13	0.34	0.14	0.35	0.13	0.34	0.14	0.34
Death of parents	0.63	0.48	0.63	0.48	0.63	0.48	0.63	0.48
Referral	0.02	0.14	0.02	0.13	0.02	0.13	0.02	0.14
Inability caregiver	0.22	0.41	0.21	0.41	0.22	0.41	0.21	0.41
Observations	11,693		11,693		2,871		26,898	

Notes: Data describes final sample after application of sampling restrictions (see appendix); Variable definitions: Age at entry = age at which the child has been admitted; Gender = Variable that is one for girls and zero for boys; No. siblings bio. only = number of biological siblings; No. all siblings = number of all siblings in alternative care family; Lifeshare spent in SOS care = Number of years in SOS alternative care divided by age; Referral = referral from another care placement; Inability caregiver = inability of caregiver to take care of child.

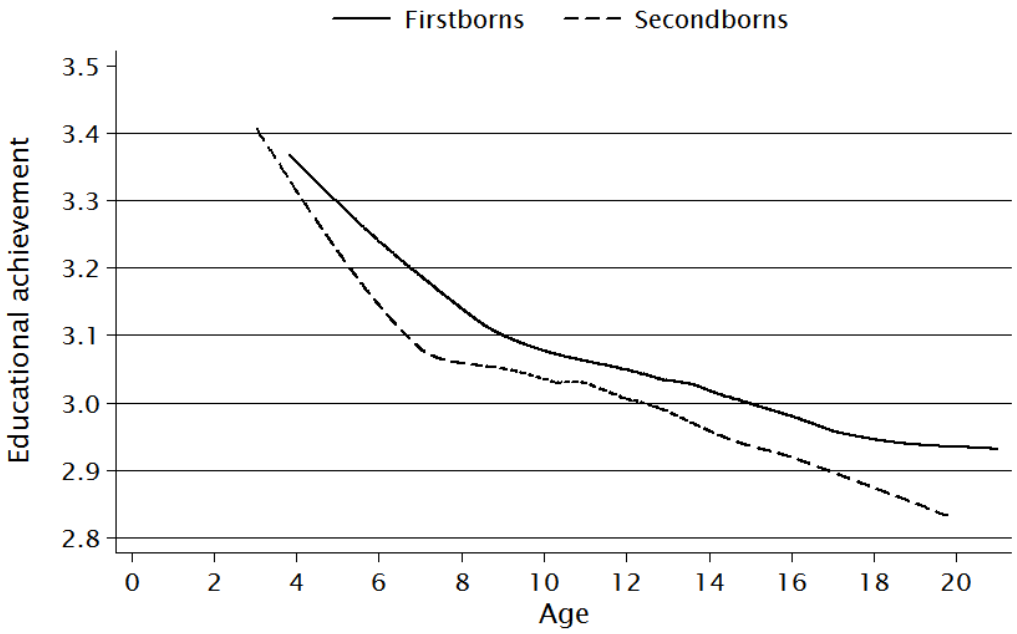
Siblings are assigned to biological sibships based on the codified identifier that the database operator assigned per surname per household. The accuracy of the matching of identified sibships is assured through several quality checks. First, the date of entry and reason for admission must be identical across all siblings. If both variables are not identical across the sibship, all those sibships are excluded, who either carry the same surname as their caregiver or that show contradicting reasons for admission within a sibship. Individuals who share a surname with their caregiver are excluded

7. The percentage figure refers to time since admission to any village divided by age.

because abandoned children and those who were referred from another care placement are sometimes given the name of their primary SOS caregiver. This leads to a shared surname amongst non-biological siblings. The process for the exclusion of all cases of contradicting reasons for admission is further detailed in appendix section C.2.

Figure 3.1 displays the central relationship of interest analyzed in this paper. The figure shows the average educational achievement of children based on age and split by birth order rank. Firstborns show consistently higher average educational achievement by age than secondborns.

Figure 3.1.: Lowess smoothing of educational achievement vs. age for first- and secondborns



Notes: **Dependent variable:** average educational achievement by age; Lowess smoothing.

3.3.2. Models

Estimation of baseline model. In the *baseline model*, educational achievement is regressed on a biological birth order vector and a set of controls. A pooled OLS estimation approach is used with biological sibship dummies and controls for individual, biological sibship and alternative care family features as shown in baseline equation 3.1. (Results based on cross section estimations are presented in the appendix.)

$$Edu_{i,j,k,t} = \beta_0 + \beta_1 BIRTHORDER_i + \gamma_j + \delta_t + \beta_2 X_i + \beta_3 XT_{i,t} + \rho_1 AltCareFam_{k,t} + \varepsilon_{i,t} \tag{3.1}$$

The dependent variable $Edu_{i,j,k,t}$ is the educational achievement of child i in sibship j in alternative care family k at time t. The variable can take on four possible values from 1 to 4, where 4 is good. The grading scheme uses the four values to indicate poor, below average, satisfactory and outstanding performance respectively. The appendix section C.1 presents the exact wording of all four values.

The birth order vector $BIRTHORDER_i$ contains the variables of interest and includes birth order dummies for all birth order ranks but the first one, which is omitted. For sibships of three siblings, this vector contains dummies for being second-born and being third-born. For estimations of biological families of more than three members, this includes dummies for being second-born and being third-born or of higher rank.

γ_j is a biological sibship dummy for sibship j that is time-invariant and captures unobserved biological family characteristics. To account for potential grading trends over time, I introduce time dummies, denoted by δ_t for each quarter. The vector X_i accounts for an individual's time-invariant attributes: gender and reason for admission, while $XT_{i,t}$ denotes time-variant attributes: age and relative lifeshare spent inside SOS care. Age is coded with a vector of dummies containing one dummy for each possible age from 3 through 21. Lifeshare spent in SOS is calculated as the number of years in SOS alternative care divided by age. Finally, the $AltCareFam_{k,t}$ variable denotes the number of children in an individual's alternative care family, both biologically related and unrelated. The baseline model will be estimated for all biological sibship groups of $N < 5$ members. An alternative specification will estimate it for sibship pairs of $N = 2$. Standard errors are clustered on the level of the individual in all models to account for the serial correlation of errors within an individual over time.

In addition to the baseline model, I estimate two derivative models to study hetero-

geneity in birth order effects. The first model divides the full sample by experience of economic or emotional hardship. The second model introduces a gender interaction effect and divides the sample by region. In a third derivative model, I estimate a factorial family model that predicts the relationship between biologically unrelated siblings.

Estimation of hardship model. The hardship model is based on the baseline model equation (1). I split the sample by experience of extreme hardship. The main question to answer is whether birth order effects persist for children who have suffered extreme hardship. While all children in this sample have suffered some form of hardship, some children have experienced extreme hardship. I assess this experience of extreme hardship based on the *detailed reason for admission* that has been provided for the majority of children upon admission.[8] The list of all detailed reasons for admission is presented in appendix table C.13.

Two dummies indicate extreme economic and extreme emotional hardship respectively. Experience of extreme hardship is coded as 0 by default in both cases.[9] If the detailed reason of admission is indicative of either type of hardship, the respective hardship dummy (economic or emotional) is coded as 1. For the model, the sample is first divided into children who have experienced extreme hardship during their childhood and those who have not according to the detailed reason for admission. The group without extreme hardship experience is divided into children whose parents passed away and all others. For the group with extreme hardship experience, I run three models. The first model includes all children with emotional hardship experience. The second model includes all children with economic hardship experience. The third model includes all children with either of the two types of experience. Importantly, children can have suffered both economic and emotional hardship.

Estimation of gender split model. The gender split model is based on the baseline model equation (1). It additionally interacts gender with being first-born. I estimate it for the full sample and by region. The regional breakdown is motivated by previous literature finding a strong preference for male offspring in Asia in general and India in

8. The general reason for admission groups are mutually exclusive, that is children will belong to either of six groups. The detailed reasons are not. There can be multiple detailed reasons for admission.

9. The default coding of 0 is based on the hypothesis that a missing *detailed reason for admission* value indicates that another reason for admission other than the stated one is not given. One might argue that a missing value is a missing assessment rather than the absence of hardship. In appendix table C.21 an alternative coding that codes all missing values as missing and drops these observations is presented.

specific.

Estimation of the factorial family model. The third derivative model is used to study educational achievement differences of biologically unrelated siblings, who are living together in alternative care families. Specifically, I investigate whether it matters to be the first, second or third oldest child of a specific biological birth order rank in an alternative care family. This is coded via tuples that identify a children's position in their biological and alternative care family.

I identify all potential positions of an individual within both its initial, biological and its new, alternative care family with the tuple (b,a). The first tuple entry (b) is equivalent to an individual's biological birth order rank. This biological birth order rank reflects an individual's position in its initial, biological family before admission to alternative care. All biologically first-born children in the sample are assigned (1,a), all secondborns are assigned (2,a), and all individuals with biological birth order rank three or higher are assigned (3,a). The second tuple entry (a) reflects an individual's relative position in an alternative care family, given its biological birth order rank. There are multiple biologically firstborns in an alternative care family, as there are multiple secondborns, and so forth. Tuple entry (a) denotes the rank of an individual *within* their alternative care family, *within* the group of children of the same biological birth order rank. Hence, the oldest biologically firstborn in an alternative care family is assigned rank 1, resulting in the tuple (1,1). The second oldest biologically firstborn is assigned rank (1,2). All biologically firstborns who are younger than the second oldest biologically firstborn are assigned rank (1,3).

Assume that there is a small alternative care family with four children {X,x,Y,y}. Within this family, {X,x} and {Y,y} form two biological sibships. The age structure of {X,x,Y,y} in years is assumed to be {20,15,10,5} respectively. X will be assigned tuple (1,1) for being biologically firstborn and being the oldest firstborn in its alternative care family. Individual x is assigned (2,1) for being a biological secondborn but being the oldest one in its alternative care family. Y is assigned tuple (1,2), and y is assigned (2,2). As both tuple values (b) and (a) can take on values from 1 to 3 only, this leads to 9 potential factorial combinations ranging from (1,1) to (3,3). Indicator vector $DTUPLE_{i,t}$ contains one dummy for each of these combinations as shown in equation 3.2.

$$DTUPLE_i \quad = \quad BioBirthorderRank_i \ * \ AlternativeCareFamilyRank_{i,t} \quad (3.2)$$

This indicator vector substitutes the former birth order vector in the baseline model equation 3.1, resulting in model equation 3.3. Importantly, I still control for age and employ biological sibship dummies. This implies that β_1 absorbs only the effect of relative rank within an alternative care family.

$$Edu_{i,j,k,t} = \beta_0 + \beta_1 DTUPLE_{i,t} + \gamma_j + \delta_t + \beta_2 X_i + \beta_3 XT_{i,t} + \rho_1 AltCareFam_{k,t} + \varepsilon_{i,t} \quad (3.3)$$

3.4. Results

3.4.1. Results of the baseline model: Birth order effects between biological siblings

Table 3.2 shows this study's baseline results. The estimates suggest the existence of negative birth order effects, implying lower educational achievement of secondborns and laterborns. Educational achievement is regressed on biological birth order, a set of controls and biological sibship dummies with a pooled OLS model. Standard errors are clustered at the level of the individual.

The minimum control model is shown in column (1). It estimates birth order effects with a gender dummy, time dummies, and age dummies only. Its estimate of birth order effects is only significant at the 10 percent level. Column (2) shows the baseline model results with full controls. The effect size in this baseline model indicates a decrease in achievement of 2.4 percent from first- to secondborns.[10] Column (3) replicates this for sibling pairs and excludes all sibships of more than two siblings. Being second-born is further associated with a slightly higher achievement compared to being later-born. However, this difference is not statistically significant at the five percent level in itself.

Across all models, age is negatively associated with educational achievement. The age dummies are not presented for the sake of brevity.[11] Female children outperform male children, ceteris paribus. The results impute no association of the number of non-biological siblings in alternative care families with educational achievement per se. The reason for admission is not found to influence later educational achievement in the overall sample. The insignificance of the reason for admission dummies is expected as the reason for admission rarely varies between siblings.[12] There is no statistically significant evidence at the five percent level for the association between the percentage of one's lifetime spent in SOS alternative care and educational achievement.

10. This figure is calculated by the effect size of -0.079 and the baseline of 3.2 in average educational achievement (This baseline value of 3.2 applies to the youngest children at the age of three and is lower for older individuals. It is calculated by dividing the effect size by the baseline value.)

11. Figure 3.1 presents the negative relationship between age and education.

12. Due to the employment of sibship dummies, the *reason for admission* dummies only capture within-sibship variance.

Table 3.2.: Baseline model: birth order effects between biological siblings

	(1) Baseline model with minimum controls	(2) Baseline model with full controls	(3) Limited baseline model: sibling pairs full controls
Secondborn	-0.051*	-0.079***	-0.096***
	(0.027)	(0.021)	(0.026)
Thirdborn or higher	-0.005	-0.091**	
	(0.040)	(0.041)	
Gender = female	0.104***	0.138***	0.119***
	(0.024)	(0.020)	(0.023)
No. all siblings		-0.003	-0.000
		(0.003)	(0.003)
Lifeshare spent in SOS care		0.158*	0.183*
		(0.090)	(0.111)
Abandonment		0.087	0.139*
		(0.060)	(0.072)
Death of parents		0.103*	0.014
		(0.059)	(0.069)
Referral		-0.034	0.033
		(0.115)	(0.148)
Constant	3.433***	3.233***	3.206***
	(0.131)	(0.158)	(0.185)
Sibship dummies		✓	✓
Time dummies	✓	✓	✓
Age dummies	✓	✓	✓
Observations	26,898	26,898	17,644
$Adj.R^2$	0.016	0.66	0.70
Clusters	4362	4362	2963
Robust standard errors clustered at individual level	Yes	Yes	Yes

Notes: **Dependent variable:** Educational achievement; **Baseline group:** firstborn male children; Sibship dummies; Minimum controls defined as control for sibship dummies, quarter dummies and age dummies only; Robust standard errors in parentheses are clustered at individual level; ***/**/* indicate significance at the 1%/5%/10% level.

3.4.2. Hardship model and gender split model: Sources of heterogeneity in birth order effects

Results of the hardship model

Table 3.3 shows the baseline model results split by the experience of hardship. I find no birth order effects for sibships, who experienced extreme emotional or economic hardship before admission, irrespective of whether their parents passed away or not. Within the subsample without hardship experience, I find birth order effects in both subsamples. Birth order effects are robust and of comparable magnitude for individuals who have experienced parental death (column (1)) and those who have been admitted on other grounds than parental death (column (2)).

I do not find significant birth order effects for the groups with hardship experience (columns (3) through (5)) with the exception of thirdborns who have suffered from emotional hardship. The general absence of birth order effects persists when splitting the model into subregions, as displayed in appendix table C.20.

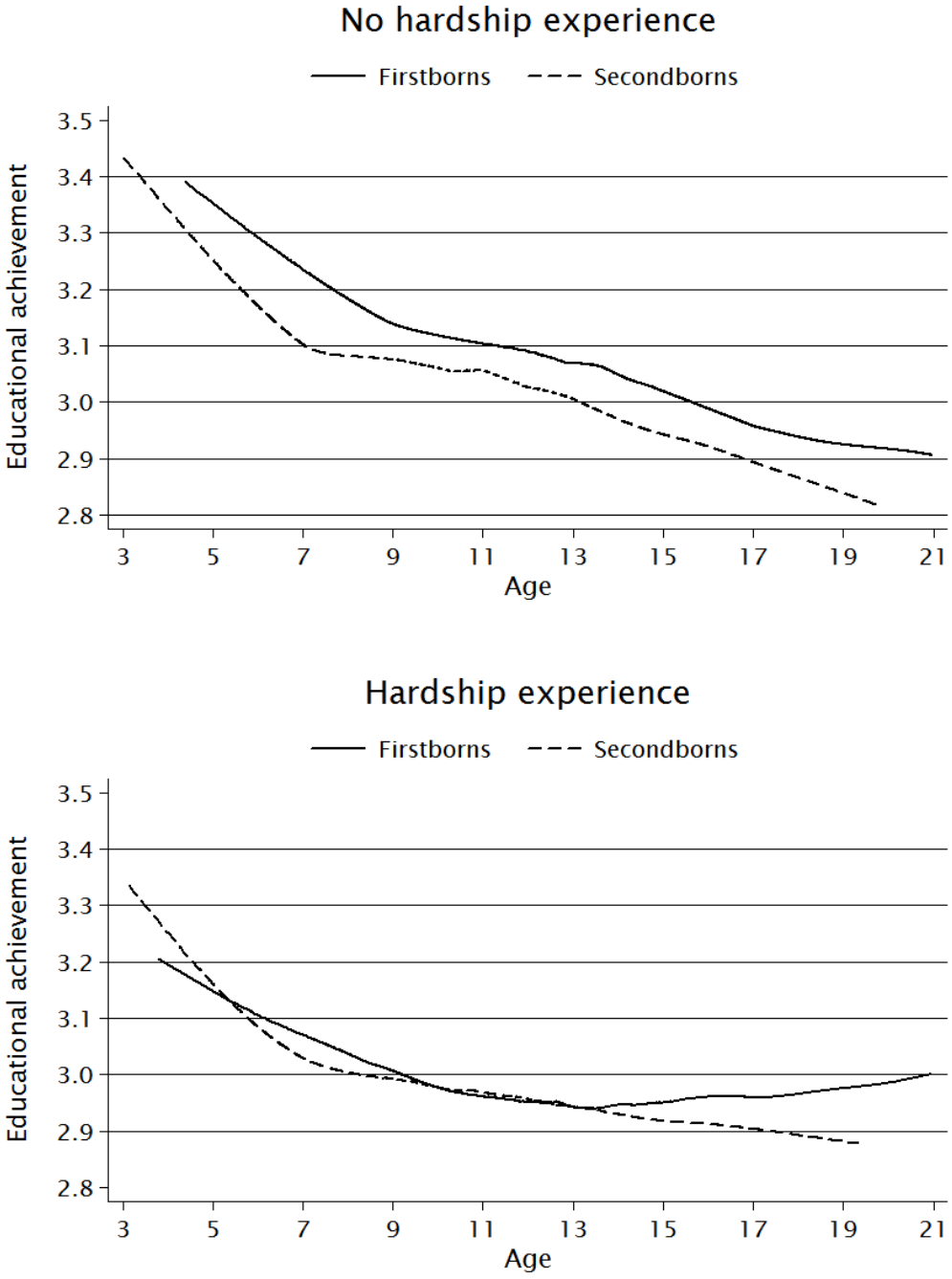
The heterogeneity in birth order effects based on hardship is also supported by the descriptive data displayed in Lowess graphs in figure 3.2. Both figures display the educational achievement of first- and secondborns by age. Birth order effects are more evident in this depiction of raw data for those siblings who have not suffered extreme hardship (top graph) than it is for children who have suffered hardship (bottom graph).

Table 3.3.: Hardship model: baseline model split by experience of hardship

	No extreme hardship		Extreme hardship		
	(1)	(2)	(3)	(4)	(5)
	Parental death	Other reasons	Financial	Emotional	Financial and emotional
Secondborn	-0.091***	-0.091**	0.012	-0.075	-0.030
	(0.032)	(0.046)	(0.052)	(0.050)	(0.037)
Thirdborn or higher	-0.148**	-0.019	0.073	-0.179**	-0.063
	(0.061)	(0.083)	(0.108)	(0.088)	(0.071)
Gender	0.092***	0.107***	0.125**	0.255***	0.216***
	(0.030)	(0.038)	(0.056)	(0.048)	(0.038)
No. all siblings	-0.003	-0.007	0.006	-0.005	-0.002
	(0.004)	(0.007)	(0.006)	(0.007)	(0.005)
Lifeshare spent in SOS care	0.004	0.463**	0.275	-0.210	0.065
	(0.132)	(0.181)	(0.242)	(0.247)	(0.183)
Constant	3.807***	3.053***	1.679**	3.888***	3.588***
	(0.338)	(0.256)	(0.692)	(0.325)	(0.324)
Sibship dummies, time dummies, age dummies	✓	✓	✓	✓	✓
Reason for admission dummies	✓	✓	✓	✓	✓
Observations	13,273	5,665	3,701	4,642	7,960
$Adj.R^2$	0.68	0.71	0.75	0.67	0.70
Clusters	2000	965	601	855	1404
Standard errors clustered at individual level	Yes	Yes	Yes	Yes	Yes

Notes: **Dependent variable:** educational achievement; **Baseline group:** firstborn male children; Sibship dummies; Robust standard errors in parentheses are clustered at individual level; ***/**/* indicate significance at the 1%/5%/10% level.

Figure 3.2.: Children without and with hardship experience: Lowess smoothing of educational achievement vs. age for first- and secondborns



Dependent variable: average educational achievement by age; Top graph: individuals without experience of extreme hardship only; Bottom graph: individuals with experience of personal hardship only; Lowess smoothing.

Results of the gender split model

Table 3.4 reports the results of a gender-specific birth order effect estimation. Based on the baseline estimation, I introduce an additional dummy interacting birth order and gender. The reference group is being a male secondborn. I find that the achievement gap between firstborns and secondborns is attenuated for girls (see column (1) of table 3.4). The regional split shows that Asia drives the moderate global effect (see table 3.4, columns (2) through (4)). First-born girls in Asia exhibit a mitigation of their *firstborn advantage*. Within Asia, this mitigation is entirely attributable to India and Nepal, as shown in the appendix table C.22. In India and Nepal, the full firstborn advantage is revoked for women.

Table 3.4.: Gender split model: baseline model with gender interaction term

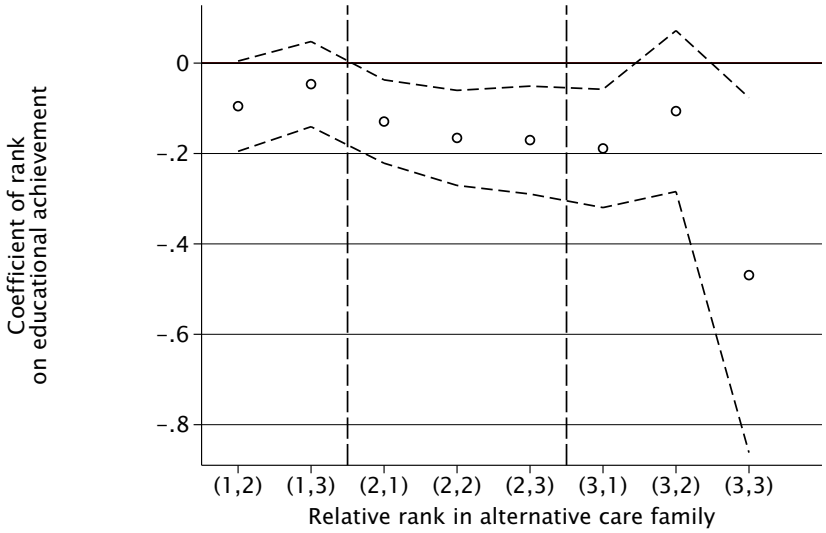
	(1) Full sample	(2) Africa	(3) Latin America	(4) Asia
Firstborn=1	0.121***	0.010	0.111*	0.144***
	(0.026)	(0.069)	(0.062)	(0.032)
Gender = female=1	0.174***	0.144**	0.240***	0.152***
	(0.025)	(0.062)	(0.053)	(0.031)
Firstborn=1 X Gender = female=1	-0.086***	-0.128	-0.053	-0.098**
	(0.033)	(0.082)	(0.077)	(0.041)
Thirdborn	-0.016	-0.025	-0.066	0.006
	(0.028)	(0.072)	(0.064)	(0.035)
Constant	3.405***	3.593***	3.658***	2.165***
	(0.273)	(0.409)	(0.359)	(0.686)
Sibship dummies, time dummies, age dummies	✓	✓	✓	✓
Reason for admission dummies and lifeshare spent in SOS	✓	✓	✓	✓
Observations	26,898	4,200	5,121	17,577
$Adj.R^2$	0.66	0.68	0.56	0.66
Clusters	4362	706	915	2741
Standard errors clustered at individual level	Yes	Yes	Yes	Yes

Notes: **Dependent variable:** educational achievement; **Baseline group:** secondborn male children; Sibship dummies; Robust standard errors in parentheses are clustered at individual level; ***/**/* indicate significance at the 1%/5%/10% level.

Results of the factorial family model

I assess differences between children of the same biological birth order based on their relative rank in their alternative care family via a pooled OLS model. Figure 3.3 summarizes the results. The figure shows the coefficients of the indicator vector describing the relative position within a non-biological alternative care family. The shown data are for sibships without the experience of extreme hardship only (This omission is based on the insight that this group does not show birth order effects). Appendix table C.1 presents the estimates shown in figure 3.3 and for the full sample, including individuals with hardship experience (The effects are smaller if one includes these children). I do not find differences that are statistically significant at the 5 percent level between children who are of the same biological rank but differ in their alternative care family rank. While large standard errors, also due to splitting the sample into nine subgroups, prevent effect estimation of differences within the biological birth order rank groups, figure 3.3 reveals a pattern which hints at the fact that relative alternative care family rank might play a role in educational achievement.

Figure 3.3.: Factorial family model: On the importance of the relative rank in the alternative care family



Notes: **Dependent variable:** educational achievement; **Baseline group:** male children of alternative care family relative rank (1,1); Alternative care family relative rank: Holding biological birth order rank position constant, position in alternative care family based on age. (1,1) is the oldest biological firstborn of an alternative care family, (1,2) is the second oldest biological firstborn and so forth; Alternative care family dummies; Robust standard errors are clustered at individual level; Lines indicate 95 percent confidence interval.

The omitted baseline group is composed of firstborns, who are the oldest firstborn in their alternative care family (assigned tuple (1,1)). Controlling for age, they are expected to perform better than the second oldest biologically firstborn (1,2). However, this effect is only statistically significant at the 10 percent level. Biologically firstborns are expected to perform better than any of the secondborns. Estimates indicate that biological secondborns benefit from ranking higher in their alternative care family, too. Compared to the baseline firstborn, the penalty for being the oldest secondborn is smaller than the one of the third oldest secondborn. Again, the difference between biologically secondborns is not statistically significant. The lowest performing children are those with a biological birth order rank of three or higher, who also come third or later in their alternative care family amongst children of the same biological birth order (3,3). This relationship holds true although the youngest children tend to show the highest achievement.

3.5. Robustness checks and validity

3.5.1. Robustness checks

The main birth order results predicting negative birth order effects as shown in table 3.2 are robust to different choices of dummies on country or family level and the removal of all dummies (see appendix table C.17). I find more pronounced effects for Latin America and Asia when splitting the baseline model by region (see appendix table C.14). As shown in the regional breakdown, results are not driven by an individual country alone (see appendix tables C.15 and C.16). The baseline model specification is also robust to the estimation of single periods as shown in appendix tables C.18 and C.19. For the hardship model, results are also robust to a split by region (see appendix table C.20). Furthermore, I estimate the hardship model in a specification that excludes all individuals with missing information regarding their *detailed reason for admission*. The main finding with respect to the absence of birth order effects amongst children who have experienced hardship remains unaffected by definition. Birth order effects for individuals without hardship experience remain partially unaffected (see appendix table C.21).[13]

13. This is likely be driven by a loss of observations. Particularly children who have lost both parents often have no information regarding the detailed reason for admission. This is expected as the loss of two parents justifies an admission without further explanation of experience of hardship.

3.5.2. Internal and external validity

Internal validity. The endogeneity of fertility decision making poses a challenge to birth order research in general, whenever it is not possible to instrument for fertility decisions. However, the robustness of birth order effects in light of endogeneity has already been shown by Bagger et al. (2013) and Black, Devereux, and Salvanes (2010). Moreover, the results of this study also hold true for different sibship size subsets, such as sibling pairs (see column three of table 3.2). The factorial family model shown in table 3.3 remains unaffected as well, since alternative care siblings' features are statistically independent of biological parents' fertility decisions.

I match siblings based on a codified last name. This method could bias birth order effects downwards. It might exclude *real* biological siblings for example due to differently spelled names and hence differently codified last names. These false negative cases would reduce the sample size. Bias would arise if false negative cases exhibited birth order effects that were systematically different from correctly identified sibships. For this to hold true, one would need to assume an unlikely association between false rejection of a sibship and their (unmeasured) birth order effect size. Meanwhile, the opposite scenario of a false positive scenario is more likely and expected to downward bias the effects: children who by chance carry the same last name are not expected to exhibit any birth order effects and will downward bias the estimation. I use quality control checks to exclude these cases. These are described in appendix section C.2.

A final bias threatening internal validity could arise from non-random patterns in reporting of achievement data. For a considerable share of all children, achievement data has not been reported. These individuals have not been included in this analysis. For results to be unbiased, one needs to assume that a lack of reporting is independent from a child's performance *relative* to its siblings. This assumption seems to be reasonable, in particular as the provision of educational achievement data varies rather at village than at individual level.

External validity. The external validity of this study depends on whether this sample can be considered representative of its underlying population or, alternatively, whether deviations from population averages will bias results. Björklund, Lindahl, and Plug (2004) provide a framework to assess the external validity of studies that employ data from adoptive settings. While this study's setting is different from an adoptive context, it shares important characteristics: non-biological caregivers, non-biological

siblings and a change of the care context from the child's perspective. Björklund, Lindahl, and Plug (2004) argue that three assumptions need to hold true to work with adoption data and to extrapolate findings to biological sibships: (i) Children need to be as good as randomly assigned to their adoptive families, (ii) they need to be adopted early on in their lives and one needs to assume that (iii) studies on adoptive child-parent relationships can be extrapolated to biological child-parent relationships (This last assumption is based on the hypothesis of non-differential treatment of adoptees by adoptive parents as well as the general similarity of individuals in adoptive settings vs. non-adoptive settings concerning unobserved traits). I discuss whether the assumptions (i) through (iii) are met by this study in appendix section C.5.4. I conclude that the sample's traits can downward bias effect sizes but will not lead to a reversal in the effect sign.

3.6. Discussion

Baseline model, gender model, and hardship model. This study is the first one to deliver cross-continental evidence for birth order effects on educational achievement in low- and middle-income countries. I propose individual hardship and parental gender preference as two explanations for previously documented heterogeneity in findings. This within-sample heterogeneity relates the findings to the mixed-picture found in previous studies using data from low- and middle-income countries. The effect size for the baseline model is comparable to previous research. For example, Black, Devereux, and Salvanes (2011) document IQ differences of around 3 points between firstborns and secondborns.

Children who have been exposed to higher degrees of adversity prior to admission to a village show mitigated and potentially reversed birth order effects. A lack of resources in the case of economic hardship and the presence of adversity in the case of emotional hardship provide two explanations.[14] The relevance of hardship experience, and thereby absence of parental resources, ties in with previous literature documenting the role of parental education in high-income countries (Lafortune and Lee, 2014).

14. The absence of child labor within alternative care is expected to dampen effects partially. Previous studies show that the existence of child labor typically disfavors earlier born children. The parental anticipation of earlier born children engaging in child labor in the future can create this disadvantage already previous to the oldest sibling engaging in child labor. The absence of child labor in villages is hence expected to disperse this effect partially compared to settings with child labor.

This finding supports the resource dilution hypothesis by showing that the absence of parental resources can lead to a mitigation of birth order effects.

The second source of heterogeneity is that of cultural gender preferences. A birth order effect model with gender interaction, as shown in Table 3.4, recommends gender as a far-reaching determinant concerning the development of differences between children based on birth order. The disproportionately high share of female children amongst abandoned children in Asia supports the hypothesis of a male preference.[15] Less pronounced birth order effects for female children accord with studies in countries with a preference for male offspring (Jayachandran and Kuziemko, 2011; Fors and Lindskog, 2017).

Factorial family model. This study provides novel insight into interactions of children within non-biological families. The factorial family model suggests that besides birth order, children benefit from having younger siblings in their biologically unrelated family. In this sample's setting, later- and last-born children receive more tutoring opportunities than they would in their biological families. This alludes to existing explanations that propose tutoring opportunities as one driver of the advantageous intellectual development of firstborns (Zajonc, 2001; Zajonc, Markus, and Markus, 1979; Zajonc and Markus, 1975). The estimates in table 3.3 imply a sizable advantage for being in a higher alternative care family rank if biologically later-born. The pattern is only suggestive as the standard errors do not allow to make statements on statistically significant differences. However, I argue that the most likely reason for this pattern is the interaction between children within their alternative care families, with tutoring as a suggested mechanism, as proposed in confluence theory. Eskreis-Winkler, Fishbach, and Duckworth (n.d., in press) show that troubled children, who are asked to motivate others, benefit from mentoring. In a randomized trial setting, the motivation of struggling children to do homework increased more than those of troubled peers that received expert advise. The authors propose that it is a higher self-confidence that leads to higher accomplishment.

The interaction between biologically unrelated siblings is expected to benefit older siblings disproportionately. Exclusion of other potential influences suggests this mechanism. Prenatal factors and postnatal differences in biological parental resource di-

15. In the overall sample, 52 percent of all children are female. The equivalent figure for abandoned children is 52 percent as well. In Asia, 53 percent of all children are female, whereas 56 percent of all abandoned children are female.

lution cannot account for the observed patterns. Biological parents can discriminate between their *own* children before admission but are not in a position to intervene in the period following admission. Selection-based explanations would imply that the parents of older biological sibship groups are systematically different from those of younger ones. I find the relative share of *lifetime spent inside SOS Children's Villages' care* to be statistically insignificant at the five percent level.[16] If older cohorts differed systematically, the share of a child's life spent inside SOS care should capture this effect. So should age control variables.

Discriminatory behavior by the alternative care mother is highly unlikely, too. Theoretically, deliberate discriminatory behavior could induce these differences. If alternative care mothers were actively discriminating in favor of the respective oldest firstborn of a sibship cohort, the relative lifeshare spent inside alternative care should capture this effect. However, it is insignificant throughout all models. Another factor rendering this unlikely is the admission process itself. The alternative care mother is experiencing a *flow* of children over time as younger siblings enter the family while older siblings leave it. The oldest sibling of an alternative care family hence used to be amongst its youngest. The *social* role of being alternative care family firstborn is consequently temporary and developing over time. If being older was beneficial per se, one would find age to have a positive sign and effect on grades – as opposed to its current negative sign. And, if having more siblings was beneficial, one would find the variable reflecting the number of total alternative care family siblings to be significant.

Policy implications. This study provides insight into the educational achievement of vulnerable children in low- and middle-income countries who are without parental care or at risk of losing it. This group is the target of many local and global development programs and policies.

Results carry particular significance for families living in a context which is comparable to the one of SOS Children's Villages. This applies for example to children living with relatives or non-kin families (foster care, youth facilities, and boarding schools) and young asylum seekers/unaccompanied minors who are living in group homes. The suggested interaction between biologically unrelated siblings via tutoring requires acknowledgment by policymakers. Eskreis-Winkler, Fishbach, and Duckworth (n.d., in press) documents the effectiveness of such tutoring interventions.

16. Definition of lifeshare spent in SOS care: Lifetime share spent inside a village divided by the age of an individual.

Acknowledging sources of relative achievement differences within families is important for two reasons. First, building on this and other studies, policymakers will be able to better identify family members in highest need for intervention. Secondly, knowledge about the mechanics at work can inform the type of necessary intervention.

3.7. Conclusion

This study contributes to the debate on the formation of human capital and determinants of educational achievement in low- and middle-income countries. It does so by drawing from a dataset from three continents. The findings suggest strong heterogeneity in human capital formation, with economic and emotional hardship, parental gender preferences and sibling interaction as mediating factors.

This study also contributes to the on-going theoretical dispute on the relevance of tutoring effects to explain birth order effects. The results can inform policy interventions by identifying the most vulnerable members of families and describing the drivers of the development of within-family differences. Tutoring is pointed out as a potentially under-appreciated mediator of personal growth. While these findings advance the discussion, more multi-country evidence is needed to understand the mediators of birth order effects more profoundly.

Appendix A. Diverging perceptions of female decision-making power

A.1. Decision domains and contraception typology

Table A.1 presents the list of decision domains. The lead question for every domain listed in the following is "In your household, who makes decisions about:". Individuals can then circle letters representing single household members, such as A for the respondent, B for the spouse, etc.

Table A.1.: List of single decision domains

Code	Decisions	Considered in analysis	Reason for omission
A1	Expenditure on food eaten at home	Yes	Not applicable
A2	Choice of food eaten at home	No	Considered as indicator for housekeeping role
B	Routine purchases for the household of items such as cleaning supplies	Yes	Not applicable
C	Your clothes	No	Indicator for housekeeping role; unclear interpretation ("choice" vs. "expenditure")
D	Your spouse's clothes	No	
E	Your children's clothes	No	
F	Your children's education	Yes	Not applicable
G	Your children's health	Yes	Not applicable
H	Large expensive purchases for the household (i.e., refrigerator or TV)	Yes	Not applicable
I	Giving money to your parents/family	Yes	Not applicable
J	Giving money to your spouse's parents/family	Yes	Not applicable
K	Gifts for parties/weddings	Yes	Not applicable
L	Money for monthly arisan (savings lottery)	Yes	Not applicable
M	Money for monthly savings	Yes	Not applicable
N	Time the husband spends socializing	Yes	Not applicable
O	Time the wife spends socializing	Yes	Not applicable
P	Whether you/your spouse works?	Yes	Not applicable
Q	Whether you and your spouse use contraception?	Yes	Not applicable

Table A.2 presents the typology of statement combinations used in the contraceptive use decision domain. Our main estimates are based on couples of the first three

concordance categories cCM, cCF and cCB and the first three discordance categories cDFM, cDMF and cDOBOF. The grey-colored categories (cNM, cNF, cFN and cMN) are not used in our models since they describe an outcome and decision at the same time.

Table A.2.: Typology of responses: contraceptive use domain

Code	Combination of husband's and wife's statement	
	Concordance	
cCM	Wife and Husb: (m)	Both partners agree on the husband as the sole decision maker
cCF	Wife and Husb: (f)	Both partners agree on the wife as the sole decision maker
cCB	Wife and Husb: (mf)	Both partners agree on joint decision making
cCN	Wife and Husb: (no use)	Both partners agree that they do not use contraceptive methods
	Discordance	
cDFM	Wife: (f) $\vee$ (mf); Husb: (m)	Strong Discordance: wife perceives female decision-making power (individually or jointly), husband does not
cDMF	Husb: (f) $\vee$ (mf); Wife: (m)	Strong Discordance: husband perceives female decision-making power (individually or jointly), wife does not
cDOBOF	One Spouse: (mf); Other Spouse: (f)	Weak Discordance: both spouses perceive female decision-making power, one of them perceives a sole female decision maker, the other joint decision making
cNM	Wife: (no use); Husb: (m)	Wife states that couple does not use contraception, husband perceives male decision making power only
cNF	Wife: (no use); Husb: (f) $\vee$ (mf)	Wife states that couple does not use contraception, Husband perceives female decision making power (individually or jointly)
cFN	Wife: (f) $\vee$ (mf); Husb: (no use)	Wife perceives female decision making power (individually or joint), husband states that couple does not use contraception
cMN	Wife: (m); Husb: (no use)	Wife perceives male decision making power only, husband states that couple does not use contraception

A.2. Definitions and codings of variables

A.2.1. Distribution and coding of educational achievement

Table A.3.: Distribution and coding (see table notes) of educational achievement

	What is the highest educational level attended?	
	Women	Men
Educational level	– in percent of column	– in percent of column
2:Elementary school	36.1	34.9
3:Junior high general	17.4	15.4
4:Junior high vocational	0.5	0.8
5:Senior high general	14.4	15.8
6:Senior high vocational	9.7	12.5
11:Adult education A	0.0	0.1
12:Adult education B	0.3	0.4
13:Open university	0.1	0.1
14:Islamic School (pesantren)	0.2	0.2
15:Adult education C	0.5	1.0
60:College (D1,D2,D3)	3.8	3.1
61:University S1	7.9	9.0
62:University S2	0.4	0.8
63:University S3	0.0	0.1
72:Islamic Elementary School (Madrasah Ibtidaiyah)	1.7	1.3
73:Islamic Junior/High School (Madrasah Tsanawiyah)	4.3	2.4
74:Islamic Senior/High School (Madrasah Tsanawiyah)	2.6	2.0
90:Kindergarten	0.0	
95:Other	0.0	0.0
98:Don't Know		0.0
Observations	8,277	8,419

Notes: **Data:** IFLS-5 wave (2014), cross-sectional data, one observation is one couple; Categories 2 and 72 are coded as elementary education, categories 3 to 6 and 11 to 15 and 73 and 74 are coded as secondary education, categories 60 to 63 are coded as tertiary education. All remaining categories, for instance, group 90, are grouped in the *no education* category. Individuals who reported that they never received education are not listed in any category but also assigned to the *no education* group of our model.

A.2.2. Classification of covert methods

There are different approaches to classify covert methods. Ashraf, Field, and Lee (2014) limit their analysis to injectables and test the results to implants and IUDs. Gasca and Becker (2017, p. 7) estimate the rate of indirect, covert use as the share of couples in which women use *"female modern contraceptive methods"* and men reported not a modern method. They consider "female sterilization, contraceptive pill, implant, injectable, IUD, diaphragm/foam/jelly and the female condom" as a modern method. Biddlecom and Fapohunda (1998) report method utilization rates for both women who report open and covert use. For example, injectables are preferred by women who report covert use while only a small share of women who openly use contraceptives report their use. The following table A.4 is based on previous studies. The second column presents our reasoning for the classification.

Table A.4.: Classification of covert methods

Covert methods	**Reasoning**
Injection (1,2,3 months)	Classified as covert method by Gasca and Becker (2017).
IUD (Intrauterin device)/ AKDR/Spiral	Classified as covert method by Gasca and Becker (2017).
Norplant / Implant	Discussed by Gasca and Becker (2017) but not considered in their context due to limited availability of method.
Female Sterilization	Can be used without husband taking notice.
Non-Covert methods	
Pill	It is possible to covertly use the pill. However, Chikovore et al., 2002 and Ashraf, Field, and Lee (2014) describe the difficulties of women to hide them consistently from men. We run a robustness check for an alternative classification.
Condom	Not useable without husband taking notice.
Rhythm / Calendar	Difficult to use without husband taking notice.
Coitus Interruptus	Not useable without husband taking notice.
Male Sterilization	Not useable without husband taking notice.
Methods not considered	
Traditional Herbs	Traditional methods excluded analogous to Gasca and Becker (2017)
Traditional Massage	Traditional methods excluded analogous to Gasca and Becker (2017)
Femidom	Ambiguity of covert/non-covert status leads to exclusion. (However Gasca and Becker (2017) classifies female condoms as female modern contraceptive method.)
Intravag	Ambiguity of covert/non-covert status
Other	Ambiguity of covert/non-covert status

Notes: AKDR = Alat Kontrasepsi Dalam Rahim (Contraception Installation Tool Rahim)

A.2.3. Definition of the instrumental variable

The instrumental variable approach is based on household exposure to natural disasters and the reactions of households to them. Our knowledge of event experience and reactions is based on responses to two questions in the IFLS-3 (2000) Survey. The first question GE01 asks respondents about the experience of the list of events. It elicits whether the household has experienced a crop loss or a natural disaster. The question and options are listed in table A.5. The second question is about how the household reacted to this event. The question and options are listed in table A.6. We only consider the events of C crop loss and D natural disaster of table A.5 and the options Q, S, T, U, V, W of table A.6. We do not include reactions O and P as the inclusion leads to underidentification in some model specifications (see statistics of appendix table A.5.4).

Table A.5.: List of possible household shocks

GE01: Has this household gone through [...] in the last 5 years? (Yes/No)	
A	Death of a householder or a family member who is not a householder
B	Sickness of a householder or a family member who is not a householder that necessitated hospitalization or continuous medical treatment
C	Crop loss, reason (Other) [Blank space]
D	Household/business loss due to earthquake, fire or other natural disaster
E	Any of the householders lost a job or failed in business?
F	Decrease of household income, due to decrease of production or very low price of products?
G1	Other, [Blank space]
G2	Other, [Blank space]

Notes: Question and responses taken from the IFLS-3 Survey (2000) supplementary materials; [Blank space]: Space for interviewer to fill text.

Table A.6.: List of possible household reactions

GE03: What steps have been taken by household members in response to this difficulty?			
A	Eat less food	O	Close/reduce business activities
B	Buy less food	P	Changed business
C	Take child out of school	Q	Changed job/quit job
D	Reduce current spending	R	Take a border into the household
E	Delay plans to spend money	S	Start working for pay
F	Save less money	T	Increase working hours
G	Use savings	U	Take an additional job
H	Sell possessions	V	Expand business activities
I	Borrow money	W	Start a business
J	Move within village	X	Pray
K	Move to new village	Y	Other [Blank space]
L	Receive assistance from friends/family	Z	Nothing
M	Receive assistance from government		
N	Receive assistance from other group		

Notes: Question and responses taken from the IFLS-3 Survey (2000) supplementary materials; [Blank space]: Space for interviewer to fill text.

A.3. Auxiliary tables

A.3.1. Differences in attributes of couples that give discordant responses

Table A.7.: Couple attributes by discordance type in labor market decision domain

	DFM(1) vs. DOBOF (2)			DMF(1) vs DOBOF(2)		
	Mean(1)	Mean(2)	Diff.	Mean(1)	Mean(2)	Diff.
Individual attributes:						
Age	38.58	43.61	-5.03***	39.46	43.61	-4.15***
Sp: age	42.96	48.43	-5.47***	44.27	48.43	-4.16***
Elementary education	0.38	0.40	-0.02	0.40	0.40	0.00
Any secondary education	0.49	0.42	0.07**	0.47	0.42	0.05*
Any college education	0.08	0.11	-0.02	0.08	0.11	-0.03
Sp: Elementary education	0.37	0.43	-0.07**	0.36	0.43	-0.07**
Sp: Any secondary education	0.50	0.44	0.06*	0.48	0.44	0.05
Sp: Any college education	0.10	0.09	0.01	0.13	0.09	0.03
Household attributes:						
N HH adults	2.40	2.47	-0.08	2.39	2.47	-0.09*
N HH children	1.70	1.57	0.13*	1.75	1.57	0.18**
Migration indicator	0.21	0.16	0.05*	0.22	0.16	0.06**
Household economy:						
Log HH income	16.65	16.56	0.09	16.56	16.56	0.00
Any land	0.32	0.33	-0.02	0.34	0.33	0.01
Female income share	0.16	0.36	-0.20***	0.13	0.36	-0.24***
Log h worked pa	6.91	7.17	-0.26***	6.74	7.17	-0.43***
Sp: Log h worked pa	7.41	7.26	0.15**	7.39	7.26	0.13*
Decision share:						
Wife: own decision share	0.79	0.76	0.03**	0.61	0.76	-0.15***
Husb: wife's decision share	0.59	0.78	-0.19***	0.78	0.78	-0.00
Wife: husband's decision share	0.62	0.45	0.16***	0.62	0.45	0.17***
Husb: own decision share	0.65	0.58	0.07***	0.69	0.58	0.11***

Notes: T-Test; **Data:** IFLS-5 wave (2014), cross-sectional data, one observation is one couple; Individual level attributes are wife's attributes by default, Spouse (Sp:) refers to husband's attributes; **Decision making domain:** "Whether you/your spouse works"; **Variable definitions:** N HH children: number of children in household; N HH adults: number of adults in household; Migration indicator: household moved since last wave (yes = 1/no = 0); Log HH income: Log of annual income in IDR; Wife:/Husb.: provides wife's/husband's perspective on decision making power respectively; **Typology:** DFM: wife perceives female decision making power, husband does not, DMF: husband perceives female decision making power, wife does not, DOBOF: both spouses perceive female decision making power, one of them perceives a sole female decision maker, the other joint decision making; ***/**/* indicate significance of difference at the 1%/5%/10% level.

A.3.2. Contraception and covert method use estimates

The wife reports the use of contraception and the used method in a separate interview. Table A.8 displays the association between responses to the question of who makes decisions about "whether you and your spouse use contraception" and contraceptive use as the dependent variable.

Table A.9 displays the association between responses to the question of who makes decisions about "whether you and your spouse use contraception" and covert method use as the dependent variable, conditional on contraceptive use.

Appendix table A.40 provides the average reported contraceptive and covert method rates for women whose spouses were and were not present during this interview.

Table A.8.: Estimate of association between statement combinations and contraceptive use

	Dependent variable: Couple uses contraception				
	(1)	(2)	(3)	(4)	(5)
	No controls	Baseline	Baseline with exclusion restriction 1	Baseline with exclusion restriction 2	Baseline with exclusion restrictions 1 and 2
Concordance:					
cCM: Wife and Husb: (m)	-0.226***	-0.234***	-0.145*	-0.233***	-0.112
	(0.041)	(0.045)	(0.075)	(0.048)	(0.074)
cCB: Wife and Husb: (mf)	-0.071***	-0.061***	-0.045***	-0.053***	-0.030***
	(0.018)	(0.018)	(0.015)	(0.017)	(0.010)
Discordance:					
cDFM: Wife: (f) ∨ (mf); Husb: (m)	-0.078**	-0.106***	-0.109**	-0.094***	-0.099***
	(0.034)	(0.028)	(0.038)	(0.028)	(0.032)
cDMF: Husb: (f) ∨ (mf); Wife: (m)	-0.058**	-0.077***	-0.038	-0.078***	-0.041*
	(0.022)	(0.022)	(0.025)	(0.019)	(0.022)
cDOBOF: One spouse: (f); Other spouse: (mf)	-0.022	-0.029	-0.011	-0.022	0.001
	(0.021)	(0.019)	(0.018)	(0.014)	(0.016)
Constant	0.756***	1.037***	1.198***	1.169***	1.276***
	(0.018)	(0.191)	(0.188)	(0.169)	(0.199)
Region dummies		✓	✓	✓	✓
Control for (spousal) age, HH income and education		✓	✓	✓	✓
Control for N HH adults, N HH children, migration		✓	✓	✓	✓
Observations	5,682	5,167	2,571	4,716	2,475
R^2	0.0075	0.069	0.052	0.076	0.070
Clusters	18	18	18	18	18
Robust standard errors clustered at regional level	Yes	Yes	Yes	Yes	Yes

Note: Standard OLS estimate; **Dependent variable:** contraceptive use; **Baseline group:** concordant couples, perceiving wife as the sole decision maker; Individual level attributes are wife's attributes by default, Spouse (Sp:) refers to husband's attributes; **Data:** IFLS-5 wave (2014), cross-sectional data, one observation is one couple; **Decision making domain:** "whether you and your spouse use contraception"; Exclusion restriction 1: exclusion of women who respond that they personally want to receive any additional children ("Do you personally wish to have another child?"); Exclusion restriction 2: exclusion of women who state that they do not use contraceptives, due to divorcee/widow status, recent birth (being pre-menstrual or absent from sex) or current breastfeeding; **Typology:** cCM: concordant report of husband as sole decision maker, cCB: concordant report of joint decision making, cDFM: wife perceives female decision making power, husband does not, cDMF: husband perceives female decision making power, wife does not, cDOBOF: both spouses perceive female decision making power, one of them perceives a sole female decision maker, the other joint decision making; Robust standard errors in brackets; Robust standard errors in parentheses are clustered at regional level; ***/**/* indicate significance at the 1%/5%/10% level.

Table A.9.: Estimate of association between statement combinations and covert method use

	Dependent variable: Couple uses covert method conditional on contraceptive use				
	(1)	(2)	(3)	(4)	(5)
	No controls	Baseline	Baseline with exclusion restriction 1	Baseline with exclusion restriction 2	Baseline with exclusion restrictions 1 and 2
Concordance:					
cCM: Wife and Husb: (m)	-0.050	-0.077	-0.147	-0.077	-0.147
	(0.063)	(0.060)	(0.089)	(0.060)	(0.089)
cCB: Wife and Husb: (mf)	0.089***	0.056***	0.029	0.056***	0.029
	(0.027)	(0.014)	(0.019)	(0.014)	(0.019)
Discordance:					
cDFM: Wife: (f) ∨ (mf); Husb: (m)	0.100***	0.088***	0.110***	0.088***	0.110***
	(0.024)	(0.023)	(0.024)	(0.023)	(0.024)
cDMF: Husb: (f) ∨ (mf); Wife: (m)	0.084***	0.052	0.014	0.052	0.014
	(0.028)	(0.030)	(0.049)	(0.030)	(0.049)
cDOBOF: One Spouse: (f); Other Spouse: (mf)	0.051***	0.052***	0.040**	0.052***	0.040**
	(0.016)	(0.016)	(0.019)	(0.016)	(0.019)
Constant	0.685***	1.133***	1.144***	1.133***	1.144***
	(0.023)	(0.140)	(0.190)	(0.140)	(0.190)
Region dummies		✓	✓	✓	✓
Control for (spousal) age, HH income and education		✓	✓	✓	✓
Control for N HH adults, N HH children, migration		✓	✓	✓	✓
Observations	4,016	3,640	2,047	3,640	2,047
R^2	0.0072	0.069	0.071	0.069	0.071
Clusters	18	18	18	18	18
Robust standard errors clustered at regional level	Yes	Yes	Yes	Yes	Yes

Note: Standard OLS estimate; **Dependent variable:** covert contraceptive method use; **Baseline group:** concordant couples, perceiving wife as the sole decision maker; **Data:** IFLS-5 wave (2014), cross-sectional data, one observation is one couple; **Decision making domain:** "whether you and your spouse use contraception"; Exclusion restriction 1: exclusion of women who respond that they personally want to receive any additional children ("Do you personally wish to have another child?"); Exclusion restriction 2: exclusion of women who state that they do not use contraceptives, due to divorce/widow status, recent birth (being premenstrual or absent from sex) or current breastfeeding; **Typology:** cCM: concordant report of husband as sole decision maker, cCB: concordant report of joint decision making, cDFM: wife perceives female decision making power, husband does not, cDMF: husband perceives female decision making power, wife does not, cDOBOF: both spouses perceive female decision making power, one of them perceives a sole female decision maker, the other joint decision making; Robust standard errors in brackets; Robust standard errors in parentheses are clustered at regional level; ***/**/* indicate significance at the 1%/5%/10% level.

A.3.3. Fixed effects analysis: Single domains

Tables A.10, A.11, A.12, A.13, A.14, A.15, and A.16 employ a panel model approach with individual fixed effects to survey the relationship of changes in female economic and decision making power over time – from both perspectives, the husband's and the wife's. Female power is surveyed across multiple domains of eight economic household decisions – each table presents two. Economic power is measured as the wife's contribution to household income. Across all tables, the dependent variable is whether the wife or husband agree that the wife has any say in the respective decision domain, irrespective of the spouse's position.

If the wife contributes more to the household income, husbands, as well as wives, tend to attribute more power to the wife over time in the dimensions of expensive purchases and labor market participation. A similar picture of less statistical significance emerges for expenditure for food. As the female income share increases, neither husbands or wives do attribute more decision making power to women with respect to savings and giving money to either of the spouses' families. Women associate higher relative income of themselves with more say over routine purchases and gifts for parties/weddings.

Table A.10.: Female share of income and female decision making power across domains (1/7)

	Dependent variable: Wife has decision-making power in domain of expenditure for food eaten at home		Dependent variable: Wife has decision-making power in domain of routine purchases for the HH of items such as cleaning supplies	
	(1) Wife says yes	(2) Husband says yes	(3) Wife says yes	(4) Husband says yes
Female income share	0.058**	0.074***	0.048***	0.017
	(0.023)	(0.024)	(0.014)	(0.015)
Constant	0.305	1.302***	0.535***	1.024***
	(0.349)	(0.421)	(0.127)	(0.184)
Control for (spousal) age and HH income	✓	✓	✓	✓
Control for N HH adults, N HH children, migration	✓	✓	✓	✓
Observations	14,802	14,802	20,855	20,855
Time dummies and individual fixed effects	Yes	Yes	Yes	Yes
Robust standard errors	Yes	Yes	Yes	Yes

Notes: Fixed effects estimate; **Dependent variable:** wife (husband) states that wife has decision making power in specific domain; Individual level attributes are wife's attributes by default, Spouse (Sp:) refers to husband's attributes; **Data:** IFLS waves 3, 4 and 5, panel data 2000 to 2014, one observation is one couple; ***/**/* indicate significance at the 1%/5%/10% level.

Table A.11.: Female share of income and female decision making power across domains (2/7)

	Dependent variable: Wife has decision-making power in domain of children's education		Dependent variable: Wife has decision-making power in domain of children's health	
	(1) Wife says yes	(2) Husband says yes	(3) Wife says yes	(4) Husband says yes
Female income share	0.032*	0.055***	0.004	0.008
	(0.019)	(0.020)	(0.016)	(0.018)
Constant	0.657***	0.331	0.646***	0.495*
	(0.247)	(0.320)	(0.243)	(0.259)
Control for (spousal) age and HH income	✓	✓	✓	✓
Control for N HH adults, N HH children, migration	✓	✓	✓	✓
Observations	20,855	20,855	20,851	20,851
Time dummies and individual fixed effects	Yes	Yes	Yes	Yes
Robust standard errors	Yes	Yes	Yes	Yes

Notes: Fixed effects estimate; **Dependent variable:** wife (husband) states that wife has decision making power in specific domain; Individual level attributes are wife's attributes by default, Spouse (Sp:) refers to husband's attributes; **Data:** IFLS waves 3, 4 and 5, panel data 2000 to 2014, one observation is one couple; ***/**/* indicate significance at the 1%/5%/10% level.

Table A.12.: Female share of income and female decision making power across domains (3/7)

	Dependent variable: Wife has decision-making power in domain of large expensive purchases (i.e., refrigerator or TV)		Dependent variable: Wife has decision-making power in Domain of Gifts for parties/weddings	
	(1) Wife says yes	(2) Husband says yes	(3) Wife says yes	(4) Husband says yes
Female income share	0.054***	0.048**	0.037***	0.014
	(0.018)	(0.019)	(0.013)	(0.015)
Constant	0.332	0.698***	1.029***	0.884***
	(0.213)	(0.246)	(0.163)	(0.171)
Control for (spousal) age and HH income	✓	✓	✓	✓
Control for N HH adults, N HH children, migration	✓	✓	✓	✓
Observations	20,740	20,740	20,851	20,851
Time dummies and individual fixed effects	Yes	Yes	Yes	Yes
Robust standard errors	Yes	Yes	Yes	Yes

Notes: Fixed effects estimate; **Dependent variable:** wife (husband) states that wife has decision making power in specific domain; Individual level attributes are wife's attributes by default, Spouse (Sp:) refers to husband's attributes; **Data:** IFLS waves 3, 4 and 5, panel data 2000 to 2014, one observation is one couple; ***/**/* indicate significance at the 1%/5%/10% level.

Table A.13.: Female share of income and female decision making power across domains (4/7)

	Dependent variable: Wife has decision-making power in domain of giving money to wife's family		Dependent variable: Wife has decision-making power in domain of giving money to husband's family	
	(1) Wife says yes	(2) Husband says yes	(3) Wife says yes	(4) Husband says yes
Female income share	0.013	-0.017	0.006	-0.018
	(0.018)	(0.018)	(0.018)	(0.020)
Constant	0.939***	0.979***	0.761***	0.727***
	(0.223)	(0.241)	(0.275)	(0.232)
Control for (spousal) age and HH income	✓	✓	✓	✓
Control for N HH adults, N HH children, migration	✓	✓	✓	✓
Observations	20,850	20,850	20,851	20,851
Time dummies and individual fixed effects	Yes	Yes	Yes	Yes
Robust standard errors	Yes	Yes	Yes	Yes

Notes: Fixed effects estimate; **Dependent variable:** wife (husband) states that wife has decision making power in specific domain; Individual level attributes are wife's attributes by default, Spouse (Sp:) refers to husband's attributes; **Data:** IFLS waves 3, 4 and 5, panel data 2000 to 2014, one observation is one couple; ***/**/* indicate significance at the 1%/5%/10% level.

Table A.14.: Female share of income and female decision making power across domains (5/7)

	Dependent variable: Wife has decision-making power in domain of money for monthly Arisan (lottery)		Dependent variable: Wife has decision-making power in domain of monthly savings	
	(1) Wife says yes	(2) Husband says yes	(3) Wife says yes	(4) Husband says yes
Female income share	0.052**	0.028	0.036	0.025
	(0.022)	(0.023)	(0.023)	(0.024)
Constant	-0.149	-0.093	-0.295	-0.578**
	(0.299)	(0.287)	(0.275)	(0.276)
Control for (spousal) age and HH income	✓	✓	✓	✓
Control for N HH adults, N HH children, migration	✓	✓	✓	✓
Observations	20,855	20,855	20,855	20,855
Time dummies and individual fixed effects	Yes	Yes	Yes	Yes
Robust standard errors	Yes	Yes	Yes	Yes

Notes: Fixed effects estimate; **Dependent variable:** wife (husband) states that wife has decision making power in specific domain; Individual level attributes are wife's attributes by default, Spouse (Sp:) refers to husband's attributes; **Data:** IFLS waves 3, 4 and 5, panel data 2000 to 2014, one observation is one couple; ***/**/* indicate significance at the 1%/5%/10% level.

Table A.15.: Female share of income and female decision making power across domains (6/7)

	Dependent variable: Wife has decision-making power in domain of time husband spends socializing		Dependent variable: Wife has decision-making power in domain of time wife spends socializing	
	(1) Wife says yes	(2) Husband says yes	(3) Wife says yes	(4) Husband says yes
Female income share	-0.013	0.075***	-0.014	0.010
	(0.025)	(0.024)	(0.014)	(0.016)
Constant	0.330	0.503	0.820***	0.861***
	(0.291)	(0.347)	(0.157)	(0.179)
Control for (spousal) age and HH income	✓	✓	✓	✓
Control for N HH adults, N HH children, migration	✓	✓	✓	✓
Observations	20,855	20,855	20,855	20,855
Time dummies and individual fixed effects	Yes	Yes	Yes	Yes
Robust standard errors	Yes	Yes	Yes	Yes

Notes: Fixed effects estimate; **Dependent variable:** wife (husband) states that wife has decision making power in specific domain; Individual level attributes are wife's attributes by default, Spouse (Sp:) refers to husband's attributes; **Data:** IFLS waves 3, 4 and 5, panel data 2000 to 2014, one observation is one couple; ***/**/* indicate significance at the 1%/5%/10% level.

Table A.16.: Female share of income and female decision making power across domains (7/7)

	Dependent variable: Wife has decision-making power in domain of labor market decision-making		Dependent variable: Wife has decision-making power in domain of contraception use	
	(1) Wife says yes	(2) Husband says yes	(3) Wife says yes	(4) Husband says yes
Female income share	0.291***	0.334***	0.017	-0.027
	(0.019)	(0.021)	(0.020)	(0.021)
Constant	0.002	0.351	1.078***	0.517**
	(0.258)	(0.250)	(0.262)	(0.253)
Control for (spousal) age and HH income	✓	✓	✓	✓
Control for N HH adults, N HH children, migration	✓	✓	✓	✓
Observations	20,855	20,855	20,847	20,847
Time dummies and individual fixed effects	Yes	Yes	Yes	Yes
Robust standard errors	Yes	Yes	Yes	Yes

Notes: Fixed effects estimate; **Dependent variable:** wife (husband) states that wife has decision making power in specific domain; Individual level attributes are wife's attributes by default, Spouse (Sp:) refers to husband's attributes; **Data:** IFLS waves 3, 4 and 5, panel data 2000 to 2014, one observation is one couple; ***/**/* indicate significance at the 1%/5%/10% level.

A.3.4. Fixed effects analysis: overall decision share

Table A.17.: Female share of income and perceived female decision making power across all domains

	(1) Wife's perspective	(2) Husband's perspective
Female income share	0.045***	0.048***
	(0.008)	(0.009)
Log HH income	0.015***	0.013***
	(0.002)	(0.003)
Constant	0.463***	0.516***
	(0.098)	(0.119)
Control for (spousal) age and HH income	✓	✓
Control for N HH adults, N HH children	✓	✓
Observations	20,855	20,855
Year, Regional FE	Yes	Yes
Robust SEs	Yes	Yes

Notes: Fixed effects estimate; **Dependent variable:** female overall decision share; **Data:** IFLS waves 3, 4 and 5, panel data 2000 to 2014, one observation is one couple; Definition: Overall female/male decision making share: Number of decisions with female/male involvement divided by number of overall decisions; Robust standard errors in brackets; ***/**/* indicate significance at the 1%/5%/10% level.

Table A.18.: Without labor domain: female share of income and perceived female decision making power across all domains

	(1) Wife's perspective	(2) Husband's perspective
Female income share	0.023***	0.022**
	(0.008)	(0.009)
Log HH income	0.014***	0.011***
	(0.002)	(0.003)
Constant	0.505***	0.531***
	(0.100)	(0.122)
Control for (spousal) age and HH income	✓	✓
Control for N HH adults, N HH children	✓	✓
Observations	20,855	20,855
Year, Regional FE	Yes	Yes
Robust SEs	Yes	Yes

Notes: Fixed effects estimate; **Dependent variable:** female overall decision share; **Data:** IFLS waves 3, 4 and 5, panel data 2000 to 2014, one observation is one couple; Definition: Overall female/male decision making share: Number of decisions with female/male involvement divided by number of overall decisions excluding labor domain decision making; Robust standard errors in brackets; ***/**/* indicate significance at the 1%/5%/10% level.

A.3.5. Instrumental variable first stage estimates

Table A.19.: Instrumental variable first stage estimate: effects of labor supply shocks on female income share

	Dependent variable: Female income share			
	(1) Land-owners only		(2) Full sample	
CL: Start working for pay	-0.063***	(0.023)	-0.040	(0.024)
CL: Changed job/quit job	0.171	(0.121)	0.145*	(0.074)
CL: Increased working hours	0.015	(0.033)	0.014	(0.031)
CL: Take an additional job	-0.050*	(0.030)	-0.044*	(0.024)
CL: Expand business acitvities	-0.057**	(0.025)	-0.057**	(0.023)
CL: Start a business	0.023	(0.032)	0.037	(0.029)
ND: Start working for pay	-0.039	(0.087)	0.063	(0.090)
ND: Changed job/quit job	0.164***	(0.020)	0.155***	(0.015)
ND: Increased working hours	0.031	(0.145)	-0.124	(0.099)
ND: Take an additional job	-0.120***	(0.027)	-0.071**	(0.030)
ND: Expand business acitvities	0.000	(.)	0.000	(.)
ND: Start a business	0.000	(.)	0.000	(.)
Any land			-0.046***	(0.008)
Constant	-0.351***	(0.086)	-0.369***	(0.059)
Control for (spousal) age and education	✓		✓	
Control for HH income, N HH adults, N HH children	✓		✓	
Observations	2,276		6,053	
Regional dummies	Yes		Yes	
Robust standard errors	Yes		Yes	

Note: Instrumental variable first stage estimate; **Data:** IFLS-3 wave (2000), cross-sectional data, one observation is one couple; Independent variables express changes in labor supply due to either ND: natural disaster or CL: crop loss; Estimates do not vary between husband's perspective and wife's perspective as only wife's income and not decision share is estimated in first stage; Robust standard errors in brackets are clustered at regional level; ***/**/* indicate significance at the 1%/5%/10% level.

A.4. Descriptive statistics

A.4.1. Contraception

Table A.20.: Frequency of use of contraceptive methods over time

	Which birth control device/method do you/does your husband use now?			
	2000	2007	2014	Full Sample
Contraceptive method	in percent of column	in percent of column	in percent of column	in percent of column
Pill	24.1	23.2	20.9	22.6
1_Month_Injection	1.7	5.7	8.1	5.5
2_Month_Injection	0.3	0.5	0.5	0.4
3_Month_Injection	35.5	47.1	45.8	43.5
Intravag	0.1	0.1		0.1
Condom	1.7	1.8	2.9	2.2
IUD/AKDR/Spiral	15.0	8.1	8.1	10.0
Norplant/Implant	8.7	4.0	5.9	6.0
Female_Sterilization/Tubectomy	8.6	5.0	5.2	6.1
Male_Sterilization	0.5	0.3	0.3	0.4
Rhythm/calendar	2.1	2.6	1.4	2.0
Coitus_Interruptus	0.3	1.0	0.3	0.6
Traditional_Herbs	1.1	0.4	0.1	0.5
Traditional_Massage	0.1	0.1		0.1
Other		0.1	0.3	0.2
Observations	3,094	4,017	4,228	11,339

Notes: Three residual categories merged into other section; One couple is one observation; The rate of covert contraceptive use is calculated as the number of women who use contraceptive methods divided by all women who use contraceptives.

Table A.21.: Contraception domain: contraceptive use and covert method use by response type

	Frequencies of responses by couple if wife reports in seperate interview...				Full sample
	(1)	(2)	(3)	(4)	(5)
	... no use	... use	... no covert method	... covert method	Full sample
	– in percent of column	– in percent of column	– in percent of column	– in percent of column	– in percent of column
Concordance:					
cCM: Wife and Husb: (m)	0.02	0.01	0.02	0.01	0.02
cCF: Wife and Husb: (f)	0.10	0.19	0.23	0.18	0.15
cCB: Wife and Husb: (mf)	0.22	0.29	0.26	0.31	0.25
cCN: Wife and Husb: (no use)	0.16	0.00	0.01	0.00	0.09
Sum concordance	0.49	0.50	0.51	0.50	0.50
Discordance:					
cDFM: Wife: (f) $\vee$ (mf); Husb:(m)	0.06	0.08	0.07	0.09	0.07
cDMF: Husb: (f) $\vee$ (mf); Wife:(m)	0.05	0.08	0.07	0.08	0.06
cDOBOF: One Spouse: (f); Other Spouse: (mf)	0.17	0.29	0.30	0.29	0.23
cNM: Wife: (no use); Husb: (m)	0.02	0.00	0.00	0.00	0.01
cNF: Wife: (no use); Husb: (f) $\vee$ (mf)	0.11	0.01	0.01	0.01	0.06
cFN: Wife: (f) $\vee$ (mf); Husb: (no use)	0.08	0.03	0.03	0.03	0.06
cFM: Wife: (m); Husb: (no use)	0.01	0.00	0.00	0.00	0.01
Sum discordance	0.51	0.50	0.49	0.50	0.50
Observations	2,665	4,228	1,096	3,113	8,662

Notes: **Data:** IFLS-5 wave (2014), cross-sectional data, one observation is one couple; **Decision making domain:** "whether you and your spouse use contraception"; **Typology:** cCM: concordant report of husband as sole decision maker, cCB: concordant report of joint decision making, cDFM: wife perceives female decision making power, husband does not, cDMF: husband perceives female decision making power, wife does not, cDOBOF: both spouses perceive female decision making power, one of them perceives a sole female decision maker, the other joint decision making.

A.4.2. Instrumental variable approach: geographical distribution of shocks

Table A.22.: Distribution of instrument events across provinces

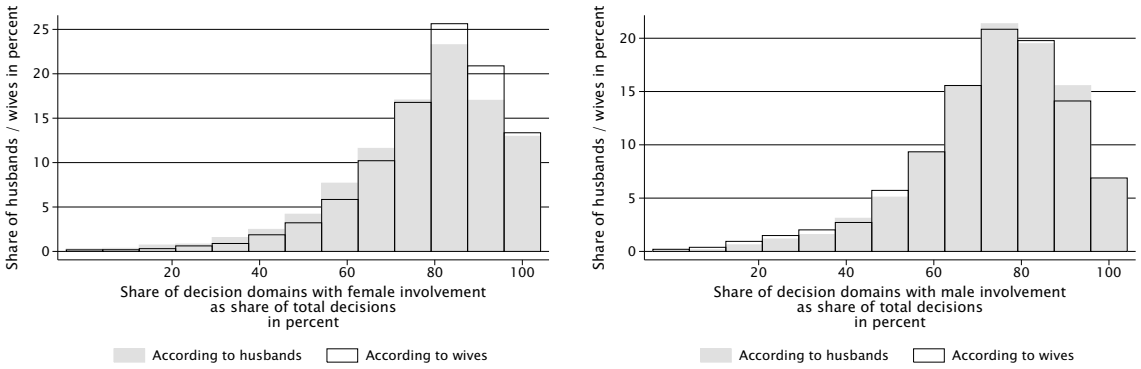
	Household loss due to earthquake fire, or other natural disaster		Crop loss	
	Loss	No Loss	Crop loss	No crop loss
Province	– in percent of column	– in percent of column	– in percent of column	– in percent of column
North_Sumartra	8.4	6.0	6.8	5.9
West_Sumatra	5.6	4.5	3.2	4.7
Riau	0.9	0.3		0.3
South_Sumartra	8.4	4.9	4.3	5.0
Lampung	0.9	4.5	12.1	3.3
Jakarta	10.3	7.4	0.5	8.5
West_Java	19.6	18.0	17.3	18.1
Central_Java	10.3	12.9	15.5	12.5
Yogyakarta	1.9	5.9	5.7	5.8
East_Java	15.0	14.2	8.4	15.0
Bali	3.7	5.6	3.2	5.9
West_Nusa_Tenggara	1.9	6.6	6.6	6.6
South_Kalimantan	7.5	4.6	11.8	3.6
South_Sulawesi	5.6	4.6	4.6	4.6
Central_Kalimantan		0.0		0.1
Observations	107	6,424	808	5,723

Notes: ***Data:*** *IFLS-3 wave (2000), cross-sectional data, one observation is one couple.*

A.5. Robustness checks

A.5.1. Estimates using 2007 data

Figure A.1.: Histograms of perceived female (left graph) and male (right graph) decision-making power with 2007 data



Note: **Data:** IFLS-4 wave (2007), cross-sectional data, one observation is one couple; Graph: two histograms on male and female perception of male and female decision-making power. Grey area: husband's perception. Black lined bars: wife's perception. **Scale:** left (right) graph: decision-making power is captured as number of household decision domains the wife (husband) is involved in over total number of household decision domains; Value 1 on x-axis indicates that wife (husband) has a say in all household decisions.

Table A.23.: Type of statement combinations and labor outcomes with 2007 data

	Dependent variables:		
	(1) Wife working (Yes = 1/ No = 0)	(2) Hours worked (in h worked pa)	(3) Unpaid family worker (Yes = 1/ No = 0)
Concordance:			
CF: Wife and Husb: (f)	0.519***	965.260***	-0.270***
	(0.026)	(151.683)	(0.043)
CB: Wife and Husb: (mf)	0.483***	1046.705***	-0.089***
	(0.021)	(51.778)	(0.027)
CN: Wife and Husb: (none)	0.603***	1513.645***	0.528***
	(0.033)	(85.859)	(0.037)
Discordance:			
DFM: Wife: (f) ∨ (mf); Husb: (m) ∨ (none)	0.278***	540.894***	-0.093***
	(0.014)	(50.542)	(0.021)
DMF: Husb: (f) ∨ (mf); Wife: (m) ∨ (none)	0.199***	375.495***	-0.039*
	(0.025)	(59.323)	(0.022)
DOBOF: One Spouse: (f); Other Spouse: (mf)	0.448***	1008.505***	-0.184***
	(0.024)	(94.427)	(0.027)
DONOM: One Spouse: (none); Other Spouse: (m)	0.014	371.239	-0.307***
	(0.109)	(335.969)	(0.104)
Household attributes:			
Log HH income	0.041***	221.271***	-0.053***
	(0.008)	(25.788)	(0.014)
Constant	-0.334**	-3.4e+03***	1.407***
	(0.119)	(392.292)	(0.190)
Control for (spousal) age and education	✓	✓	✓
Control for N HH adults, N HH children, migration	✓	✓	✓
Region dummies	✓	✓	✓
Observations	7,562	7,562	4,802
R^2	0.22	0.16	0.14
Clusters	19	19	19
Robust standard errors clustered at regional level	Yes	Yes	Yes

Notes: Standard OLS estimate; **Dependent variable:** employment status of wife (binary), reported hours worked per year (weekly hours multiplied with weeks worked) and status as unpaid family worker (binary); **Baseline group:** concordant couples, perceiving husband as the decision maker; Individual level attributes are wife's attributes by default, Spouse (Sp:) refers to husband's attributes; **Data:** IFLS-4 wave (2007), cross-sectional data, one observation is one couple; **Typology:** CM: concordant report of husband as sole decision maker, CF: concordant report of wife as sole decision maker, CB: concordant report of joint decision making; CN: concordant report of neither partner as decision maker, DFM: wife perceives female decision making power, husband does not, DMF: husband perceives female decision making power, wife does not, DOBOF: both spouses perceive female decision making power, one of them perceives a sole female decision maker, the other joint decision making, DONOM: one of the spouses perceives that neither spouse makes the decision, the other spouse perceives male decision maker; Robust standard errors in parentheses are clustered at regional level; ***/**/* indicate significance at the 1%/5%/10% level.

Table A.24.: Type of statement combinations and contraception use with 2007 data

	Dependent variable: Couple uses contraception		
	(1)	(2)	(3)
	No controls	Baseline	Baseline with exclusion restriction 2
Concordance:			
cCM: Wife and Husb: (m)	-0.102	-0.113	-0.112
	(0.089)	(0.077)	(0.089)
cCB: Wife and Husb: (mf)	-0.083*	-0.066	-0.047
	(0.044)	(0.046)	(0.034)
Discordance:			
cDFM: Wife: (f) ∨ (mf); Husb:(m)	-0.060**	-0.018	-0.022
	(0.025)	(0.027)	(0.028)
cDMF: Husb: (f) ∨ (mf); Wife:(m)	-0.041	-0.047	-0.035
	(0.041)	(0.048)	(0.034)
cDOBOF: One Spouse: (f); Other Spouse: (mf)	-0.047	-0.047	-0.037
	(0.043)	(0.040)	(0.031)
Constant	0.843***	0.696***	0.985***
	(0.034)	(0.163)	(0.137)
Income, age, HH and education control		✓	✓
Region dummies		✓	✓
Observations	4,838	4,734	4,371
R^2	0.0048	0.052	0.056
Clusters	17	17	17
Robust standard errors clustered at regional level	Yes	Yes	Yes

Notes: Standard OLS estimate; **Dependent variable:** contraceptive use; **Baseline group:** concordant couples, perceiving wife as the sole decision maker; Individual level attributes are wife's attributes by default, Spouse (Sp:) refers to husband's attributes; **Data:** IFLS-4 wave (2007), cross-sectional data, one observation is one couple; **Decision making domain:** "whether you and your spouse use contraception"; **Typology:** cCM: concordant report of husband as sole decision maker, cCB: concordant report of joint decision making, cDFM: wife perceives female decision making power, husband does not, cDMF: husband perceives female decision making power, wife does not, cDOBOF: both spouses perceive female decision making power, one of them perceives a sole female decision maker, the other joint decision making; Information for exclusion restriction 1 unavailable in 2007; ***/**/* indicate significance at the 1%/5%/10% level.

Table A.25.: Type of statement combinations and their association with covert method use with 2007 data

	Dependent variable: Couple uses covert method conditional on contraceptive use		
	(1) No controls	(2) Baseline	(3) + Exclusion restriction 2
Concordance:			
cCM: Wife and Husb: (m)	-0.165	-0.100	-0.100
	(0.119)	(0.093)	(0.093)
cCB: Wife and Husb: (mf)	0.002	0.003	0.003
	(0.027)	(0.026)	(0.026)
Discordance:			
cDFM: Wife: (f) ∨ (mf); Husb:(m)	-0.000	0.007	0.007
	(0.033)	(0.029)	(0.029)
cDMF: Husb: (f) ∨ (mf); Wife:(m)	0.048	0.029	0.029
	(0.058)	(0.043)	(0.043)
cDOBOF: One Spouse: (f); Other Spouse: (mf)	0.017	0.017	0.017
	(0.027)	(0.029)	(0.029)
Constant	0.715***	0.920***	0.920***
	(0.022)	(0.131)	(0.131)
Income, age, HH and Education control		✓	✓
Region dummies		✓	✓
Observations	3,753	3,671	3,671
R^2	0.0012	0.046	0.046
Clusters	17	17	17
Robust standard errors clustered at regional level	Yes	Yes	Yes

Notes: Standard OLS estimate; **Dependent variable:** covert contraceptive use conditional on contraceptive use; **Baseline group:** concordant couples, perceiving wife as the sole decision maker; Individual level attributes are wife's attributes by default, Spouse (Sp:) refers to husband's attributes; **Data:** IFLS-4 wave (2007), cross-sectional data, one observation is one couple; **Decision making domain:** "whether you and your spouse use contraception"; **Typology:** cCM: concordant report of husband as sole decision maker, cCB: concordant report of joint decision making, cDFM: wife perceives female decision making power, husband does not, cDMF: husband perceives female decision making power, wife does not, cDOBOF: both spouses perceive female decision making power, one of them perceives a sole female decision maker, the other joint decision making; Information for exclusion restriction 1 unavailable in 2007; Robust standard errors in parentheses are clustered at regional level; ***/**/* indicate significance at the 1%/5%/10% level.

A.5.2. Alternative specification: covert methods

Table A.26.: Exclusion of all pill users: types of statement combinations and covert method use

	Dependent variable: Couple uses covert method conditional on contraceptive use				
	(1) No controls	(2) Baseline	(3) + Exclusion restriction 1	(4) + Exclusion restriction 2	(5) + Exclusion restriction 1 and 2
Concordance:					
cCM: Wife and Husb: (m)	-0.264***	-0.272***	-0.317***	-0.272***	-0.317***
	(0.069)	(0.067)	(0.089)	(0.067)	(0.089)
cCB: Wife and Husb: (mf)	-0.046**	-0.046***	-0.048**	-0.046***	-0.048**
	(0.017)	(0.010)	(0.017)	(0.010)	(0.017)
Discordance:					
cDFM: Wife: (f) ∨ (mf); Husb:(m)	-0.043***	-0.044***	-0.029	-0.044***	-0.029
	(0.012)	(0.010)	(0.024)	(0.010)	(0.024)
cDMF: Husb: (f) ∨ (mf); Wife:(m)	-0.046***	-0.060***	-0.069***	-0.060***	-0.069***
	(0.013)	(0.014)	(0.016)	(0.014)	(0.016)
cDOBOF: One spouse: (f); Other spouse: (mf)	-0.017***	-0.017***	-0.021**	-0.017***	-0.021**
	(0.005)	(0.006)	(0.009)	(0.006)	(0.009)
Constant	0.978***	1.281***	1.276***	1.281***	1.276***
	(0.006)	(0.093)	(0.108)	(0.093)	(0.108)
Region dummies		✓	✓	✓	✓
Income, age, HH and education ccntrol		✓	✓	✓	✓
Observations	3,155	2,849	1,612	2,849	1,612
R^2	0.025	0.088	0.082	0.088	0.082
Clusters	18	18	18	18	18
Robust standard errors clustered at regional level	Yes	Yes	Yes	Yes	Yes

Notes: Standard OLS estimate; **Dependent variable:** covert contraceptive use conditional on contraceptive use; **Baseline group:** concordant couples, perceiving wife as the sole decision maker; **Data:** IFLS-5 wave (2014), cross-sectional data, one observation is one couple; **Decision making domain:** "whether you and your spouse use contraception"; **Typology:** cCM: concordant report of husband as sole decision maker, cCB: concordant report of joint decision making, cDFM: wife perceives female decision making power, husband does not, cDMF: husband perceives female decision making power, wife does not, cDOBOF: both spouses perceive female decision making power, one of them perceives a sole female decision maker, the other joint decision making; Robust standard errors in parentheses are clustered at regional level ***/**/* indicate significance at the 1%/5%/10% level.

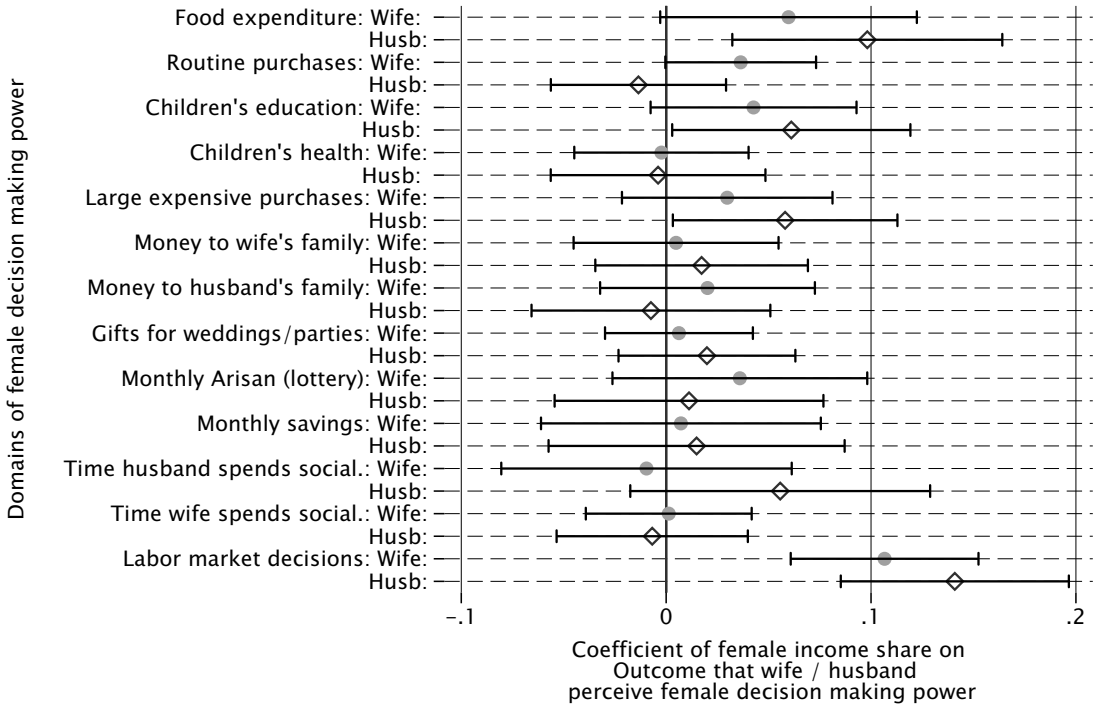
Table A.27.: Classification of pill as covert method: types of statement combinations and covert method use

	Dependent variable: Couple uses covert method conditional on contraceptive use				
	(1)	(2)	(3)	(4)	(5)
	No controls	Baseline	Baseline with exclusion restriction 1	Baseline with exclusion restriction 2	Baseline with exclusion restriction 1 and 2
Concordance:					
cCM: Wife and Husb: (m)	-0.239***	-0.240***	-0.276***	-0.240***	-0.276***
	(0.061)	(0.058)	(0.074)	(0.058)	(0.074)
cCB: Wife and Husb: (mf)	-0.041***	-0.039***	-0.039**	-0.039***	-0.039**
	(0.014)	(0.008)	(0.014)	(0.008)	(0.014)
Discordance:					
cDFM: Wife: (f) ∨ (mf); Husb:(m)	-0.039***	-0.039***	-0.030	-0.039***	-0.030
	(0.010)	(0.009)	(0.021)	(0.009)	(0.021)
cDMF: Husb: (f) ∨ (mf); Wife:(m)	-0.040***	-0.051***	-0.058***	-0.051***	-0.058***
	(0.010)	(0.011)	(0.012)	(0.011)	(0.012)
cDOBOF: One spouse: (f); Other spouse: (mf)	-0.015***	-0.014***	-0.017**	-0.014***	-0.017**
	(0.004)	(0.004)	(0.007)	(0.004)	(0.007)
Constant	0.985***	1.221***	1.204***	1.221***	1.204***
	(0.004)	(0.071)	(0.083)	(0.071)	(0.083)
Region dummies		✓	✓	✓	✓
Income, age, HH and education control		✓	✓	✓	✓
Observations	4,016	3,640	2,047	3,640	2,047
R^2	0.024	0.076	0.069	0.076	0.069
Clusters	18	18	18	18	18
Robust standard errors clustered at regional level	Yes	Yes	Yes	Yes	Yes

Notes: Standard OLS estimate; **Dependent variable:** covert contraceptive use conditional on contraceptive use; **Baseline group:** concordant couples, perceiving wife as the sole decision maker; **Data:** IFLS-5 wave (2014), cross-sectional data, one observation is one couple; **Decision making domain:** "whether you and your spouse use contraception"; **Typology:** cCM: concordant report of husband as sole decision maker, cCB: concordant report of joint decision making, cDFM: wife perceives female decision making power, husband does not, cDMF: husband perceives female decision making power, wife does not, cDOBOF: both spouses perceive female decision making power, one of them perceives a sole female decision maker, the other joint decision making; Robust standard errors in parentheses are clustered at regional level ***/**/* indicate significance at the 1%/5%/10% level.

A.5.3. Alternative specification: household income share

Table A.28.: Fixed effects model: alternative coding of household income share



Note: Alternative household share calculation; Fixed effects estimate; **Data:** IFLS waves 3, 4 and 5, panel data 2000 to 2014, one observation is one couple, **Dependent variable:** wife (husband) states that wife has decision making power in specific domain; Ticks indicate 95 percent confidence interval.

Table A.29.: IV model: alternative coding of household income share

	Dependent variable: female decision-making share from ...			
	(1) ...Wife's perspective (Land-owners only)	(2) ...Husband's perspective (Land-owners only)	(3) ...Wife's perspective (Full sample)	(4) ...Husband's perspective (Full sample)
Female income share	0.481**	0.482**	0.538**	0.574**
	(0.206)	(0.202)	(0.242)	(0.238)
Any land			0.049**	0.061***
			(0.023)	(0.023)
Constant	0.687***	0.771***	0.555***	0.645***
	(0.151)	(0.154)	(0.121)	(0.125)
Control for (spousal) age and education	✓	✓	✓	✓
Control for income, N HH adults, N HH children, migration	✓	✓	✓	✓
Observations	1,499	1,499	3,377	3,377
Underid. F-Statistic	15.7	15.7	17.0	17.0
(p-Value)	0.110	0.110	0.0755	0.0755
Weak id. F-Statistic	9.578	9.578	25.26	25.26
Region dummies	Yes	Yes	Yes	Yes
Robust standard errors	Yes	Yes	Yes	Yes

Notes: Instrumental variable reduced results; **Dependent variable:** female share in household decision making across all decision domains; **Data:** IFLS-3 wave (2000), cross-sectional data, one observation is one couple; Individual level attributes are wife's attributes by default, Spouse (Sp:) refers to husband's attributes

A.5.4. Alternative specification: instrumental variable coding

Table A.30 repeats the instrumental variable estimation with an alternative instrument coding. Before, we coded an indicator variable vector reflecting all possible six reactions [Q, S, T, U, V, W] of table A.6 to the two events (crop loss and natural disasters). We now include an extended list of reactions, namely [O, P, Q, S, T, U, V, W]. Keeping our two possible events, this results in a vector of 16 indicator variables. This instrument is not valid as suggested by the underidentification and overidentification statistics provided in table A.30.

Table A.30.: IV model: reduced form results with alternative set of instruments

	Dependent variable: female decision-making share from ...			
	(1) ...Wife's perspective (Landowners only)	(2) ...Husband's perspective (Landowners only)	(3) ...Wife's perspective (Full sample)	(4) ...Husband's perspective (Full sample)
Female income share	0.457**	0.492***	0.510**	0.437**
	(0.189)	(0.179)	(0.199)	(0.184)
Any land			0.033***	0.034***
			(0.011)	(0.010)
Constant	0.523***	0.632***	0.457***	0.505***
	(0.096)	(0.098)	(0.086)	(0.083)
Control for (spousal) age and education	✓	✓	✓	✓
Control for HH income, N HH adults, N HH children, migration	✓	✓	✓	✓
Observations	2,276	2,276	6,053	6,053
Underid. F-Statistic	22.9	22.9	20.3	20.3
Underid. P-val.	0.0431	0.0431	0.0878	0.0878
Weak id. F-Statistic	16.81	16.81	11.31	11.31
Overid. F-Statistic	9.795	15.87	10.60	26.72
Region dummies	Yes	Yes	Yes	Yes
Robust standard errors	Yes	Yes	Yes	Yes

Notes: Reduced form IV results; **Dependent variable:** female share in household decision making across all decision domains; Individual level attributes are wife's attributes by default, Spouse (Sp:) refers to husband's attributes; alternative coding of instrument; **Data:** IFLS-3 wave (2000), cross-sectional data, one observation is one couple; ***/**/* indicate significance at the 1%/5%/10% level.

A.6. Further analyses

A.6.1. Relative education and power

The four tables A.31, A.32, A.33 and A.34 describe the relationship between perceived power from both partners' perspective and educational achievement of both partners. Tables A.31, A.32, A.33 and A.34 show that holding male education constant, higher female education is associated with higher female power. This is true for the husband's and the wife's perspective. A comparison across tables also allows to see that holding female education constant, higher education of husbands is associated with higher self-perceived power (The group of not formally educated husbands forms an exception in this regard).

We also observe that the male (female) assessment of female (male) power is almost always lower than the female (male) self-assessment. The gap between not formally educated women and women that received a tertiary education is substantial. For wives of husbands with primary or secondary education, the difference amounts to between nine and ten percentage points – from the wife's perspective. This difference is even larger (eleven percentage points) for women whose husbands have received tertiary education.

Table A.31.: Education and power: wives with husbands without formal education

	Husband without formal education, wife with:		
	no formal education	elementary education	secondary education
	–	–	–
	mean share	mean share	mean share
Perception of female power:			
Wife: own decision share	0.65	0.73	0.83
Husb: wife's decision share	0.59	0.63	0.71
Perception of male power:			
Wife: husband's decision share	0.51	0.49	0.63
Husb: own decision share	0.61	0.59	0.71
Observations	109	123	13

Notes: **Data:** IFLS-5 wave (2014), cross-sectional data, one observation is one couple.

Table A.32.: Education and power: wives with husbands with primary education

	Husband with primary education, wife with:			
	no formal education	elementary education	secondary education	tertiary education
	– mean share	– mean share	– mean share	– mean share
Perception of female power:				
Wife: own decision share	0.70	0.70	0.75	0.80
Husb: wife's decision share	0.64	0.68	0.70	0.80
Perception of male power:				
Wife: husband's decision share	0.53	0.57	0.62	0.57
Husb: own decision share	0.63	0.63	0.66	0.60
Observations	252	1,988	794	18

Notes: **Data:** IFLS-5 wave (2014), cross-sectional data, one observation is one couple.

Table A.33.: Education and power: wives with husbands with secondary education

	Husband with secondary education, wife with:			
	no formal education	elementary education	secondary education	tertiary education
	– mean share	– mean share	– mean share	– mean share
Perception of female power:				
Wife: own decision share	0.71	0.72	0.75	0.80
Husb: wife's decision share	0.64	0.69	0.72	0.76
Perception of male power:				
Wife: husband's decision share	0.60	0.61	0.66	0.67
Husb: own decision share	0.71	0.67	0.70	0.69
Observations	24	966	2,867	414

Notes: **Data:** IFLS-5 wave (2014), cross-sectional data, one observation is one couple.

Table A.34.: Education and power: wives with husbands with tertiary education

	Husband with tertiary education, wife with:			
	no formal education	elementary education	secondary education	tertiary education
	– mean share	– mean share	– mean share	– mean share
Perception of female power:				
Wife: own decision share	0.44	0.66	0.77	0.82
Husb: wife's decision share	0.67	0.69	0.72	0.78
Perception of male power:				
Wife: husband's decision share	0.58	0.65	0.69	0.72
Husb: own decision share	0.78	0.72	0.73	0.75
Observations	3	50	466	574

Notes: **Data:** IFLS-5 wave (2014), cross-sectional data, one observation is one couple.

A.6.2. Couple attributes associated with the perception of female power

Table A.35 estimates the relationship between the overall female decision making share and individual as well as household attributes. We do so for both perspectives, the husband's and the wife's. We run four distinct models for four variables of particular interest: female income share, dowry and female ownership over the family's house or the family's business. We find that from the wife's perspective, female business ownership is positively associated with female decision making power, while it is not from the husband's perspective. From both perspectives, female education is positively associated with female decision making power as is the overall household income.

Table A.35.: Four potential sources of female power from male and female perspectives

	Dependent variable: Female perception of female power Independent variable of interest:				Dependent Variable: Male perception of female power Independent variable of interest			
	(1) Female income share	(2) Dowry	(3) Female owner house	(4) Female owner business	(5) Female Income Share	(6) Dowry	(7) Female Owner House	(8) Female Owner Business
Individual attributes:								
Age	0.002***	0.003***	0.001	0.002**	0.001	0.003*	-0.001	0.000
Spousal age	-0.002***	-0.002***	-0.001	-0.002***	-0.000	-0.000	0.001	0.000
Elementary education	0.013*	-0.010	0.032	0.021	0.032***	0.038	0.037	0.019
Any secondary education	0.038***	0.022	0.078***	0.041**	0.045***	0.063	0.060	0.034
Any college education	0.068***	0.061	0.132***	0.070***	0.079***	0.095*	0.096**	0.065***
Sp: Elementary education	-0.025*	-0.019	-0.037*	-0.054**	0.036	0.064	0.035	0.005
Sp: Any secondary education	-0.018	-0.007	-0.038	-0.046**	0.048*	0.065	0.041	0.011
Sp: Any college education	-0.017	0.001	-0.040	-0.047*	0.041	0.058	0.041	-0.012
Household attributes:								
Log HH income	0.021***	0.019***	0.018***	0.015***	0.017***	0.012	0.011	0.017***
N HH members	-0.011***	-0.016**	-0.013***	-0.011***	-0.013***	-0.009	-0.006	-0.009***
N HH children	0.019***	0.042***	0.015**	0.017***	0.025***	0.034***	0.015	0.017***
Migration indicator	-0.014**	-0.015	-0.009	-0.011	-0.010**	-0.007	-0.011	-0.022*
Multi-religious HH	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Variables of interest:								
Female income share	0.089***				0.076***			
Log Dowry		0.003				0.003		
Female house ownership			0.000				0.000	
Female business ownership				-0.021***				-0.016
Constant	0.381***	0.321***	0.435***	0.567***	0.344***	0.242	0.467***	0.419***
Regional FE	✓	✓	✓	✓	✓	✓	✓	✓
Observations	7,302	1,988	1,992	3,035	7,302	1,988	1,992	3,035
R^2	0.11	0.12	0.14	0.080	0.10	0.11	0.15	0.069
Clusters	19	19	17	18	19	19	17	18
SEs Clustered at Regional Level	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Standard OLS estimate; **Data:** IFLS-5 wave (2014), cross-sectional data, one observation is one couple; Definition: Overall female/male decision making share: Number of decisions with female/male involvement divided by number of overall decisions.

A.6.3. Cross tabulation of labor supply and contraceptive use domains

Table A.36.: Cross tabulation of labor supply and contraceptive use domains

	Labor supply *Concordance* in percent of column			Labor supply *Discordance* in percent of column			Full sample
Contraceptive use	CM	CF	CB	DFM	DMF	DOBOF	
Concordance:							
cCM	5.3		0.7	1.4	1.7		2.0
cCF	26.2	65.0	12.3	19.0	18.4	27.3	19.0
cCB	17.4	8.3	49.3	25.3	27.7	21.7	32.1
Discordance:							
cDFM	13.1	5.0	3.8	17.8	3.6	5.3	9.2
cDMF	13.1		4.5	5.6	16.0	11.0	8.4
cDOBOF	24.9	21.7	29.4	31.0	32.6	34.8	29.3
Observations	1,492	60	2,415	1,560	776	374	6,688

Notes: **Data:** IFLS-5 wave (2014), cross-sectional data, one observation is one couple; **Typology:** CM: concordant report of husband as sole decision maker, CF: concordant report of wife as sole decision maker, CB: concordant report of joint decision making; CN: concordant report of neither partner as decision maker, DFM: wife perceives female decision making power, husband does not, DMF: husband perceives female decision making power, wife does not, DOBOF: both spouses perceive female decision making power, one of them perceives a sole female decision maker, the other joint decision making, DONOM: one of the spouses perceives that neither spouse makes the decision, the other spouse perceives male decision maker; One observation is one couple.

A.6.4. Presence of spouse at interview

This section provides insight into the endogenous relationship between the presence of spouse at the other spouse's interview and reported decision making power and outcomes. We assume that this relationship is endogenous because we expect that the presence of a spouse is not independent of couples' unobserved features.

able A.37 relates statement combinations to the average reported decision-making power of husbands and wives. Tables A.38 and A.39 do so for the labor supply and contraceptive decision domain respectively. The first column of both tables presents average values for all couples at which both spouses were present at the interview of their spouse. Columns two and three present equivalent values for only one spouse being present at the other spouse's interview with the other spouse being interviewed alone. The last column presents results for the case that both spouses were alone during their interviews.

In table A.37, we find that the average rates of total discordance across all dimensions are highest where neither spouse was present during the other spouse's interview. Compared to cases in which both partners are present at each other's interview, we find less frequent concordant reporting of joint decision making and higher rates of any form of discordance. In table A.38, we find similar patterns for the specific question on labor supply decision making. There is no clear emerging picture concerning the relationship between spouse presence and rates of discordance in the domain of contraceptive use decision making (see table A.39). In table A.40, we find that the reported rates of contraceptive use and covert method use do not differ at a statistically significant level between women whose partners were present and those whose partners were not present during their interview. In table A.41, we find that the reported rates of overall female decision shares are higher if husbands were present during the wives' interviews. Same holds for the husband's reported decision share from the wife's perspective. The husband's self-reported decision share is uninformed by the fact whether he was or was not present during his wife's interview.

Table A.37.: Concordance and discordance by presence of spouse: averages across all domains

	(1) Both present at both interviews	(2) Wife at husb.'s interview Wife alone at own interview	(3) Husb. at wife's interview Husb. alone at own interview	(4) No presence of spouse	(5) Full sample
	– in percent of column	– in percent of column	– in percent of column	– in percent of column	– in percent of column
Concordance:					
CM: Wife and Husb: (m)	0.08	0.10	0.09	0.09	0.09
CF: Wife and Husb: (f)	0.08	0.10	0.09	0.10	0.09
CB: Wife and Husb: (mf)	0.34	0.30	0.31	0.29	0.31
CN: Wife and Husb: (none)	0.05	0.05	0.05	0.05	0.05
Sum concordance	0.54	0.54	0.54	0.53	0.54
Discordance:					
DFM: Wife: (f) ∨ (mf); Husb: (m) ∨ (none)	0.16	0.16	0.16	0.17	0.16
DMF: Husb: (f) ∨ (mf); Wife: (m) ∨ (none)	0.12	0.12	0.13	0.12	0.12
DOBOF: One spouse: (f); Other spouse: (mf)	0.17	0.17	0.17	0.17	0.17
DONOM: One spouse: (none); Other spouse: (m)	0.01	0.01	0.01	0.01	0.01
Sum discordance	0.46	0.46	0.46	0.47	0.46
Observations	2,803	2,258	1,195	2,406	8,662

Notes: **Data:** IFLS-5 wave (2014), cross-sectional data, one observation is one couple; **Typology:** CM: concordant report of husband as sole decision maker, CF: concordant report of wife as sole decision maker, CB: concordant report of joint decision making; CN: concordant report of neither partner as decision maker, DFM: wife perceives female decision making power, husband does not, DMF: husband perceives female decision making power, wife does not, DOBOF: both spouses perceive female decision making power, one of them perceives a sole female decision maker, the other joint decision making, DONOM: one of the spouses perceives that neither spouse makes the decision, the other spouse perceives male decision maker

Table A.38.: Concordance and discordance by presence of spouse: labor supply domain

	(1) Both present at both interviews	(2) Wife at husb.'s interview Wife alone at own interview	(3) Husb. at wife's interview Husb. alone at own interview	(4) No presence of spouse	(5) Full sample
	– in percent of column	– in percent of column	– in percent of column	– in percent of column	– in percent of column
Concordance:					
CM: Wife and Husb: (m)	0.21	0.24	0.20	0.21	0.21
CF: Wife and Husb: (f)	0.01	0.01	0.01	0.01	0.01
CB: Wife and Husb: (mf)	0.40	0.34	0.39	0.36	0.37
CN: Wife and Husb: (none)	0.00	0.00	0.00	0.00	0.00
Sum concordance	0.61	0.58	0.60	0.57	0.59
Discordance:					
DFM: Wife: (f) $\vee$ (mf); Husb: (m) $\vee$ (none)	0.22	0.24	0.21	0.24	0.23
DMF: Husb: (f) $\vee$ (mf); Wife: (m) $\vee$ (none)	0.11	0.12	0.13	0.12	0.12
DOBOF: One spouse: (f); Other spouse: (mf)	0.06	0.06	0.07	0.07	0.06
DONOM: One spouse: (none); Other spouse: (m)	0.00	0.00	0.00	0.00	0.00
Sum discordance	0.39	0.42	0.40	0.43	0.41
Observations	2,803	2,258	1,195	2,406	8,662

Notes: **Data:** IFLS-5 wave (2014), cross-sectional data, one observation is one couple; **Typology:** CM: concordant report of husband as sole decision maker, CF: concordant report of wife as sole decision maker, CB: concordant report of joint decision making; CN: concordant report of neither partner as decision maker, DFM: wife perceives female decision making power, husband does not, DMF: husband perceives female decision making power, wife does not, DOBOF: both spouses perceive female decision making power, one of them perceives a sole female decision maker, the other joint decision making, DONOM: one of the spouses perceives that neither spouse makes the decision, the other spouse perceives male decision maker; Decision on *"whether you / your spouse works"*.

Table A.39.: Concordance and discordance by presence of spouse: decision making on contraceptive use

	(1) Both present at both interviews	(2) Wife at husb.'s interview Wife alone at own interview	(3) Husb. at wife's interview Husb. alone at own interview	(4) No presence of spouse	(5) Full sample
	– in percent of column	– in percent of column	– in percent of column	– in percent of column	– in percent of column
Concordance:					
cCM: Wife and Husb: (m)	0.01	0.02	0.01	0.01	0.02
cCF: Wife and Husb: (f)	0.12	0.17	0.14	0.16	0.15
cCB: Wife and Husb: (mf)	0.27	0.24	0.24	0.24	0.25
cCN: Wife and Husb: (no use)	0.09	0.10	0.08	0.08	0.09
Sum concordance	0.50	0.52	0.47	0.50	0.50
Discordance:					
cDFM: Wife: (f) $\vee$ (mf); Husb:(m)	0.06	0.08	0.07	0.07	0.07
cDMF: Husb: (f) $\vee$ (mf); Wife:(m)	0.06	0.07	0.06	0.07	0.06
cDOBOF: One Spouse: (f); Other Spouse: (mf)	0.22	0.21	0.25	0.24	0.23
cNM: Wife: (no use); Husb: (m)	0.01	0.01	0.01	0.01	0.01
cNF: Wife: (no use); Husb: (f) $\vee$ (mf)	0.07	0.05	0.07	0.05	0.06
cFN: Wife: (f) $\vee$ (mf); Husb: (no use)	0.07	0.06	0.06	0.05	0.06
cFM: Wife: (m); Husb: (no use)	0.01	0.01	0.01	0.01	0.01
Sum discordance	0.50	0.48	0.53	0.50	0.50
Observations	2,803	2,258	1,195	2,406	8,662

Notes: **Data:** IFLS-5 wave (2014), cross-sectional data, one observation is one couple; **Typology:** CM: concordant report of husband as sole decision maker, CF: concordant report of wife as sole decision maker, CB: concordant report of joint decision making; CN: concordant report of neither partner as decision maker, DFM: wife perceives female decision making power, husband does not, DMF: husband perceives female decision making power, wife does not, DOBOF: both spouses perceive female decision making power, one of them perceives a sole female decision maker, the other joint decision making, DONOM: one of the spouses perceives that neither spouse makes the decision, the other spouse perceives male decision maker; Decision on *"whether you and your spouse use contraception"*.

Table A.40.: Presence of spouse: contraceptive use

	Contraceptive use and covert methods use by presence of husband during book 4 interview				
	Husband not present	Husband present	Diff.	N Husband not present	N Husband present
Contraceptive use	0.62	0.60	0.02	3,840	3,053
Use of covert method	0.74	0.74	-0.00	2,375	1,834
Total observations	6,893				

Notes: **Data:** IFLS-5 wave (2014), cross-sectional data, one observation is one couple; **Decision making domain:** "whether you and your spouse use contraception"; Covert: Injections, IUD, AKDR, Spiral, Norplant, Implant, Female Sterilization. Non-covert: Pill, Condom, Rhythm, Calendar, Coitus Interruptus, Male Sterilization. Not considered: Femidom, Intravag, Traditional Massage, Traditional Herbs, others.

Table A.41.: Presence of spouse: overall perception of power

	Overall perception of power by presence of husband during interview				
	Husband not present	Husband present	Diff.	N Husband not present	N Husband present
Perception of female power:					
Wife: Own decision share	0.73	0.75	-0.01***	4,664	3,998
Husb: Wife's decision share	0.70	0.71	-0.02***	4,664	3,998
Perception of male power:					
Wife: Husband's decision share	0.61	0.65	-0.03***	4,664	3,998
Husb: Own decision share	0.67	0.68	-0.01	4,664	3,998
Total Observations	8,662				

Notes: **Data:** IFLS-5 wave (2014), cross-sectional data, one observation is one couple; Definition: Overall female/male decision making share: Number of decisions with female/male involvement divided by number of overall decisions.

Appendix B. Domestic violence in Indonesia following two volcano eruptions

B.1. Definitions and coding of variables

Living conditions variable scale
The scale is reversed from the original scaling in the raw data to offer a more intuitive interpretation.

4 = Family lives in conditions that are adequate, as per local standards (defined on community level in consultation with key stakeholders).

3 = Family lives in conditions that are fairly adequate, as per local standards.

2 = Family lives in conditions that are below local standards, but not compromising the personal well-being of individual (and/or family).

1 = Family lives in conditions that are below local standards, and are compromising the personal well-being of individual (and/or family).

Emotional well-being scale
The scale is reversed from the original scaling in the raw data to offer a more intuitive interpretation.

4 = Care-giver is pro-active in addressing the situation of her/his family, and is emotionally stable, with a generally positive outlook.

3 = Care-giver often takes action to address the situation of her/his family, but is struggling to cope with stress or is emotionally unstable.

2 = Care-giver rarely takes action to address the situation of her/his family, and is struggling to cope with stress or is emotionally unstable.

1 = Care-giver is passive (not taking any action to address the situation of her/his family) and/or is emotionally unstable (showing signs of anger, irritability, aggression or depression).

Alcoholism variable

Social workers report whether alcohol and/or drug use negatively affects the family. If either one of the two caregivers abuses alcohol or drugs, it is coded as 1, if none of the two abuse alcohol or drugs, it is coded as 0.

B.2. Auxiliary tables

B.2.1. Household expenditure data and treatment status by regency/city

Table B.1 shows total household expenditure figures in IDR of all regencies/cities considered in the study. Table B.2 shows the development of household expenditure figures in IDR over time and the assignment of regions to treatment and control status. Table B.2 displays household expenditure growth over time in columns one and two. Numbers are based on World Bank data. In column three the table indicates whether a given regency/city is located on Java or not. Column four shows whether a regency is part of the treatment group. The treatment group is located in the areas of DI Yogyakarta and Yogyakarta. Regencies that are on Java but not in the treatment group form the control group. Regencies that are neither on Java nor in the treatment group are not part of the main analysis but part of the synthetic control approach. Regencies and cities on Java (see column 3 for information on location) display a bandwidth ranging from -14 percent to positive 13 percent for 2013 on 2014. The rural DI Yogyakarta region around the city of Yogyakarta (Yogyakarta, Kota) displays the lowest overall growth figure of negative 14 percent. It is also the region in which a large share of the treatment group is located.

Table B.1.: Household expenditure by year and by regency/city

	Household expenditure per capita in IDR			
	2011	2012	2013	2014
Regency/city				
Aceh Barat, Kab.	540,932	638,680	632,962	684,435
Aceh Besar, Kab.	634,162	713,185	648,954	713,994
Bogor, Kab.	600,637	738,744	683,325	774,349
Bogor, Kota	763,232	816,762	811,084	883,508
DI Yogyakarta, Prop.	649,901	721,349	692,732	748,303
Jawa Barat, Prop.	587,951	674,459	680,911	781,065
Medan, Kota	717,967	861,019	976,107	957,268
Semarang, Kab.	522,294	661,918	630,186	739,148
Semarang, Kota	749,405	760,646	1,023,720	1,058,218
Sikka, Kab.	332,327	379,175	447,326	462,701
Tabanan, Kab.	755,171	830,672	838,093	993,577
Yogyakarta, Kota	913,793	904,525	1,088,371	940,194
Observations	26,879	26,879	26,879	26,879

Notes: **Data:** Indonesia Database for Policy and Economic Research (INDO-DAPOER), World Bank Group; Total household expenditure per capita over time in IDR; Abbreviations Kab and Kota refer to regencies (*Kabupaten*) and cities (*Kota*); Prop. refers to areas with available province-level data, only.

Table B.2.: Household expenditure growth (in IDR) and treatment/control status by regency/city

	Household expenditure per capita		Treatment status	
Regency/city	Growth 2011 to 2014 in percent	Growth 2013 to 2014 in percent	Share of group living on Java in percent	Share of group in treatment group in percent
Aceh Barat, Kab.	0.27	0.08	0.00	0.00
Aceh Besar, Kab.	0.13	0.10	0.00	0.00
Bogor, Kab.	0.29	0.13	1.00	0.00
Bogor, Kota	0.16	0.09	1.00	0.00
DI Yogyakarta, Prop.	0.15	0.08	1.00	1.00
Jawa Barat, Prop.	0.33	0.15	1.00	0.00
Medan, Kota	0.33	-0.02	0.00	0.00
Semarang, Kab.	0.42	0.17	1.00	0.00
Semarang, Kota	0.41	0.03	1.00	0.00
Sikka, Kab.	0.39	0.03	0.00	0.00
Tabanan, Kab.	0.32	0.19	0.00	0.00
Yogyakarta, Kota	0.03	-0.14	1.00	1.00
Total	0.29	0.05	0.45	0.25
Observations	26,879	26,879	26,879	26,879

Notes: **Data:** Indonesia Database for Policy and Economic Research (INDO-DAPOER), World Bank Group; Assignment to Java / Non Java groups and treatment group according to geographical location; Abbreviations Kab and Kota refer to regencies (*Kabupaten*) and cities (*Kota*); Prop. refers to areas with available province level data, only.

B.2.2. Household expenditure data and treatment status by SOS Children's Villages community

Table B.3 shows the development of household expenditure at the level of SOS Children's Villages communities. The first letter indicates the programme with which a community is associated (see also table notes).

Table B.3.: Java only: Household expenditure growth over time by SOS Children's Villages community and by treatment status

	Household expenditure per capita			Treatment status
SOS Children's Villages community name	Household expend. 2014 in IDR	Growth 2013 to 2014 in percent	Growth 2011 to 2014 in percent	Share of group in treatment group in percent
J: Bogor (Summarized)	854,322.24	0.10	0.19	0.00
L: Andesde	748,303.42	0.08	0.15	1.00
L: Cakrawala	748,303.42	0.08	0.15	1.00
L: Huntap	748,303.42	0.08	0.15	1.00
L: Huntap Glagaharjo	748,303.42	0.08	0.15	1.00
L: Jayagiri	781,065.46	0.15	0.33	0.00
L: Jogoluhur	748,303.42	0.08	0.15	1.00
L: Kota Jogja	940,194.33	-0.14	0.03	1.00
L: Pakridhan Yogawidagdo	748,303.42	0.08	0.15	1.00
L: Puri Manunggal B.A.B.	748,303.42	0.08	0.15	1.00
L: Pusaka	748,303.42	0.08	0.15	1.00
L: Sapa Jiwa	748,303.42	0.08	0.15	1.00
L: Seputar	781,065.46	0.15	0.33	0.00
L: Sumedang	781,065.46	0.15	0.33	0.00
L: Turus Becik	748,303.42	0.08	0.15	1.00
L: Wangunsari	781,065.46	0.15	0.33	0.00
S: Ambarawa	739,147.83	0.17	0.42	0.00
S: Sayung	1058217.76	0.03	0.41	0.00
S: Tambak Rejo	1058217.76	0.03	0.41	0.00
Total	843,459.94	0.06	0.22	0.55
Observations	12,169	12,169	12,169	12,169

Notes: **Data:** Indonesia Database for Policy and Economic Research (INDO-DAPOER), World Bank Group; Total household expenditure per capita over time in IDR; The first letter indicates the SOS Children's Village programme with which a programme is associated. J for SOS Social Centre Jakarta, L for SOS Social Centre Lembang, S for SOS Family Strengthening Programme Semarang; Abbreviations Kab and Kota refer to regencies (*Kabupaten*) and cities (*Kota*); Prop. refers to regions on province level.

B.3. Descriptive statistics

B.3.1. Number of observations by quarter (full sample and Java only)

Table B.4 presents the total number of observations over time for the full sample. This sample is used for the synthetic control approach. Table B.5 presents the subsample of communities on Java. This sample is used for the difference-in-differences estimation.

Table B.4.: Number of observations by quarter: full sample

Quarters	Number total observations	Relative share in percent	Cumulative share in percent
1	1,989	7.4	7.4
2	2,024	7.5	14.9
3	2,029	7.5	22.5
4	2,027	7.5	30.0
5	2,025	7.5	37.6
6	2,024	7.5	45.1
7	2,019	7.5	52.6
8	2,018	7.5	60.1
9	2,018	7.5	67.6
10	1,931	7.2	74.8
11	1,737	6.5	81.3
12	1,715	6.4	87.6
13	1,685	6.3	93.9
14	1,638	6.1	100.0
Observations	26,879		

Notes: One observation is one family; Quarter 1/2/3/4 = data collected during first, second, third, fourth quarter of 2013 respectively; Quarter 5/6/7/8 = data collected during first, second, third, fourth quarter of 2014 respectively; Quarter 9/10/11/12 = data collected during first, second, third, fourth quarter of 2015 respectively; Quarter 13/14 = data collected during first and second quarter of 2016 respectively; For the estimations, I only consider families that have been admitted to the support program before 2013 and for which reports are given until at least 2015; All observations of families that do not fulfill these prerequisites are dropped.

Table B.5.: Number of observations by quarter: Java only

Quarters	Number observations Java only	Relative share in percent	Cumulative share in percent
1	897	7.4	7.4
2	906	7.4	14.8
3	910	7.5	22.3
4	909	7.5	29.8
5	908	7.5	37.2
6	906	7.4	44.7
7	901	7.4	52.1
8	900	7.4	59.5
9	900	7.4	66.9
10	831	6.8	73.7
11	813	6.7	80.4
12	815	6.7	87.1
13	808	6.6	93.7
14	765	6.3	100.0
Observations	12,169		

Notes: Java only. One observation is one family; Quarter 1/2/3/4 = data collected during first, second, third, fourth quarter of 2013 respectively; Quarter 5/6/7/8 = data collected during first, second, third, fourth quarter of 2014 respectively; Quarter 9/10/11/12 = data collected during first, second, third, fourth quarter of 2015 respectively; Quarter 13/14 = data collected during first and second quarter of 2016 respectively; For the estimations, I only consider families that have been admitted to the support program before 2013 and for which reports are given until at least 2015; All observations of families that do not fulfill these prerequisites are dropped.

B.3.2. Descriptive statistics for treatment group by domestic violence status

Table B.6 presents descriptive statistics for those families living in the treatment area. The treatment group is divided into whether there are domestic violence reports or not.

Table B.6.: Descriptive statistics for treatment group by presence of domestic violence

	Families of treatment group with:		
	no domestic violence	domestic violence	Diff.
Outcome variables:			
Share of fam. with domestic vio.	0.00	1.00	-1.00
Log household expenditure per capita	13.58	13.60	-0.02
Living conditions (1 to 4=good)	2.51	2.49	0.02
Emotional well-being (1 to 4=good)	2.67	2.27	0.40***
Share families with alcohol/drug abuse	0.01	0.15	-0.14***
Cargegiver characteristics:			
Caregiver education (1 to 4)	5.41	6.15	-0.74*
Age female caregivers	41.86	37.35	4.52*
Age male caregivers	44.22	41.45	2.77
Share biological children	0.99	1.00	-0.01
Support variables:			
Time since program admission in years	5.52	4.27	1.25***
Support: Food	0.65	0.83	-0.18*
Support: Healthcare	0.79	0.93	-0.13*
Support: Material	0.30	0.56	-0.26***
Support: Economic	0.65	0.66	-0.01
Support: Living conditions	0.57	0.68	-0.11
Support: Psychosocial	0.78	0.93	-0.15*
Support: Childcare	0.76	0.93	-0.17*
Support: Legal	0.45	0.66	-0.21**

Notes: **Data:** Only treatment area; Only data for three last quarters of 2014 (post-treatment); Treatment group divided into families without report of domestic violence (column 1) and with report of domestic violence (column 2) during this period; **Definition variables:** Education caregiver on scale 1 to 4 indicates 1) no formal education, 2) primary education, 3) lower secondary education, and 4) higher achievement than lower secondary education respectively; Support domains indicate support given by SOS in respective domains; See appendix section B.1 for exact other scales and definitions of variables.

B.3.3. Descriptive statistics by location on Java

Table B.7 presents descriptive statistics for those families that live on Java and those that live not on Java. Communities on Java are used for all difference-in-differences estimates. The families that are not living on Java are only used for the synthetic control approach (together with families living on Java).

Table B.7.: Descriptive statistics by location on Java

	Java		Non-Java		Full sample	
	Mean	SD	Mean	SD	Mean	SD
Outcome variables:						
Share of fam. with domestic vio.	0.002	0.047	0.002	0.049	0.002	0.048
Log household expenditure per capita	13.581	0.19	13.241	0.32	13.394	0.32
Living conditions (1 to 4=good)	2.677	0.63	2.368	0.49	2.496	0.57
Emotional well-being (1 to 4=good)	2.843	0.61	2.910	0.30	2.882	0.46
Share families with alcohol/drug abuse	0.006	0.074	0.016	0.13	0.011	0.11
Cargegiver characteristics:						
Caregiver education (1 to 4)	2.609	0.75	2.317	0.52	2.443	0.65
Age female caregivers	46.449	11.0	41.319	9.24	42.723	10.0
Age male caregivers	44.529	10.2	44.759	11.4	44.644	10.8
Share biological children	0.988	0.11	0.997	0.052	0.993	0.083
Support variables:						
Time since program admission in years	5.901	1.99	6.265	1.57	6.101	1.78
Support: Food	0.547	0.50	0.561	0.50	0.555	0.50
Support: Healthcare	0.682	0.47	0.926	0.26	0.816	0.39
Support: Material	0.254	0.44	0.150	0.36	0.196	0.40
Support: Economic	0.482	0.50	0.609	0.49	0.552	0.50
Support: Living conditions	0.514	0.50	0.207	0.41	0.345	0.48
Support: Psychosocial	0.785	0.41	0.434	0.50	0.592	0.49
Support: Childcare	0.760	0.43	0.615	0.49	0.680	0.47
Support: Legal	0.404	0.49	0.305	0.46	0.349	0.48
Observations	2,713		3,329		6,042	

Notes: **Data:** Statistics based on data from first three quarters of 2013; **Definition variables:** Education caregiver on scale 1 to 4 indicates 1) no formal education, 2) primary education, 3) lower secondary education, and 4) higher achievement than lower secondary education respectively; Support domains indicate support given by SOS in respective domains; See appendix section B.1 for exact other scales and definitions of variables.

B.4. Robustness checks

B.4.1. Main specification estimated with male primary caregivers only

Table B.8.: Treatment effects on domestic violence with male primary caregivers only

	Dependent variable: domestic violence		
	(1) No controls	(2) Minimum controls	(3) Full controls (Baseline model)
Treatment	0.012***	0.011***	0.011***
	(0.000)	(0.000)	(0.000)
Constant	0.002***	-0.001	-0.011*
	(0.000)	(0.001)	(0.004)
Time dummies		✓	✓
Programme dummies		✓	✓
Age group dummies		✓	✓
Support dummies			✓
Time since admission and no. of children			✓
Observations	8,773	8,732	8,732
$Adj.R^2$	0.0081	0.015	0.015
Clusters	4	4	4
Family fixed effects	Yes	Yes	Yes
Standard erros clustered at regional level	Yes	Yes	Yes

Notes: Difference-in-differences estimate with family fixed-effects; **Dependent variable:** domestic violence (yes = 1/no = 0); Robust standard errors in parentheses are clustered at regional level; ***/**/* indicate significance at the 1%/5%/10% level.

B.4.2. Alternative outcomes estimated with fixed effects

Table B.9.: Treatment effects on alternative outcomes with fixed effects

	Alternative outcomes: (fixed effects)			
	(1) Household expenditure	(2) Household living conditions	(3) Emotional well-being	(4) Alcohol/ drug abuse
Treatment	-0.054*	-0.060	-0.067**	0.005***
	(0.019)	(0.053)	(0.018)	(0.000)
Time s. Prog. Adm.	-0.007***	0.021	0.003	0.004*
	(0.001)	(0.010)	(0.018)	(0.002)
Constant	13.651***	2.595***	2.953***	-0.031*
	(0.015)	(0.054)	(0.200)	(0.012)
Time dummies	✓	✓	✓	✓
Programme dummies	✓	✓	✓	✓
Age group dummies	✓	✓	✓	✓
No. of children	✓	✓	✓	✓
Observations	7,207	11,134	11,583	12,128
$Adj.R^2$	0.37	0.018	0.057	0.010
Clusters	4	4	4	4
Family level fixed effects	Yes	Yes	Yes	Yes
Standard erros clustered at regional level	Yes	Yes	Yes	Yes

Notes: Difference-in-differences estimate with family fixed-effects; **Dependent variable:** column 1: log of household expenditure in IDR (measured at regency-/city-level), column 2: living conditions (measured at micro level, scale of 1 to 4, where 4 is good), column 3: emotional well-being (scale of 1 to 4, where 4 is good), column 4: alcohol/drug abuse affects family (yes = 1/no = 0) respectively, see appendix section B.1 for definitions of variables; Robust standard errors in parentheses are clustered at regional level; ***/**/* indicate significance at the 1%/5%/10% level.

B.4.3. Omission of Kabupaten Gunungkidul region from baseline model

Table B.10.: Omission of Kabupaten Gunungkidul region from baseline model

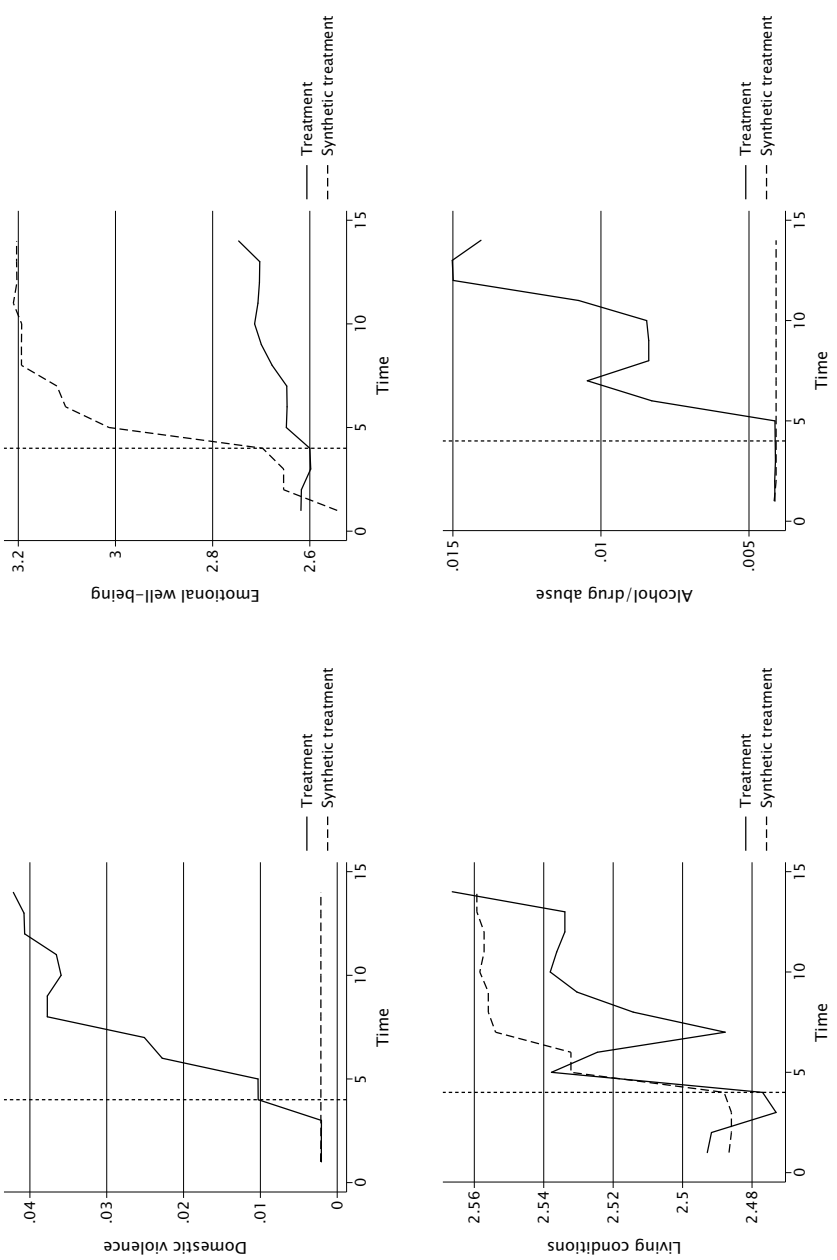
	Fixed effects model			
	(1)	(2)	(3)	(4)
	No controls with smaller sample	Minimum controls with smaller sample	Full controls with smaller sample	Full controls with full sample (Baseline Model)
Treatment	0.033***	0.024***	0.024***	0.022***
	(0.000)	(0.003)	(0.002)	(0.002)
Constant	0.002***	-0.098**	-0.282**	-0.217***
	(0.000)	(0.028)	(0.049)	(0.026)
Quarter dummies		✓	✓	✓
Programme dummies		✓	✓	✓
Age group dummies		✓	✓	✓
Support dummies			✓	✓
Time s. Adm. and No. of Children			✓	✓
Observations	10,305	10,270	10,270	12,128
R^2	0.016	0.077	0.091	0.069
$Adj.R^2$	0.0160	0.0750	0.0888	0.0673
Clusters	4	4	4	4
Family FE	Yes	Yes	Yes	Yes
SEs clustered at regional level	Yes	Yes	Yes	Yes

Notes: Difference-in-differences estimate with family fixed-effects; **Dependent variable:** domestic violence (yes = 1/no = 0); Columns one to three without Kabupaten Gunungkidul. Column four presents baseline model with full sample; Robust standard errors in parentheses are clustered at regional level; ***/**/* indicate significance at the 1%/5%/10% level.

B.4.4. Synthetic control approach: verifying robustness of previous results

I employ a synthetic control approach to create a control group that allows evaluating how the treated region could have developed if it would not have been affected by a volcano eruption. To conduct the analysis, the sample is collapsed on the regency/city level. The primary caregiver's age group, the level of primary caregiver education, the primary caregiver sex, living conditions and household expenditure are used as pre-treatment regressors to approximate a synthetic control group. Synthetic control estimations are run for all outcome variables, except for the household expenditure macro data (This is only available on annual level). The pre-treatment period is defined as the first 3 quarters of 2013. The treatment period is defined as the first quarter of 2014 (The last quarter of 2014 is spared due to its exposure to the first eruption of Mount Merapi). The finding is confirming the hypothesis, that there was indeed an adverse development caused by the eruption.

Figure B.1.: Synthetic control approach based estimation of outcomes



Notes: Synthetic control results. Synthetic treatment (control group) marked by dashed line; treatment group marked by black line; **Dependent variable:** upper left graph: domestic violence (yes = 1/no = 0), upper right graph: emotional well-being (scale of 1 to 4, where 4 is good), lower left graph: living conditions (measured at micro level, scale of 1 to 4, where 4 is good), lower right graph: alcohol/drug abuse affects family (yes = 1/no = 0); Time expressed in quarters. Quarter 1 equals first quarter of 2013.

B.5. Disaster information and map

B.5.1. Detailed account of Mount Merapi eruption

The following report of the Mount Merapi eruption is published by Wunderman (2014b) and describes the eruption and its consequences. Begin of quote (Italics, bold text by Wunderman (2014b)):

> "***Actvitiy during 2014.*** BPPTK noted that on 17 January at 1615, a white plume rose to 50 m above the summit, heading E. At 1854 LT on 10 March 2014, Merapi erupted forming an ash plume that blew W. The event was captured on an automated closed-circuit video (CCTV Pasarbubar) and was followed by two more blasts within a minute (the first at 1855). At 1908, BPPTK noted a volcanic earthquake (with a maximum amplitude of 20 cm). Another video monitor (CCTV Bubar) recorded brown eruptive columns that rose straight up, reaching up to ~1.5 km above the summit. During 1925 to 1930, the eruption gradually stopped. Around this time, ash fell on several villages including Umbulharjo, Kepuharjo, Sidorejo, and Balerante, areas located ~6–7 km to the S of Merapi. During 14-20 March 2014, thick gas plumes rose to ~600 m above the summit. On 17 March, the BPPTK recorded one such event at 0530. On 27 March 2014, an eruption lasted from 1312 to 1316 LT. The VAAC detected volcanic ash to ~9.8 km altitude, using multi-spectral MTSAT-2 imagery, and the Aviation Color Code was raised to Red. A pilot reported that the "large ash cloud [was] moving NW." Darwin VAAC received a SACS SO2 alert at 2150 for the plume, and atmospheric SO2 gas was detected SE of Merapi. By 2232, the volcanic ash appeared to be dissipating; the advisory was terminated at 0830 on 28 March. The 27 March eruption was the subject of a Jakarta Post news article by Muryanto and Ayuningtyas (2014), who indicated that ash fell in the Kemalang and Balerante Klaten regency and that it was 1 mm thick in some areas. The article also noted an M 5.4 tectonic earthquake that struck ~115 km SE of Malang regency, East Java on 23 March. The ash discharge had apparently been occurring regularly since the 2010 eruption but authorities had not taken this as a sign of an escalation in activity, and they urged locals to remain calm. However, according to the article, Sukiman, a resident of the nearby Deles district, said villagers responded to half an hour of ash falling by hitting "kenton-

gan [bamboo drums] to warn others of the danger." On 15 April, BPPTK reported that a thick white plume rose to a maximum of 300 m above the summit. Several tectonic earthquakes occurred in April 2014. On 18 April at 2033, BPPTK recorded tectonic earthquakes 151 km SW of Merapi at a depth of 10 km. On 19 April, four more tectonic earthquakes occurred between 0800 and 2000, and an earthquake lasting 20 minutes was recorded at 0421 from a station on the peak of Merapi. On 20 April from 0426 to 0440, rumbling could be heard within a radius of 8 km around the volcano. The BPPTK reported that on 20 April at 1600, an ash plume traveled W towards the village of Sewukan, amid foggy conditions. The associated eruption was followed by a widely heard roar and a later thin-to-thick plume rose to 400 m above the summit at 1800. The activity ultimately led to ashfall in Sewukan and in sectors to the SE, S, and SW, up to 15 km away from Merapi's summit. The ash from this eruption was also detected by Darwin VAAC, who stated that the ash plume rose to ~10.7 km and extended ~260 km W to NW. The ash was difficult to distinguish from meteorological clouds, and at 1004 LT on 21 April, the VAAC terminated the advisory. In a news article, Minggu (2014) added further details on the eruption omitted here. The BPPTK conducted a field expedition on 22 April to Merapi's crater. The expedition found that the eruption on 20 April had changed the summit crater morphology (figure 63). The slit that cut through the lava dome trending NE had widened by 70 m to the W, and reddish material that the team judged as indicative of oxidation was visible around the center of the lava dome. They also found new eruptive products along the crater's W side and evidence of new growth at the lava dome."

Detailed account of Mount Kelud eruption

The following account of the Mount Kelud eruption has been published by Wunderman (2014a) and describes the eruption. Begin of quote (Italics, bold text by Wunderman (2014a)):

"***Synopsis.*** On 13 February 2014, the Indonesian National Board for Disaster Management (Badan Nasional Penanggulangan Bencana-BNPB) reported that a major eruption occurred at Kelut (also known as Kelud)

volcano in East Java, Indonesia. Ground-based observers had little insight about the ash plume height, but a number of satellite observations helped to constrain the height and other eruption parameters such the direction of plume movement. CALIPSO satellite data revealed that a rapidly rising portion of the plume ejected material up to an altitude exceeding ~26 km, well into the tropical stratosphere. Most of the less rapidly rising portions of the plume remained lower, at 19-20 km altitude. The 2014 eruption destroyed a dome emplaced in the volcano's caldera during the previous eruption in 2007 (*BGVN* 33:03 and 33:07). According to BNPB in a report issued on 18 February 2014, ~7 people were killed and ~100,000 evacuated. At least one commercial aircraft flew into the plume, later landing successfully but incurring costly engine damage.

This report discusses the pre- and syn-eruption observations from the early January through 25 February 2014. Much of the detailed reporting used here describing Kelut's behavior came from the Indonesian Centre for Volcanology and Geological Hazard Mitigation (CVGHM; also known as Pusat Vulkanologi dan Mitigasi Bencana Geologi, PVMBG). Kelut is located just S of Surabaya (Surabaja), Indonesia's second largest city (see a map, figure 8 in *BGVN* 33:03). [...]

According to CVGHM, ash plumes rose to an altitude of 17 km and caused ashfall in areas NE, NW, W, and elsewhere as far as Pacitan (133 km WSW), Kulon Progo (236 km W), Temanggung (240 km WNW), and Banyuwangi (228 km E). As ash began to blanket parts of the region, 40 airline flights were cancelled; impacted airports included Juanda (81 km NE), Adi Sucipto Yogya (208 km W), and Adi Sumarmo Solo (175 km WNW). News articles reported that flights in and out of seven airports were cancelled or rerouted. [...]

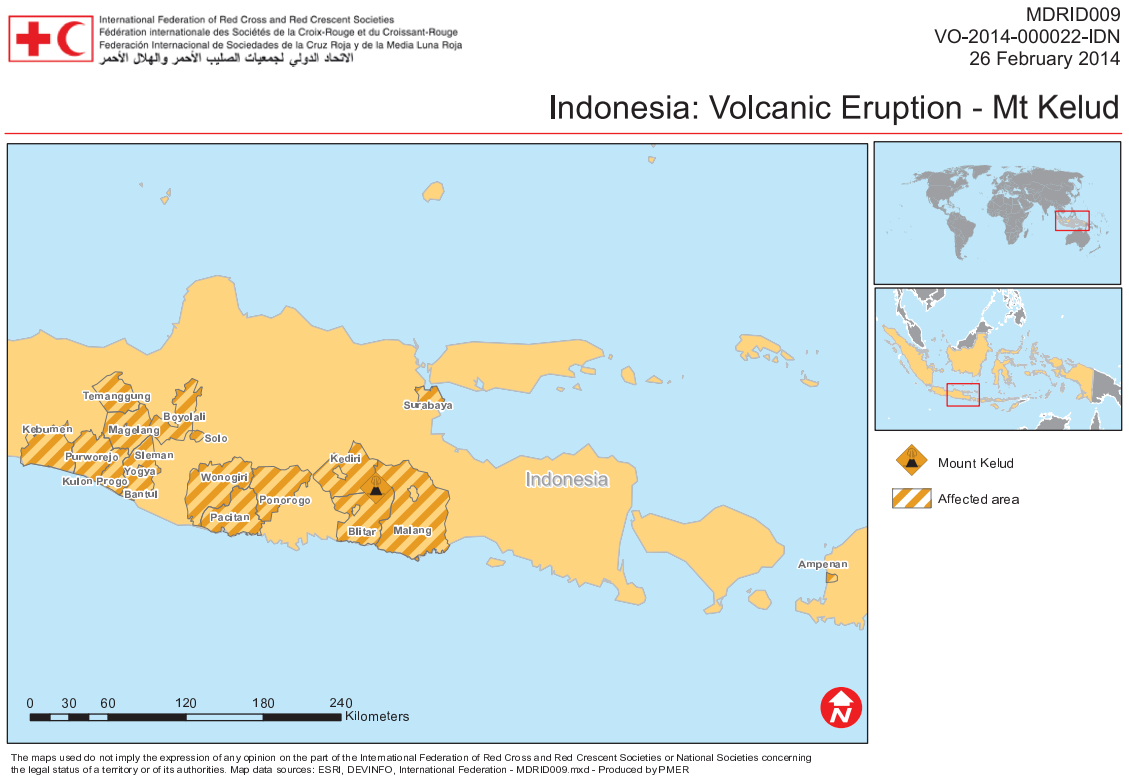
On 14 February BNPB reported that the eruption had killed four people (but later estimates were higher): one died due to a collapsing wall, one from ash inhalation, and two from "shortness of breath." All four victims lived within 7 km of Kelut in the regency of Malang, an area that received ashfall up to 20 cm thick. By 0600 on 14 February, BNPB reported that the number of displaced people reached 100,248, but the report also noted that volcanic activity had declined. Later that day BNPB noted that 76,388 people remained evacuated. Seismicity continued to decline and was at moderate levels during 15-17 February. During 16-20 February white plumes rose as high as 1 km and drifted N, NE, and E. Heavy rain

on 18 February caused lahars in Ngobo, Mangli (Kediri, 35 km WNW), Bladak (Blitar, 20 km SW), and Konto (Malang, 35 km E). BNPB noted that the lahars flooded five houses and one mosque, and destroyed two homes and one bridge. An 18 February BNPB report noted that a total of 7 people in Malang regency had died, and that the ashfall had affected farms, including cattle health and dairy production, and the water supply. Damage to infrastructure in Malang included 3,782 houses, 20 government buildings, 251 schools, 9 hospitals, and 36 churches. Data from satellite instruments provided a 14 February 2014 image on sulfur dioxide (SO2) from Kelut (figure 20). The plume had spread primarily W of the volcano. [...]

Summary of damage. According to the International Federation of Red Cross and Red Crescent Societies (IFRC) (2014), "over the first few days the eruption affected 201,228 people (58,341 families) from 35 villages in three districts: Blitar, Kediri, and Malang. . . As of 14 February 2014, there had been seven fatalities and 70 people in hospitals in serious condition suffering from ash inhalation. Around this time the number of internally displaced persons (IDPs) had reduced to 100,248 people who had evacuated and camped across the province in 172 IDP camps set up to cater for their basic needs." "In addition to the volcanic ash, heavy rain fell and produced cold lahar flooding in Malang, Kediri and Blitar districts. This caused further damage to buildings, farm lands, and roads." Table 4 gives data on damage to structures in the 3 affected districts surrounding Kelut through February 2014. The figures are expected to increase once a more thorough assessment is made. [...]

Lutfi was reported to have stated further that the "impact of Mount Kelud's eruption will extend far beyond the initial cleanup efforts. Fruit farmers reportedly lost more than Rp 24 billion ($2 million) in revenue as ash and debris destroyed whole fields of apples, durian and rambutan that were ready for harvest. The trees, covered in a thick coating of ash, had withered from lack of sunlight."

Figure B.2.: Hazard map of Mount Kelud eruption in 2014



Notes: Map describes Mount Kelud eruption in 2014. Source is International Federation of Red Cross and Red Crescent Societies (2014).

Appendix C. Birth order effects and educational achievement

C.1. Coding of variables

Educational achievement scale

In the database, the main outcome variable is named educational performance. To align with the terminology used in the previous literature, I use the term educational achievement, when referring to educational outcomes in this study. This is done because the description of the raw data fits the definition of educational achievement by York, Gibson, and Rankin (2015).

The scale is reversed from the original scaling in the raw data to offer a more intuitive interpretation.

4 = Outstanding performance=Child is learning very well, and progressing as expected by caregivers, teachers, and other leaders.

3 = Satisfactory performance=Child is learning well, but caregivers, teachers, or other leaders have a few concerns about progress.

2 = Below average performance=Child is learning and progressing poorly or is falling behind.

1 =Poor performance=Child has serious problems with learning.

Note: The scale for South America consists of five instead of only four potential outcomes. The two highest ones are merged based on their wording. Their joint share among all remaining four outcomes is comparable to the other two continents.

C.2. Sampling restrictions

General sampling restrictions. Applying all of the following sampling restrictions reduces the sample size of 53,907 individuals to 4,362 individuals, which I then observe over time. Besides, four countries are excluded.

On country-level, I exclude all countries that show double-peaked grading patterns in educational achievement, that is distributions with the two most common grades being separated by another grade. Local, cultural understanding of grading is expected to be different from the rest of the sample. China, Costa Rica, Ecuador, and Honduras have grading curves with two peaks. For example, China has 67 percent of individuals ranked on the highest and best grade, 5 percent ranked on the second highest grade and then again 17 percent on the third highest grade and 12 percent on the lowest grade (sample average is 25, 54, 19, 2 percent respectively). This informs the notion that in these countries, all lower grades are rather considered to be a punishment than grading on a continuous curve. Baseline model results are not reported but robust to including these countries.[1]

On alternative care family level, I exclude SOS families with either more than 15 or less than two alternative care siblings. The reason for doing so is to find a compromise between including the largest number of *regular families* and excluding so-called Youth Facilities, which are run by SOS but do not operate a traditional alternative care family model.

On biological sibship level, I only include sibships with less than five members. I run alternative specifications for sibship groups with two members. This is done to prevent false identification of non-biological siblings that have been assigned a *placeholder* name that is not the name of the primary caregiver.

On an individual level, I drop a full biological sibship if one of its individuals is bound by any of the following individual-level restrictions. I use a wide age frame of three to 21 years to assure inclusion of all individuals who are receiving schooling. I exclude only children as they are out of focus for birth order analysis. Also, all multiple birth siblings (for example twins) are omitted for the reason that their development is expected to be different from single-birth children with the same birth order rank

1. The baseline model run on those countries only shows significant birth order effects for third-borns, but not for secondborns.

(see Barclay (2015)). I exclude all observations with missing education or age data within the sibship as both are key for all models. To assure accurate identification of biological firstborns, all individuals with siblings outside of SOS Children's Villages care are omitted.

Process for the exclusion of individuals based on conflicting reasons for admission within a sibship. I conduct a quality check on all children who have contradicting reason for admission. If individuals of a sibship differ concerning their reason for admission, this can be a sign of false positive identification on the grounds of a shared surname within an alternative care family. A conflict is assumed in two cases. The first case is given if at least one sibling of an identified sibship is registered as a half-orphan without mother and another sibling of the same sibship is registered as a half-orphan without a father. As both reasons cannot hold true in parallel, all individuals for which this case applies are excluded. The second case is given if the following sequence of events is given. The earliest admitted child of a sibship enters a village on the grounds of loss of both parents. Later, siblings with the same name enter the same family due to loss of only one parent. These statements cannot hold true in parallel in this specific sequence and thus are excluded.

C.3. Auxiliary tables

Table C.1 displays the factorial family model estimation results. The last specification in column 3 is used for the graph in figure 3.3 in the main part of this essay.

Table C.1.: Factorial family model: Relative rank in alternative care family and educational achievement

	(1) Minimum control	(2) Baseline controls	(3) Baseline controls no hardship only
(1,2) Bio. Rank 1, AC Family Relative Rank 2	-0.056	-0.047	-0.095*
	(0.041)	(0.041)	(0.051)
(1,3) Bio. Rank 1, AC Family Relative Rank 3	-0.052	-0.050	-0.046
	(0.039)	(0.038)	(0.048)
(2,1) Bio. Rank 2, AC Family Relative Rank 1	-0.110***	-0.103***	-0.129***
	(0.039)	(0.039)	(0.047)
(2,2) Bio. Rank 2, AC Family Relative Rank 2	-0.151***	-0.146***	-0.165***
	(0.044)	(0.043)	(0.054)
(2,3) Bio. Rank 2, AC Family Relative Rank 3	-0.154***	-0.145***	-0.170***
	(0.050)	(0.050)	(0.061)
(3,1) Bio. Rank 3, AC Family Relative Rank 1	-0.150***	-0.152***	-0.189***
	(0.055)	(0.055)	(0.067)
(3,2) Bio. Rank 3, AC Family Relative Rank 2	-0.089	-0.079	-0.106
	(0.068)	(0.069)	(0.091)
(3,3) Bio. Rank 3, AC Family Relative Rank 3	-0.379**	-0.437**	-0.469**
	(0.178)	(0.192)	(0.200)
Time dummies	✓	✓	✓
Age dummies	✓	✓	✓
Sibship dummies	✓	✓	✓
Baseline model controls		✓	✓
Reason for admission dummies		✓	✓
Observations	26,898	26,898	18,938
$Adj.R^2$	0.65	0.66	0.67
Clusters	4362	4362	2958
Standard errors clustered at individual level	Yes	Yes	Yes

Notes: **Dependent variable:** educational achievement; **Baseline group:** male children of alternative care family relative rank (1,1); Alternative care family relative rank: given the biological birth order rank, position in alternative care family based on age. (1,1) is the oldest biological firstborn of an alternative care family, (1,2) is the second oldest biological firstborn et cetera; Sibship dummies = biological sibship dummies; Minimum controls defined as the control for sibship dummies, quarter dummies, and age dummies only; Robust standard errors in parentheses are clustered at individual level; ***/**/* indicate significance at the 1%/5%/10% level.

C.4. Descriptive statistics

The vast majority of children (90 percent) is between seven and 18 years old. The largest age group are the 11 to 13-year-olds (compare table C.2). Table C.3 presents a cross tabulation of biological birth order and the relative alternative care family rank within the biological birth order. The fact that there are many cases of biologically firstborns falling into the last category (3) of alternative care family rank is attributable to singletons. The relative position within a family is coded before I exclude children that, eg do not have any biological siblings in SOS care.

C.4.1. Age distribution and birth order ranks statistics

Table C.2.: Number of observations by age

	Overview		
	N	Rel. share in percent	Cum. share in percent
3	28	0.1	0.1
4	196	0.7	0.8
5	412	1.5	2.4
6	744	2.8	5.1
7	1,225	4.6	9.7
8	1,787	6.6	16.3
9	2,304	8.6	24.9
10	2,667	9.9	34.8
11	2,832	10.5	45.3
12	2,817	10.5	55.8
13	2,806	10.4	66.2
14	2,487	9.2	75.5
15	1,968	7.3	82.8
16	1,626	6.0	88.9
17	1,263	4.7	93.5
18	884	3.3	96.8
19	511	1.9	98.7
20	273	1.0	99.7
21	68	0.3	100.0
Observations	26,898		

Table C.3.: Overview birth order ranks and relative alternative care family ranks

	Relative alternative care family rank within biological birth order			
Biological birth order	1	2	3	Total
1	1,821	2,139	7,733	11,693
2	5,180	4,288	2,225	11,693
3	2,823	659	30	3,512
Observations	9,824	7,086	9,988	26,898

C.4.2. Descriptive statistics by birth order rank and region

Table C.4.: Descriptive statistics for first-born children

	Attributes of firstborns			
	Africa mean	Latin America mean	Asia mean	Full sample mean
Age at entry	7.69	7.95	7.73	6.35
Age	13.36	14.50	13.49	12.02
Educational achievement	3.08	2.65	3.11	3.03
Gender = female	0.52	0.50	0.58	0.52
No. bio. siblings	2.22	2.40	2.29	2.44
No. all siblings	8.99	7.60	10.15	9.52

Notes: Share of women is statistically different from 50 percent in Asia only (at 5 percent level).

Table C.5.: Descriptive statistics for second-born children

	Attributes of secondborns			
	Africa mean	Latin America mean	Asia mean	Full sample mean
Age at entry	5.25	5.60	5.59	6.35
Age	10.67	12.00	11.04	12.02
Educational achievement	3.12	2.72	3.07	3.03
Gender = female	0.50	0.47	0.48	0.52
No. bio. siblings	2.22	2.40	2.29	2.44
No. all siblings	8.99	7.60	10.15	9.52

Notes: Share of women is statistically different from 50 percent in Asia and Latin-America only (at 5 percent level).

Table C.6.: Descriptive statistics for third- and later-born children

	Attributes of thirdborns or higher			
	Africa mean	Latin America mean	Asia mean	Full sample mean
Age at entry	7.69	7.95	7.73	6.35
Age	13.36	14.50	13.49	12.02
Educational achievement	3.08	2.65	3.11	3.03
Gender = female	0.52	0.50	0.58	0.52
No. bio. siblings	2.22	2.40	2.29	2.44
No. all siblings	8.99	7.60	10.15	9.52

Notes: Share of women is statistically different from 50 percent in Africa only (at 5 percent level).

Table C.7.: Descriptive statistics comparison by region

	Comparsion Africa (1) vs. Latin America (2)			Comparison Africa (1) vs. Asia (2)			Comparsion Latin America (1) vs. Asia (2)		
	Mean (1)	Mean (2)	Diff.	Mean (1)	Mean (2)	Diff.	Mean (1)	Mean (2)	Diff.
Personal characteristics:									
Gender = female	0.50	0.48	0.02	0.50	0.53	-0.02**	0.48	0.53	-0.04***
Age at entry	6.22	6.36	-0.14	6.22	6.37	-0.15**	6.36	6.37	-0.01
Age	11.75	12.65	-0.90***	11.75	11.91	-0.16**	12.65	11.91	0.74***
Family characteristics:									
No. bio. siblings	2.31	2.56	-0.25***	2.31	2.44	-0.13***	2.56	2.44	0.12***
No. all siblings	9.03	7.66	1.37***	9.03	10.17	-1.15***	7.66	10.17	-2.52***
Reason for admission (in percent):									
Death of parents	0.75	0.32	0.42***	0.75	0.69	0.06***	0.32	0.69	-0.37***
Inability caregiver	0.21	0.38	-0.18***	0.21	0.17	0.04***	0.38	0.17	0.22***
Abandonment	0.05	0.25	-0.21***	0.05	0.12	-0.08***	0.25	0.12	0.13***
Hardship experience (in percent):									
Emotional hardship experience	0.12	0.49	-0.37***	0.12	0.09	0.03***	0.49	0.09	0.40***
Financial hardship experience	0.00	0.00	0.00	0.00	0.21	-0.21***	0.00	0.21	-0.21***

C.4.3. Descriptive statistics by reason for admission

Table C.8.: Descriptive statistics for children by experience of hardship

	No hardship prior to admission		Hardship prior to admission		Full sample	
	Mean	SD	Mean	SD	Mean	SD
Outcome variable:						
Educational achievement	3.05	0.72	2.98	0.74	3.03	0.73
Age	12.08	3.46	11.86	3.44	12.02	3.46
Age at entry	6.40	3.13	6.21	3.14	6.35	3.13
Gender	0.51	0.50	0.53	0.50	0.52	0.50
No. bio. siblings	2.43	0.65	2.47	0.69	2.44	0.66
No. all siblings	9.62	2.15	9.27	2.22	9.52	2.17
Lifeshare spent in SOS care	0.46	0.23	0.47	0.24	0.46	0.23
Reason for admission:						
Abandonment	0.11	0.31	0.21	0.41	0.14	0.34
Death of parents	0.70	0.46	0.46	0.50	0.63	0.48
Referral	0.02	0.13	0.02	0.15	0.02	0.14
Inability caregiver	0.18	0.38	0.31	0.46	0.21	0.41
Share of regions:						
Share Africa	0.20	0.40	0.06	0.24	0.16	0.36
Share Latin America	0.14	0.34	0.31	0.46	0.19	0.39
Share Asia	0.67	0.47	0.62	0.48	0.65	0.48
Observations	18,938		7,960		26,898	

Notes: Data describes final sample after application of sampling restrictions.

C.4.4. Distribution of educational achievement by region and country

Table C.9.: Distribution of of educational achievement by region

Educational achievement	Asia	Latin America	Africa	Full sample
1	275	130	87	492
	(1.6)	(2.5)	(2.1)	(1.8)
2	2,706	1,942	562	5,210
	(15.4)	(37.9)	(13.4)	(19.4)
3	9,497	2,339	2,368	14,204
	(54.0)	(45.7)	(56.4)	(52.8)
4	5,099	710	1,183	6,992
	(29.0)	(13.9)	(28.2)	(26.0)
Observations	17,577	5,121	4,200	26,898

Notes: Column percentages in parantheses.

Table C.10.: Educational achievement distribution by country (I/II)

Country	Educational achievement distribution by country					Number of observations by country		
	1	2	3	4	Total	N	Rel. share	Cum. Share
	in percent						in percent	
Angola	0.00	20.00	60.00	20.00	100.00	70	0.26	0.26
Argentina	1.32	36.84	51.32	10.53	100.00	76	0.28	0.54
Bangladesh	0.69	6.39	73.19	19.72	100.00	720	2.68	3.22
Benin	8.13	30.08	44.72	17.07	100.00	123	0.46	3.68
Bolivia	3.34	35.72	43.61	17.33	100.00	1,408	5.23	8.91
Botswana	0.00	24.24	12.12	63.64	100.00	33	0.12	9.03
Cambodia	0.32	12.77	33.55	53.35	100.00	1,237	4.60	13.63
Centr. Afr. Rep.	10.00	40.00	50.00	0.00	100.00	90	0.33	13.97
Chad	0.00	0.00	100.00	0.00	100.00	18	0.07	14.03
Chile	0.00	39.06	55.79	5.15	100.00	233	0.87	14.90
Colombia	4.25	52.75	39.75	3.25	100.00	400	1.49	16.39
Congo, Dem. Rep.	0.00	8.75	68.75	22.50	100.00	160	0.59	16.98
Cote d'Ivoire	11.11	36.11	25.00	27.78	100.00	36	0.13	17.12
Dominican Rep.	2.91	49.82	38.91	8.36	100.00	275	1.02	18.14
El Salvador	0.82	21.31	37.70	40.16	100.00	122	0.45	18.59
Equat. Guinea	0.00	0.00	26.03	73.97	100.00	146	0.54	19.14
Ethiopia	0.00	0.88	40.98	58.14	100.00	571	2.12	21.26
Gambia	0.00	78.57	21.43	0.00	100.00	14	0.05	21.31
Guatemala	4.35	33.15	48.37	14.13	100.00	184	0.68	21.99
Guinea	7.37	1.05	53.68	37.89	100.00	95	0.35	22.35
Haiti	3.70	14.81	70.37	11.11	100.00	135	0.50	22.85
India	1.55	15.58	58.43	24.44	100.00	7,811	29.04	51.89
Indonesia	0.00	0.00	65.22	34.78	100.00	46	0.17	52.06
Jamaica	23.38	48.05	23.38	5.19	100.00	77	0.29	52.35
Kenya	4.92	6.56	65.57	22.95	100.00	183	0.68	53.03
Laos	1.66	12.31	60.51	25.52	100.00	1,869	6.95	59.97
Liberia	12.50	33.33	42.50	11.67	100.00	360	1.34	61.31
Malawi	0.00	22.73	27.27	50.00	100.00	88	0.33	61.64

Table C.11.: Educational achievement distribution by country (II/II)

Country	Educational achievement distribution by country					Number of observations by country		
	1	2	3	4	Total	N	Rel. share	Cum. share
	in percent						in percent	
Mexico	1.58	72.33	26.09	0.00	100.00	253	0.94	62.58
Namibia	0.00	17.78	40.00	42.22	100.00	45	0.17	62.75
Nepal	1.01	14.58	43.78	40.64	100.00	2,675	9.94	72.69
Nicaragua	2.37	28.46	37.15	32.02	100.00	253	0.94	73.63
Niger	0.00	50.00	50.00	0.00	100.00	20	0.07	73.71
Nigeria	0.00	0.00	93.33	6.67	100.00	45	0.17	73.88
Palestine	0.00	15.38	46.15	38.46	100.00	13	0.05	73.92
Panama	4.40	61.54	34.07	0.00	100.00	182	0.68	74.60
Paraguay	0.60	30.35	46.59	22.46	100.00	837	3.11	77.71
Peru	0.32	30.81	62.24	6.64	100.00	633	2.35	80.07
Philippines	0.00	6.16	61.85	31.99	100.00	422	1.57	81.63
Senegal	0.85	14.41	84.75	0.00	100.00	118	0.44	82.07
Sierra Leone	0.00	3.90	89.22	6.88	100.00	436	1.62	83.69
Somalia	0.00	0.00	78.95	21.05	100.00	19	0.07	83.76
Somaliland	0.00	4.60	81.03	14.37	100.00	174	0.65	84.41
South Africa	4.35	4.35	43.48	47.83	100.00	23	0.09	84.50
Sri Lanka	0.92	14.96	52.82	31.30	100.00	869	3.23	87.73
Swaziland	0.18	14.86	44.57	40.40	100.00	552	2.05	89.78
Thailand	0.00	5.98	68.41	25.61	100.00	535	1.99	91.77
Togo	0.00	15.14	41.08	43.78	100.00	185	0.69	92.46
Venezuela	0.00	25.00	75.00	0.00	100.00	40	0.15	92.61
Vietnam	5.67	34.24	41.13	18.95	100.00	1,393	5.18	97.78
Zambia	0.00	13.07	71.68	15.25	100.00	505	1.88	99.66
Zanzibar	0.00	42.86	57.14	0.00	100.00	70	0.26	99.92
Zimbabwe	0.00	19.05	61.90	19.05	100.00	21	0.08	100.00
Total	1.83	19.37	52.81	25.99	100.00	26,898	100.00	

C.4.5. Distribution of reason for admission by region

Table C.12.: Share of children by reason for admission

Reason for admission	Reason for admission by region			
	Africa in percent of column	Latin America in percent of column	Asia in percent of column	Full sample in percent of column
Abandoned	0.20	0.25	0.21	0.21
Caregivers unable to care for child	0.17	0.01	0.18	0.15
Death of both parents / no responsible caregiver	0.21	0.38	0.17	0.21
Death of father and mother unable to care for a child	0.38	0.06	0.31	0.27
Death of mother and father unable to care for a child	0.05	0.25	0.12	0.14
Referred from another care placement	0.00	0.04	0.02	0.02

Table C.13.: Share of children by detailed reason for admission

Detailed reason for admission	Detailed reason for admission by region			
	Africa in percent of column	Latin America in percent of column	Asia in percent of column	Full sample in percent of column
Emotional hardship:				
Alcohol or drug use of parents	0.16	0.37	0.11	0.19
Caregivers unwilling to care for child	0.61	0.00	0.00	0.04
Child rights violation (physical, sexual, emotional, neglect)	0.23	0.75	0.14	0.34
War or natural disasters or emergencies	0.07	0.00	0.00	0.01
Harmful cultural practices	0.01	0.00	0.10	0.07
Economic hardship:				
Severely economically underresourced households	0.00	0.00	0.75	0.46
Other reasons:				
Other	0.00	0.00	0.01	0.00
Born out of wedlock	0.00	0.00	0.02	0.01
Caregiver in conflict with law or in prison	0.01	0.01	0.03	0.02
Caregiver unable to care for child (disability, illness)	0.10	0.09	0.16	0.13
Family in crisis situation	0.00	0.11	0.00	0.04
No information	0.04	0.00	0.01	0.01
Parents separated or divorced	0.00	0.00	0.12	0.07

Notes: Columns do not add up to 100 percent as some children have multiple reasons as entries. Individuals with any hardship experience only.

C.5. Robustness checks

Tables C.14, C.15, and C.16 show baseline model estimates for different sub-regions. I find that the general effect is driven by countries in Latin America and Asia. In Asia, both main regions display birth order effects. In Latin America, observed birth order effects are driven by one main region.

C.5.1. Baseline model estimation for individual regions

Table C.14.: Baseline model estimation for individual regions

	(1) Baseline model full sample	(2) Baseline model Africa only	(3) Baseline model Latin America only	(4) Baseline model Asia only
Secondborn	-0.079***	0.054	-0.099**	-0.095***
	(0.021)	(0.057)	(0.049)	(0.025)
Thirdborn or higher	-0.091**	0.018	-0.181*	-0.093*
	(0.041)	(0.097)	(0.104)	(0.049)
Gender = female	0.138***	0.088*	0.218***	0.114***
	(0.020)	(0.048)	(0.043)	(0.025)
No. all siblings	-0.003	-0.004	-0.002	-0.003
	(0.003)	(0.006)	(0.010)	(0.003)
Lifeshare spent in SOS care	0.158*	-0.236	0.109	0.206*
	(0.090)	(0.179)	(0.245)	(0.108)
Abandonment	0.087	-0.006	0.124	0.082
	(0.060)	(0.106)	(0.137)	(0.069)
Death of parents	0.103*	-0.022	0.404***	0.070
	(0.059)	(0.122)	(0.150)	(0.071)
Referral	-0.034	0.000	0.245	-0.124
	(0.115)	(.)	(0.213)	(0.132)
Constant	3.515***	3.547***	3.829***	2.302***
	(0.281)	(0.434)	(0.383)	(0.704)
Sibship dummies, time dummies, age dummies	✓	✓	✓	✓
Observations	26,898	4,200	5,121	17,577
$Adj.R^2$	0.66	0.68	0.56	0.66
Clusters	4362	706	915	2741
Standard errors clustered at individual level	Yes	Yes	Yes	Yes

Notes: **Dependent variable:** educational achievement; **Baseline group:** firstborn male children; Sibship dummies; Robust standard errors in parentheses are clustered at individual level; ***/**/* indicate significance at the 1%/5%/10% level.

Table C.15.: Baseline model estimation with Asia in 2 subgroups

	(1) Baseline model Asia only	(2) Baseline model Asia group 1 only	(3) Baseline model Asia group 2 only
Secondborn	-0.095***	-0.092***	-0.101***
	(0.025)	(0.034)	(0.036)
Thirdborn or higher	-0.093*	-0.043	-0.201***
	(0.049)	(0.063)	(0.073)
Gender = female	0.114***	0.064*	0.172***
	(0.025)	(0.034)	(0.034)
No. all siblings	-0.003	-0.002	-0.007**
	(0.003)	(0.004)	(0.004)
Lifeshare spent in SOS care	0.206*	0.223	0.273
	(0.108)	(0.138)	(0.172)
Abandonment	0.082	-0.015	0.146**
	(0.069)	(0.104)	(0.070)
Death of parents	0.070	-0.020	0.198*
	(0.071)	(0.098)	(0.107)
Referral	-0.124	-0.177	-0.080
	(0.132)	(0.164)	(0.150)
Constant	2.302***	4.023***	2.158***
	(0.704)	(0.438)	(0.657)
Sibship dummies, time dummies, age dummies	✓	✓	✓
Observations	17,577	10,486	7,091
$Adj.R^2$	0.66	0.63	0.72
Clusters	2741	1679	1062
Standard errors clustered at individual level	Yes	Yes	Yes

Notes: **Dependent variable:** educational achievement; **Baseline group:** firstborn male children; Sibship dummies; Groups based on cultural cluster – Group 1: India, Nepal. Group 2: Bangladesh, Cambodia, Indonesia, Laos, Philippines, Sri Lanka, Thailand, Vietnam; Robust standard errors in parentheses are clustered at individual level; ***/**/* indicate significance at the 1%/5%/10% level.

Table C.16.: Baseline model with Latin America in 3 subgroups

	(1) Baseline model Latin America only	(2) Baseline model Latin America group 1 only	(3) Baseline model Latin America group 2 only	(4) Baseline model Latin America group 3 only
Secondborn	-0.099**	-0.009	-0.208***	-0.057
	(0.049)	(0.074)	(0.080)	(0.090)
Thirdborn or higher	-0.181*	-0.246*	-0.043	-0.296
	(0.104)	(0.147)	(0.172)	(0.230)
Gender = female	0.218***	0.176***	0.201***	0.284***
	(0.043)	(0.068)	(0.069)	(0.076)
No. all siblings	-0.002	-0.007	0.013	-0.029
	(0.010)	(0.015)	(0.017)	(0.029)
Lifeshare spent in SOS care	0.109	0.670	0.147	-0.476
	(0.245)	(0.427)	(0.399)	(0.290)
Abandonment	0.124	0.331*	0.334	-0.192
	(0.137)	(0.193)	(0.284)	(0.203)
Death of parents	0.404***	0.392**	0.394	0.082
	(0.150)	(0.199)	(0.304)	(0.232)
Referral	0.245	0.372*	1.502***	-0.510***
	(0.213)	(0.194)	(0.392)	(0.187)
Constant	3.829***	3.134***	2.914***	4.739***
	(0.383)	(0.468)	(0.648)	(0.394)
Sibship dummies, time dummies, age dummies	✓	✓	✓	✓
Observations	5,121	2,350	1,764	994
$Adj.R^2$	0.56	0.51	0.62	0.67
Clusters	915	391	327	184
Standard errors clustered at individual level	Yes	Yes	Yes	Yes

Notes: **Dependent variable:** educational achievement; **Baseline group:** firstborn male children; Sibship dummies; Groups based on cultural cluster – Group 1: Argentina, Chile, Bolivia, Peru; Group 2: Brazil, Colombia, Dominican Republic, Haiti, Jamaica, Paraguay, Venezuela; Group 3: El Salvador, Guatemala, Mexico, Nicaragua, Panama; Robust standard errors in parentheses are clustered at individual level; ***/**/* indicate significance at the 1%/5%/10% level.

C.5.2. Alternative specifications for baseline model

Table C.17.: Baseline estimation with different types of dummies

	(1) Sibship dummies	(2) Family dummies	(3) Village dummies	(4) Country dummies	(5) No dummies
Secondborn	-0.079***	-0.065***	-0.065***	-0.046*	-0.065**
	(0.021)	(0.020)	(0.024)	(0.026)	(0.027)
Thirdborn or higher	-0.091**	-0.054	-0.043	-0.002	-0.032
	(0.041)	(0.037)	(0.036)	(0.039)	(0.040)
Gender = female	0.138***	0.118***	0.106***	0.101***	0.104***
	(0.020)	(0.020)	(0.021)	(0.022)	(0.023)
No. all siblings	-0.003	0.004	0.014***	0.014***	0.029***
	(0.003)	(0.004)	(0.005)	(0.005)	(0.005)
Lifeshare spent in SOS care	0.158*	0.015	-0.001	0.072	0.086*
	(0.090)	(0.061)	(0.052)	(0.052)	(0.052)
Constant	3.233***	3.667***	2.431***	3.182***	3.092***
	(0.158)	(0.287)	(0.325)	(0.232)	(0.138)
Time dummies, age dummies	✓	✓	✓	✓	✓
Reason for admission dummies	✓	✓	✓	✓	✓
Sibship dummies	✓				
Family dummies		✓			
Village dummies			✓		
Country dummies				✓	
Observations	26,898	26,898	26,898	26,898	26,898
$Adj.R^2$	0.66	0.55	0.28	0.13	0.041
Clusters	4362	4362	4362	4362	4362
Standard errors clustered at individual level	Yes				
Standard errors clustered at family level		Yes			
Standard errors clustered at village level			Yes		Yes
Standard errors clustered at country level				Yes	

Notes: **Dependent variable:** educational achievement; **Baseline group:** firstborn male children; Robust standard errors in parentheses are clustered at individual level; ***/**/* indicate significance at the 1%/5%/10% level.

Table C.18.: Baseline estimation for each quarter (I/II)

	(1) 2014Q3	(2) 2014Q4	(3) 2015Q1	(4) 2015Q2	(5) 2015Q3
Secondborn	-0.107*	-0.137***	-0.155***	-0.101***	-0.066*
	(0.058)	(0.040)	(0.040)	(0.038)	(0.036)
Thirdborn or higher	-0.122	-0.153*	-0.218***	-0.180**	-0.085
	(0.113)	(0.082)	(0.081)	(0.075)	(0.068)
Gender = female	0.127***	0.146***	0.143***	0.128***	0.155***
	(0.046)	(0.034)	(0.032)	(0.030)	(0.030)
No. all siblings	1.005	0.296**	-0.101	0.434	0.445
	(0.760)	(0.116)	(0.284)	(0.304)	(0.305)
Lifeshare spent in SOS care	0.351	0.221	0.235	0.096	0.066
	(0.285)	(0.189)	(0.171)	(0.166)	(0.152)
Abandonment	0.194	0.001	0.022	0.175*	0.129
	(0.162)	(0.146)	(0.125)	(0.098)	(0.099)
Death of parents	0.254*	0.154	0.100	0.090	0.104
	(0.143)	(0.115)	(0.105)	(0.096)	(0.090)
Referral	-0.197	-0.070	-0.133	-0.084	-0.123
	(0.257)	(0.211)	(0.185)	(0.184)	(0.176)
Constant	-5.806	0.653	3.994	-0.498	-1.130
	(6.349)	(1.417)	(2.838)	(2.852)	(2.889)
Sibship dummies, time dummies, age dummies	✓	✓	✓	✓	✓
Observations	1,074	2,213	2,397	2,651	2,765
$Adj.R^2$	0.55	0.48	0.49	0.49	0.48
Clusters	1074	2213	2397	2651	2765
SEs clustered at individual Level	Yes	Yes	Yes	Yes	Yes

Notes: **Dependent variable:** educational achievement; **Baseline group:** firstborn male children; Robust standard errors in parentheses are clustered at individual level; ***/**/* indicate significance at the 1%/5%/10% level.

Table C.19.: Baseline estimation for each quarter (II/II)

	(1) 2015Q4	(2) 2016Q1	(3) 2016Q2	(4) 2016Q3	(5) 2016Q4
Secondborn	-0.065**	-0.088***	-0.060*	-0.053*	-0.082**
	(0.032)	(0.032)	(0.032)	(0.032)	(0.032)
Thirdborn or higher	-0.082	-0.120*	-0.023	-0.033	-0.097
	(0.063)	(0.064)	(0.065)	(0.066)	(0.064)
Gender = female	0.113***	0.144***	0.161***	0.118***	0.135***
	(0.028)	(0.028)	(0.029)	(0.029)	(0.028)
No. all siblings	0.876	0.917	0.151	-0.053	0.869
	(0.568)	(0.598)	(0.162)	(0.302)	(0.574)
Lifeshare spent in SOS care	0.012	0.082	0.184	0.245*	0.045
	(0.137)	(0.134)	(0.129)	(0.128)	(0.152)
Abandonment	0.115	0.078	0.088	0.158	0.135
	(0.105)	(0.111)	(0.101)	(0.102)	(0.111)
Death of parents	0.132	0.112	0.071	0.077	0.145
	(0.090)	(0.091)	(0.087)	(0.084)	(0.091)
Referral	0.015	0.019	0.005	0.103	0.049
	(0.197)	(0.196)	(0.179)	(0.196)	(0.199)
Constant	-4.334	-4.767	2.365**	4.048	-3.922
	(4.910)	(5.216)	(1.197)	(2.588)	(4.988)
Sibship dummies, time dummies, age dummies	✓	✓	✓	✓	✓
Observations	3,090	3,124	3,186	3,318	3,080
$Adj.R^2$	0.50	0.49	0.49	0.47	0.49
Clusters	3090	3124	3186	3318	3080
SEs clustered at individual level	Yes	Yes	Yes	Yes	Yes

Notes: **Dependent variable:** educational achievement; **Baseline group:** firstborn male children; Robust standard errors in parentheses are clustered at individual level; ***/**/* indicate significance at the 1%/5%/10% level.

C.5.3. Robustness checks of hardship and gender split model

Table C.20.: Hardship model estimation for individual regions

	Extreme hardship experience only			
	(1)	(2)	(3)	(4)
	Full sample	Africa only	Latin America only	Asia only
Secondborn	-0.030	-0.198	-0.081	0.001
	(0.037)	(0.147)	(0.065)	(0.046)
Thirdborn or higher	-0.063	-0.355	-0.303**	0.046
	(0.071)	(0.251)	(0.128)	(0.088)
Gender = female	0.216***	0.103	0.334***	0.127***
	(0.038)	(0.113)	(0.064)	(0.046)
No. all siblings	-0.002	-0.026	-0.008	0.002
	(0.005)	(0.031)	(0.013)	(0.004)
Lifeshare spent in SOS care	0.065	0.298	-0.154	0.052
	(0.183)	(0.551)	(0.314)	(0.233)
Constant	3.588***	3.727***	4.098***	2.033***
	(0.324)	(0.663)	(0.408)	(0.713)
Sibship dummies, time dummies, age dummies	✓	✓	✓	✓
Reason for admission dummies	✓	✓	✓	✓
Observations	7,960	500	2,507	4,953
$Adj.R^2$	0.70	0.59	0.61	0.74
Clusters	1404	113	493	798
Standard errors clustered at individual level	Yes	Yes	Yes	Yes

Notes: **Dependent variable:** educational achievement; **Baseline group:** firstborn male children; Individuals with experience of personal hardship only; Sibship dummies; Robust standard errors in parentheses are clustered at individual level; ***/**/* indicate significance at the 1%/5%/10% level.

Table C.21.: Hardship model: different coding of hardship experience

	No extreme hardship		Extreme hardship		
	(1)	(2)	(3)	(4)	(5)
	Parental death	Other reasons	Economic	Emotional	Financial and emotional
Secondborn	-0.087	-0.091**	0.012	-0.075	-0.030
	(0.063)	(0.046)	(0.052)	(0.050)	(0.037)
Thirdborn or higher	-0.178	-0.019	0.073	-0.179**	-0.063
	(0.117)	(0.083)	(0.108)	(0.088)	(0.071)
Gender = female	0.160***	0.107***	0.125**	0.255***	0.216***
	(0.056)	(0.038)	(0.056)	(0.048)	(0.038)
No. all siblings	0.014	-0.007	0.006	-0.005	-0.002
	(0.009)	(0.007)	(0.006)	(0.007)	(0.005)
Lifeshare spent in SOS care	0.188	0.463**	0.275	-0.210	0.065
	(0.304)	(0.181)	(0.242)	(0.247)	(0.183)
Constant	3.421***	3.053***	1.679**	3.888***	3.588***
	(0.429)	(0.256)	(0.692)	(0.325)	(0.324)
Sibship dummies, time dummies, age dummies	✓	✓	✓	✓	✓
Reason for admission dummies	✓	✓	✓	✓	✓
Observations	4,179	5,664	3,701	4,642	7,960
$Adj.R^2$	0.71	0.71	0.75	0.67	0.70
Clusters	637	964	601	855	1404
Standard errors clustered at individual level	Yes	Yes	Yes	Yes	Yes

Notes: **Dependent variable:** educational achievement; **Baseline group:** firstborn male children; Sibship dummies; Robust standard errors in parentheses are clustered at individual level; ***/**/* indicate significance at the 1%/5%/10% level.

Table C.22.: Gender split model: Asia only baseline model with gender interaction term: interacting gender with being first-born

	(1) Asia	(2) Nepal and India	(3) Rest of Asia
Firstborn=1	0.144***	0.209***	0.037
	(0.032)	(0.043)	(0.046)
Gender = female=1	0.152***	0.147***	0.129***
	(0.031)	(0.042)	(0.044)
Firstborn=1 X Gender = female=1	-0.098**	-0.220***	0.101*
	(0.041)	(0.054)	(0.060)
Thirdborn	0.006	0.035	-0.048
	(0.035)	(0.046)	(0.051)
No. all siblings	-0.003	-0.002	-0.007**
	(0.003)	(0.004)	(0.004)
Lifeshare spent in SOS care	0.208*	0.247*	0.258
	(0.107)	(0.137)	(0.166)
Abandonment	0.080	-0.017	0.147**
	(0.069)	(0.103)	(0.071)
Death of parents	0.070	-0.021	0.191*
	(0.071)	(0.098)	(0.107)
Referral	-0.116	-0.151	-0.087
	(0.132)	(0.165)	(0.160)
Constant	2.165***	3.888***	2.041***
	(0.686)	(0.393)	(0.669)
Sibship dummies, time dummies, age dummies	✓	✓	✓
Observations	17,577	10,486	7,091
$Adj.R^2$	0.66	0.63	0.72
Clusters	2741	1679	1062
Standard errors clustered at individual level	Yes	Yes	Yes

Notes: **Dependent variable:** educational achievement; **Baseline group:** secondborn male children; Sibship dummies; Asia only; Robust standard errors in parentheses are clustered at individual level; ***/**/* indicate significance at the 1%/5%/10% level.

C.5.4. Validity of extrapolation of findings

Björklund, Lindahl, and Plug (2004) argue that in order to extrapolate findings from adoption data to the general population, the following needs to hold: (i) Children need to be as good as randomly assigned to their adoptive families, (ii) they need to be adopted early on in their lives and one needs to assume that (iii) studies on adoptive child-parent relationships can be extrapolated to biological child-parent relationships. In the following section, I discuss whether the assumptions made by Björklund, Lindahl, and Plug (2004) hold in this setting.

(i) Children are not randomly selected into families. Consequently, SOS Children's Villages parents' attributes are not expected to be statistically independent of those of their children. Rather, children and parents have a say in whom they are paired up with and have a trial period of living together. The data do not allow to reject the hypothesis, that the sorting of children and caregivers will moderate birth order effects. However, the rotation of children in and out of the family makes it unlikely that a selection induced by caregiver behavior occurs at this level.

(ii) An additional threat to external validity stems from a comparatively high age of admission. However, the baseline estimation does not suggest that the relative share of life spent in SOS Children's Villages' acts as a significant driver of educational achievement at the five percent level.[2] It is hence not expected that treatment in SOS Children's Villages leads to a reversal in birth order effects per se.[3]

(iii) Björklund, Lindahl, and Plug (2004) assert that adoptive children and their parents, as well as their relationships, shall not carry unobservable traits that lead to bias and consequently systematic deviation from what one would expect in biological settings. Generally, the sample is selected insofar as individuals are expected to disproportionately be of underprivileged socio-economic background, relative to the average population. However, first, this group is of particular interest to policymakers as it is often a target of policy interventions. Furthermore, within-sample differences based on hardship experience are still expected to provide valuable insight. The gender and culture-specific effects found in India support the belief that the sample confirms to

2. If it did, this would imply that longer exposure to biological parents changed later educational achievement.

3. One hypothesis that I cannot reject is that of heterogeneous opposing effects that cancel out. While some children might benefit from an early admission, others might suffer, resulting in a zero net effect.

general cultural patterns found in respective populations. Concerning parental influence, SOS parents are not expected to discriminate between siblings of different birth order actively. As described in-depth in the discussion part, non-biological caregivers are exposed to a *rotation* of children rendering discrimination very unlikely. Exposure to non-discriminatory parents can downward-bias effect sizes compared to biological settings but will not lead to a reversal in signs. Concerning family composition, sample families are of above average size compared to average biological and adoptive families. Growing up in large families is likely to dampen birth order effects, as Härkönen (2014), as well as Zajonc and Sulloway (2007), find. However, considering the consistent insignificance of the alternative care sibship size variable, I conclude that family sizing is not expected to moderate educational achievement in this setting.

Bibliography

Bibliography

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller. 2010. "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program." *Journal of the American Statistical Association* 105 (490): 493–505.

Abadie, Alberto, and Javier Gardeazabal. 2003. "The economic costs of conflict: A case study of the Basque country." *American Economic Review* 93 (1): 113–132.

Adams, Paul R., and Gerald R. Adams. 1984. "Mount Saint Helens's ashfall. Evidence for a disaster stress reaction." *The American Psychologist* 39 (3): 252–260.

Aizer, Anna. 2010. "The Gender Wage Gap and Domestic Violence." *American Economic Review* 100 (4): 1847–59.

———. 2011. "Poverty, Violence and Health: The Impact of Domestic Violence During Pregnancy on Newborn Health." *The Journal of Human Resources* 46 (3): 518–538.

Alesina, Alberto, Benedetta Brioschi, and Eliana La Ferrara. 2016. "Violence Against Women: A Cross-cultural Analysis for Africa." *NBER Working Paper Series* 21901.

Allendorf, Keera. 2007. "Couples' Reports of Women's Autonomy and Health-care Use in Nepal." *Studies in Family Planning* 38 (1): 35–46.

Ambler, Kate, Cheryl Doss, Caitlin Kieran, and Simone Passarelli. 2017. "He Says, She Says Exploring Patterns of Spousal Agreement in Bangladesh." *IFPRI Discussion Paper* 01616.

Ammermueller, Andreas, and Jörn-Steffen Pischke. 2009. "Peer effects in European primary schools: evidence from PIRLS." *Journal of Labor Economics* 27 (3): 315–348.

Anastario, Michael, Nadine Shebab, and Lynn Lawry. 2009. "Increased gender-based violence among women internally displaced in Mississippi 2 years post–Hurricane Katrina." *Disaster Medicine and Public Health Preparedness* 3 (1): 18–26.

Anderberg, Dan, Helmut Rainer, Jonathan Wadsworth, and Tanya Wilson. 2015. "Unemployment and Domestic Violence: Theory and Evidence." *The Economic Journal* 126 (597): 1947–1979.

Anderson, C. Leigh, Travis W. Reynolds, and Mary Kay Gugerty. 2017. "Husband and Wife Perspectives on Farm Household Decision-making Authority and Evidence on Intra-household Accord in Rural Tanzania." *World Development* 90:169–183.

Anderson, Siwan, and Mukesh Eswaran. 2009. "What determines female autonomy? Evidence from Bangladesh." *Journal of Development Economics* 90 (2): 179–191.

Arouri, Mohamed, Cuong Nguyen, and Adel Ben Youssef. 2015. "Natural Disasters, Household Welfare, and Resilience: Evidence from Rural Vietnam." *World Development* 70:59–77.

Ashraf, By Nava, Erica Field, and Jean Lee. 2014. "Household Bargaining and Excess Fertility: An Experimental Study in Zambia." *American Economic Review* 104 (7): 2210–2237.

Baez, Javier, Alejandro *de la* Fuente, and Indhira Santos. 2010. "Do natural disasters affect human capital? An assessment based on existing empirical evidence." *IZA Discussion Paper Series* 5164.

Bagger, Jesper, Javier A. Birchenall, Hani Mansour, and Sergio Urzúa. 2013. "Education, Birth Order, and Family Size." *IZA Discussion Paper Series* 7454.

Bandiera, Oriana, Niklas Buehren, Robin Burgess, Markus Goldstein, Selim Gulesci, Imran Rasul, and Munshi Sulaiman. 2018. "Women's Empowerment in Action: Evidence from a Randomized Control Trial in Africa." *Working Paper.*

Banerjee, Abhijit, Esther Duflo, Nathanael Goldberg, Dean Karlan, Robert Osei, William Parienté, Jeremy Shapiro, Bram Thuysbaert, and Christopher Udry. 2015. "A multifaceted program causes lasting progress for the very poor: Evidence from six countries." *Science* 348 (6236): 1260799.

Barclay, Kieron J. 2015. "Birth order and educational attainment: Evidence from fully adopted sibling groups." *Intelligence* 48:109–122.

Barclay, Kieron J., and Mikko Myrskylä. 2014. "Birth order and physical fitness in early adulthood: Evidence from Swedish military conscription data." *Social Science and Medicine* 123:141–148.

Bardi, Massimo, and Silvana M. Borgognini-Tarli. 2001. "A survey on parent-child conflict resolution: Intrafamily violence in Italy." *Child Abuse and Neglect* 25 (6): 839–853.

Basu, Kaushik. 2006. "Gender and say: A model of household behaviour with endogenously determined balance of power." *Economic Journal* 116 (511): 558–580.

Beaman, Lori, Raghabendra Chattopadhyay, Esther Duflo, Rohini Pande, and Petia Topova. 2009. "Powerful women: does exposure reduce bias?" *The Quarterly Journal of Economics* 124 (4): 1497–1540.

Bech, Per, Lis Raabaek Olsen, Mette Kloller, and Niels Kristian Rasmussen. 2003. "Measuring well-being rather than the absence of distress symtoms: a comparison of the SF-36 Mental Health subscale and the WHO-Five Well-Being Scale." *International Journal of Methods in Psychiatric Research* 12 (2): 85–91.

Becker, Gary S., and H. Gregg Lewis. 1973. "On the Interaction between the Quantity and Quality of Children." *Journal of Political Economy* 81 (2 (Part 2: New Economic Approaches to Fertility)): S279–S288.

Becker, Stan, Fannie Fonseca-Becker, and Catherine Schenck-Yglesias. 2006. "Husbands' and wives' reports of women's decision-making power in Western Guatemala and their effects on preventive health behaviors." *Social Science and Medicine* 62 (9): 2313–2326.

Beckett, Celia, Barbara Maughan, Michael Rutter, Jenny Castle, Emma Colvert, Christine Groothues, Jana Kreppner, Suzanne Stevens, Thomas G. O'Connor, and Edmund J. S. Sonuga-barke. 2006. "Do the Effects of Early Severe Deprivation on Cognition Persist Into Early Adolescence? Findings From the English and Romanian Adoptees Study." *Child Development* 77 (3): 696–711.

Belley, Philippe, and Lance Lochner. 2007. "The Changing Role of Family Income and Ability in Determining Educational Achievement." *Journal of Human Capital* 1 (1): 37–89.

Belmont, Lillian, and Francis A. Marolla. 1973. "Birth Order, Family Size, and Intelligence." *Science* 182 (117): 1096–1101.

Belmont, Lillian, Zena Stein, and Patricia Zybert. 1978. "Child spacing and birth order: effect on intellectual ability in two-child families." *Science* 202 (4371): 995–996.

Berkowitz, Leonard. 1993. *McGraw-Hill series in social psychology. Aggression: Its causes, consequences, and control.* New York, NY: Mcgraw-Hill Book Company.

Bertrand, Marianne, Emir Kamenica, and Jessica Pan. 2015. "Gender Identity and Relative Income within Households." *The Quarterly Journal of Economics:* 571–614.

Biddle, Bruce J., and David C. Berliner. 2002. "Small class size and its effects." *Educational Leadership* 59 (5): 12–23.

Biddlecom, Ann E., and Bolaji M. Fapohunda. 1998. "Covert Contraceptive Use: Prevalence, Motivations, and Consequences." *Studies in Family Planning* 29 (4): 360.

Björklund, Anders, Mikael Lindahl, and Erik Plug. 2004. "Intergenerational Effects in Sweden : What Can We Learn from Adoption Data?" *IZA Discussion Paper Series* 1194.

———. 2006. "The Origins of Intergenerational Associations: Lessons from Swedish Adoption Data." *The Quarterly Journal of Economics* 121 (3): 999–1028.

Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes. 2005. "The More the Merrier? The Effect of Family Size and Birth Order on Children's Education." *The Quarterly Journal of Economics* 120 (2): 669–700.

———. 2010. "Small Family, Smart Family?: Family Size and the IQ Scores of Young Men." *The Journal of Human Resources* 45 (1): 33–58.

———. 2011. "Older and wiser? Birth order and IQ of young men." *CESifo Economic Studies* 57 (1): 103–120.

Blackburn, Susan. 2004a. "Violence." Chap. 8 in *Women and the State in Modern Indonesia,* 194–219. Cambridge: Cambridge University Press.

———. 2004 b. *Women and the State in Modern Indonesia.* Cambridge: Cambridge University Press.

Blake, Judith. 1981. "Family Size and the Quality of Children." *Demography* 18 (4): 421–442.

Bobonis, Gustavo J., Melissa González-Brenes, and Roberto Castro. 2013. "Public transfers and domestic violence: The roles of private information and spousal control." *American Economic Journal: Economic Policy* 5 (1): 179–205.

Bonesrønning, Hans, and Sofia Sandgren Massih. 2011. "Birth order effects on young students' academic achievement." *Journal of Socio-Economics* 40 (6): 824–832.

Bruins, Marianne. 2017. "Women's economic opportunities and the intra-household production of child human capital." *Labour Economics* 44 (C): 122–132.

Buckles, Kasey S., and Elizabeth L. Munnich. 2012. "Birth Spacing and Sibling Outcomes." *The Journal of Human Resources* 47 (3): 613–642.

Bueno, Cruz Caridad, and Errol A. Henderson. 2017. "Bargaining or Backlash? Evidence on Intimate Partner Violence from the Dominican Republic." *Feminist Economics* 23 (4): 90–116.

Buttell, Frederick P., and Michelle Mohr Carney. 2009. "Examining the Impact of Hurricane Katrina on Police Responses to Domestic Violence." *Traumatology* 15 (2): 6–9.

Calimeris, Lauren, and Christina Peters. 2017. "Food for thought: the birth order effect and resource allocation in Indonesia." *Applied Economics* 49 (54): 1–12.

Cameron, Lisa, and Christopher Worswick. 2003. "The Labor Market as a Smoothing Device: Labor Supply Responses to Crop Loss in Indonesia." *Review of Development Economics* 7 (2): 327–341.

Cameron, Stephen V., and James J. Heckman. 2001. "The Dynamics of Educational Attainment for Black, Hispanic, and White Males." *Journal of Political Economy* 109 (3): 455–499.

Card, David, and Gordon B. Dahl. 2011. "Family violence and football: The effect of unexpected emotional cues on violent behavior." *Quarterly Journal of Economics* 126 (1): 103–143.

Carrell, Scott E., Richard L. Fullerton, and James E. West. 2009. "Does Your Cohort Matter? Measuring Peer Effects in College Achievement." *Journal of Labor Economics* 27 (3): 439–464.

Caruso, Gérman Daniel. 2017. "The legacy of natural disasters: The intergenerational impact of 100 years of natural disasters in Latin America." *Journal of Development Economics* 127:209–233.

Cavallo, Eduardo A., Sebastian Galiani, Ilan Noy, and Juan Pantano. 2013. "Catastrophic Natural Disasters and Economic Growth." *Review of Economics and Statistics* 95 (5): 1549–1561.

Cavallo, Eduardo A., and Ilan Noy. 2010. "The Economics of Natural Disasters: A Survey." *IDB Working Paper Series,* no. 124.

Chikovore, Jeremiah, Gunilla Lindmark, Lennarth Nystrom, Michael T Mbizvo, and Beth Maina Ahlberg. 2002. "The hide-and-seek game: Men's perspectives on abortion and contraceptive use within marriage in a rural community in Zimbabwe." *Journal of Biosocial Science* 34 (3): 317–332.

Chin, Yoo Mi. 2011. "Male backlash, bargaining, or exposure reduction?: Women's working status and physical spousal violence in India." *Journal of Population Economics* 25 (1): 175–200.

Cools, Sara, and Andreas Kotsadam. 2017. "Resources and Intimate Partner Violence in Sub-Saharan Africa." *World Development* 95:211–230.

Cunha, F., and James J. Heckman. 2009. "The Economics and Psychology of Inequality and Human Development." *Journal of the European Economic Association* 7 (2-3): 320–364.

Curtis, Thom, Brent C. Miller, and E. Helen Berry. 2000. "Changes in reports and incidence of child abuse following natural disasters." *Child Abuse and Neglect* 24 (9): 1151–1162.

Darling-Hammond, Linda. 2000. "Teacher Quality and Student Achievement: A Review of State Policy Evidence Previous Research." *Education* 8 (1): 1–44.

Davis, Daniel J., Sorel Cahan Cahan, and Joseph Bashi. 1977. "Birth order and intellectual development: the confluence model in the light of cross-cultural evidence." *Science* 196 (4297): 1470–14722.

De Haan, Monique, Erik Plug, and Jose Rosero. 2014. "Birth Order and Human Capital Development: Evidence from Ecuador." *The Journal of Human Resources* 49 (2): 359–392.

Denlay, Patricia L., and Elizabeth Shrader. 2000. *Gender and post-disaster reconstruction: The case of Hurricane Mitch in Honduras and Nicaragua.* Washington, DC: LCSPG / LAC Gender Team - The World Bank.

Deryugina, Tatyana, Laura Kawano, and Steven D Levitt. 2018. "The Economic Impact of Hurricane Katrina on Its Victims: Evidence from Individual Tax Returns." *American Economic Journal: Applied Economics* 10 (2): 202–233.

Desforges, Charles, and Alberto Abouchaar. 2003. *The impact of parental involvement , parental support and family education on pupil achievements and adjustment: A literature review.* Vol. 433. Nottingham: DfES publications.

Donald, Aletheia, Gayatri Koolwal, Jeannie Annan, Kathryn Falb, and Markus Goldstein. 2017. *Measuring Women's Agency.* Washington DC: The World Bank.

Downey, Douglas B. 2001. "Number of Siblings and Intellectual Development: The Resource Dilution Explanation." *American Psychologist* 56 (6/7): 497–504.

Duflo, Esther. 2000. "Child Health and Household Resources in South Africa: Evidence from the Old Age Pension program." *The American Economic Review* 90 (2): 393–398.

———. 2012. "Women, Empowerment, and Economic Development." *Journal of Eonomic Literature* 50 (4): 1051–1079.

Duflo, Esther, and Christopher Udry. 2004. "Intrahousehold Resource Allocation in Cote D'Ivoire: Social Norms, Separate Accounts and Consumption Choices." *NBER Working Paper Series* 10498.

Dutton, Donald G., and Randall P. Kropp. 2000. "A Review of Domestic Violence Risk Instruments." *Trauma, Violence & Abuse* 1 (2): 171–181.

Ebot, Jane O. 2014. "Place Matters: Community Level Effects of Women's Autonomy on Ethiopian Children's Immunization Status." *African Population Studies* 28 (2): 1202–1215.

Edmonds, Eric V. 2006. "Understanding in child labor sibling differences." *Journal of Population Economics* 19 (2006): 795–821.

Ejrnæs, Mette, and Claus C. Portner. 2004. "Birth Order and the Intrahousehold Allocation of Time." *Review of Economics and Statistics* 86 (4): 1008–1019.

Emerson, Patrick M., and André Portela Souza. 2008. "Birth Order, Child Labor, and School Attendance in Brazil." *World Development* 36 (9): 1647–1664.

Eskreis-Winkler, Lauren, Ayelet Fishbach, and Angela Duckworth. n.d. "Dear Abby: Should I Give Advice or Receive it?" *Psychological Science* in press.

Fernandez, Cristina A., Benjamin Vicente, Brandon D. L. Marshall, Karestan C. Koenen, Kristopher L. Arheart, Robert Kohn, Sandra Saldivia, and Stephen L. Buka. 2017. "Longitudinal course of disaster-related PTSD among a prospective sample of adult Chilean natural disaster survivors." *International Journal of Epidemiology* 46 (2): 440–452.

Ferreira, Francisco H G, and Norbert Schady. 2009. "Aggregate economic shocks, child schooling, and child health." *World Bank Research Observer* 24 (2): 147–181.

Fisher, Sarah. 2010. "Violence Against Women and Natural Disasters: Findings From Post-Tsunami Sri Lanka." *Violence Against Women* 16 (8): 902–918.

Fors, Heather Congdon, and Annika Lindskog. 2017. "Within-Family Inequalities in Human Capital Accumulation in India: Birth Order and Gender Effects." *University of Gothenburg - Working Paper in Economics* 700.

Frankenberg, Elizabeth, and Duncan Thomas. 2001. "Measuring Power." *IFPRI Food Consumption and Nutrition Division Discussion Paper* 113.

Galbraith, Richard C. 1982. "Sibling spacing and intellectual development: A closer look at the confluence models." *Developmental Psychology* 18 (2): 151–173.

García-Moreno, Claudia, Christina Pallitto, Karen Devries, Heidi Stöckl, Charlotte Watts, and Naeemah Abrahams. 2013. *Global and regional estimates of violence against women: prevalence and health effects of intimate partner violence and non-partner sexual violence.* Geneva, Switzerland: World Health Organization (WHO).

Gasca, Natalie C., and Stan Becker. 2017. "Using Couples' Discordant Reports To Estimate Female Covert Use of Modern Contraception in Sub-Saharan Africa." *Journal of Biosocial Science* 50 (3): 326–346.

Ghuman, Sharon J., Helen J. Lee, and Herbert L. Smith. 2006. "Measurement of women's autonomy according to women and their husbands: Results from five Asian countries." *Social Science Research* 35:1–28.

Gignoux, Jérémie, and Marta 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.

Gitter, Seth R., and Bradford L. Barham. 2007. "Credit, Natural Disasters, Coffee, and Educational Attainment in Rural Honduras." *World Development* 35 (3): 498–511.

Goenjian, Armen K., Alan M. Steinberg, Louis M. Najarian, Lynn A. Fairbanks, Madeline Tashjian, and Robert S. Pynoos. 2000. "Prospective Study of Posttraumatic Stress, Anxiety, and Depressive Reactions After Earthquake and Political Violence." *The American Journal of Psychiatry* 157 (6): 911–916.

Goldin, Claudia. 2006. "The Quiet Revolution That Transformed Women's Employment, Education, and Family." *American Economic Review* 96 (2): 1–21.

Granbois, Donald H., and Ronald P. Willett. 1970. "Equivalence of Family Role Measures Based on Husband and Wife Data." *Journal of Marriage and Family* 32 (1): 68–72.

Gray, Jeffrey S. 1998. "Divorce-Law Changes, Household Bargaining, and Married Women's Labor Supply." *American Economic Review* 88 (3): 628–642.

Gualtieri, Thomas, and Robert E. Hicks. 1985. "An immunoreactive theory of selective male affliction." *Behavioral and Brain Sciences* 8 (3): 427–441.

Haan, M *de.* 2010. "Birth order, family size and educational attainment." *Economics of Education Review* 29 (4): 576–588.

Hammitt, James, Jin-Tan Liu, and Meng-Wen Tsou. 2012. "The intergenerational transmission of education: Evidence from Taiwanese adoptions." *Economics Letters* 115 (1): 134–136.

Hanushek, Eric A., and Ludger Woessmann. 2015. *The Knowledge Capital of Nations: Education and the Economics of Growth.* Cambridge, MA: CES ifo Book Series, The MIT Press.

Härkönen, Juho. 2014. "Birth order effects on educational attainment and educational transitions in West Germany." *European Sociological Review* 30 (2): 166–179.

Harris, Alma. 2008. "Do parents know they matter? Engaging all parents in learning." *Educational Research* 5 (3): 37–41.

Hartog, Joop, and Hessel Oosterbeek. 1998. "Health, wealth and happiness: why pursue a higher education?" *Economics of Education Review* 17 (3): 245–256.

Hayati, Elli Nur. 2013. "Domestic violence against women in rural Indonesia: Searching for multilevel prevention." Medical Dissertations, Umea University.

Hayati, Elli Nur, Malin Eriksson, Mohammad Hakimi, Ulf Högberg, and Maria Emmelin. 2013. "Elastic bands strategy: women's lived experience of coping with domestic violence in rural Indonesia." *Global Health Action* 6 (1): 18894.

Hayati, Elli Nur, Ulf Högberg, Mohammad Hakimi, Mary C. Ellsberg, and Maria Emmelin. 2011. "Behind the silence of harmony: risk factors for physical and sexual violence among women in rural Indonesia." *BMC Women's Health* 11 (1): 52.

Hayes, Brittany E., and Katharine A. Boyd. 2017. "Influence of Individual- and National-Level Factors on Attitudes toward Intimate Partner Violence." *Sociological Perspectives* 60 (4): 685–701.

Heath, Rachel. 2014. "Women's Access to Labor Market Opportunities, Control of Household Resources, and Domestic Violence: Evidence from Bangladesh." *World Development* 57:32–46.

Heckman, James J., Lance J. Lochner, and Petra E. Todd. 2006. "Earnings Functions, Rates of Return and Treatment Effects: The Mincer Equation and Beyond." Chap. 7 in *Handbook of the Economics of Education,* edited by Eric A. Hanushek and Finis Welch, 1:307–458. St. Louis, MO: Elsevier.

Hidrobo, Melissa, Amber Peterman, and Lori Heise. 2016. "The effect of cash, vouchers, and food transfers on intimate partner violence: Evidence from a randomized experiment in Northern Ecuador." *American Economic Journal: Applied Economics* 8 (3): 284–303.

Hjern, Anders, Frank Lindblad, and Bo Vinnerljung. 2002. "Suicide, psychiatric illness, and social maladjustment in intercountry adoptees in Sweden: a cohort study." *The Lancet* 360 (9331): 443–448.

Hoddinott, John, and Bill Kinsey. 2001. "Child growth in the time of drought." *Oxford Bulletin of Economics and Statistics* 63 (4): 409–436.

Hotz, V Joseph, and Juan Pantano. 2015. "Strategic Parenting, Birth Order and School Performance." *Journal of Population Economics* 28 (4): 911–936.

International Federation of Red Cross and Red Crescent Societies. 2014. *Emergency Plan of Action (EPoA) Indonesia: Volcanic Eruption – Mt Kelud.* Online Report. Accessed September 4, 2018. `https://reliefweb.int/sites/reliefweb.int/files/resources/MDRID009dref.pdf`.

Janvry, Alain *de*, Frederico Finan, Elisabeth Sadoulet, and Renos Vakis. 2006. "Can conditional cash transfer programs serve as safety nets in keeping children at school and from working when exposed to shocks?" *Journal of Development Economics* 79:349–373.

Jati, Raditya. 2015. *National progress report on the implementation of the Hyogo Framework for Action (2013-2015).* Progress Report 2013-2014. The Indonesian National Agency for Disaster Management (BNPB). Accessed September 4, 2018. `https://www.preventionweb.net/files/41507%7B%5C_%7DIDN%7B%5C_%7DNationalHFAprogress%7B%5C_%7D2013-15.pdf`.

Jayachandran, Seema, and Ilyana Kuziemko. 2011. "Why Do Mothers Breastfeed Girls Less than Boys? Evidence and Implications for Child Health in India." *The Quarterly Journal of Economics* 126 (3): 1485–1538.

Jejeebhoy, Shireen J. 2002. "Convergence and Divergence in Spouses' Perspectives on Women's Autonomy in Rural India." *Studies in Family Planning* 33 (4): 299–308.

Jensen, Robert. 2000. "Agricultural volatility and investments in children." *American Economic Review* 90 (2): 399–404.

Jensen, Robert. 2012. "Do labor market opportunities affect young women's work and family decisions? Experimental evidence from India." *The Quarterly Journal of Economics* 127 (2): 753–792.

Jensen, Robert, and Emily Oster. 2009. "The Power of TV: Cable Television and Women's Status in India." *The Quarterly Journal of Economics* 124 (3): 1057–1094.

Jewkes, Rachel. 2002. "Intimate partner violence: causes and prevention." *The Lancet* 359 (9315): 1423–1429.

Kabeer, Naila. 1999. "Resources, Agency, Achievements: Reflections on the Measurement of Women's Empowerment." *Development and Change* 30 (3): 435–464.

Kabeer, Naila, Simeen Mahmud, and Sakiba Tasneem. 2011. "Does Paid Work Provide a Pathway to Women's Empowerment? Empirical Findings from Bangladesh." *IDS Working Paper* 375.

Kanazawa, Satoshi. 2012. "Intelligence, Birth Order, and Family Size." *Personality and Social Psychology Bulletin* 38 (9): 1157–1164.

Kantarevic, Jasmin, and Stéphane Mechoulan. 2006. "Birth Order , Educational Attainment , and Earnings: An Investigation Using the PSID." *The Journal of Human Resources* 41 (4): 755–777.

Kochar, Anjini. 1999. "Smoothing consumption by smoothing income: Hours-of-work responses to idiosyncratic agricultural shocks in rural India." *The Review of Economics and Statistics* 81 (1): 50–61.

Kristensen, Petter, and Tor Bjerkedal. 2007. "Explaining the relation between birth order and intelligence." *Science* 316 (5832): 1717.

Lafortune, Jeanne, and Soohyung Lee. 2014. "All for One? Family Size and Children's Educational Distribution Under Credit Constraints." *American Economic Review: Papers and Proceedings* 104 (5): 365–369.

Lindblad, Frank, Anders Hjern, and Bo Vinnerljung. 2003. "Intercountry Adopted Children as Young Adults – A Swedish Cohort Study." *American Journal of Orthopopsychiatry* 73 (2): 190–202.

Loayza, Norman V., Eduardo Olaberría, Jamele Rigolini, and Luc Christiaensen. 2012. "Natural Disasters and Growth: Going Beyond the Averages." *World Development* 40 (7): 1317–1336.

Lundberg, Shelly J., Robert A. Pollak, and Terence J. Wales. 1997. "Do Husbands and Wives Pool Their Resources? Evidence from the United Kingdom Child Benefit." *The Journal of Human Resources* 32 (3): 463–480.

Lundberg, Shelly, and Robert A. Pollak. 1994. "Non-cooperative Bargaining Models of Marriage." *American Economic Review (Papers & Proceedings)* 84 (2): 132–137.

———. 1996. "Bargaining and Distribution in Marriage." *The Journal of Economic Perspectives* 10 (4): 139–158.

Lupri, Merlin, and Eugen Brinkerhoff. 1978. "Theoretical and Methodological Issues in the Use of Decision-Making as an Indicator of Conjugal Power: Some Canadian Observations." *The Canadian Journal of Sociology / Cahiers canadiens de sociologie* 3 (1): 1–20.

Maccini, Sharon, and Dean Yang. 2009. "Under the Weather: Health, Schooling, and Economic Consequences of Early-Life Rainfall." *American Economic Review* 99 (3): 1006–1026.

Majlesi, Kaveh. 2016. "Labor market opportunities and women's decision making power within households." *Journal of Development Economics* 119:34–47.

Malhotra, Anju, and Sidney Ruth Schuler. 2005. "Women's Empowerment as a Variable in International Development." Chap. 3 in *Measuring Empowerment: Cross-Disciplinary Perspectives,* edited by Deepa Narayan, 71–88. Washington, DC: The World Bank.

Malhotra, Anju, Sidney Ruth Schuler, and Carol Boender. 2002. "Measuring Women's Empowerment as a Variable in International Development." *Background Paper Prepared for the World Bank Workshop on Poverty and Gender: New Perspectives* (World Bank).

Mankiw, Gregory, David Romer, and David N Weil. 1992. "A Contribution to the Empirics of Economic Growth." *The Quarterly Journal of Economics* 107 (2): 407–437.

Mbaye, Linguère Mously, and Natascha Wagner. 2017. "Bride Price and Fertility Decisions: Evidence from Rural Senegal." *Journal of Development Studies* 53 (6): 891–910.

Mboane, Ramos, and Madhav P. Bhatta. 2015. "Influence of a husband's healthcare decision making role on a woman's intention to use contraceptives among Mozambican women." *Reproductive Health* 12 (1): 1–8.

McElroy, Marjorie B, and Mary Jean Horney. 1981. "Nash-Bargained Household Decisions: Toward a Generalization of the Theory of Demand." *International Economic Review* 22 (2): 333–349.

McFarlane, Alexander C. 1998. "Epidemiological evidence about the relationship between ptsd and alcohol abuse: The nature of the association." *Addictive Behaviors* 23 (6): 813–825.

Mechoulan, Stéphane, and François Charles Wolff. 2015. "Intra-household allocation of family resources and birth order: evidence from France using siblings data." *Journal of Population Economics* 28 (4): 937–964.

Miguel, Edward. 2005. "Poverty and witch killing." *Review of Economic Studies* 72 (4): 1153–1172.

Moffitt, Terrie E., Avshalom Caspi, Robert F. Krueger, Lynn Magdol, Gayla Margolin, Phil A. Silva, and Ros Sydney. 1997. "Do partners agree about abuse in their relationship? A psychometric evaluation of interpartner agreement." *Psychological Assessment* 9 (1): 47–56.

Moshoeshoe, Ramaele. 2016. "Birth Order Effects on Educational Attainment and Child Labour: Evidence from Lesotho." *ERSA Working Paper* 621.

Neria, Yuval, Arijit Nandi, and Sandro Galea. 2008. "Post-traumatic stress disorder following disasters: a systematic review." *Psychological Medicine* 38 (4): 467–480.

Nilan, Pam, Argyo Demartoto, Alex Broom, and John Germov. 2014. "Indonesian Men's Perceptions of Violence Against Women." *Violence Against Women* 20 (7): 869–888.

Noy, Ilan. 2009. "The macroeconomic consequences of disasters." *Journal of Development Economics* 88 (2): 221–231.

OECD. 2015. *OECD Economic Surveys: Indonesia 2015.* Paris: OECD Publishing.

Overstreet, Stacy, Alison Salloum, Berre Burch, and Jill West. 2011. "Challenges Associated with Childhood Exposure to Severe Natural Disasters: Research Review and Clinical Implications." *Journal of Child & Adolescent Trauma* 4 (1): 52–68.

Parkinson, Debra. 2017. "Investigating the Increase in Domestic Violence Post Disaster: An Australian Case Study." *Journal of Interpersonal Violence.*

Parkinson, Debra, and Claire Zara. 2013. "The hidden disaster: Domestic violence in the aftermath of natural disaster." *Australian Journal of Emergency Management* 28 (2): 28–35.

Paxson, Christina, Elizabeth Fussell, Jean Rhodes, and Mary Waters. 2012. "Five years later: Recovery from Post Traumatic Stress and Psychological Distress Among Low-Income Mothers Affected by Hurricane Katrina." *Social Science & Medicine* 74 (2): 150–7.

Picardo, Carla W., Shirley Burton, and John Naponick. 2010. "Physically and Sexually Violent Experiences of Reproductive-Aged Women Displaced by Hurricane Katrina." *The Journal of the Louisiana State Medical Society* 162 (5): 282–290.

Pierotti, Rachael S. 2013. "Increasing Rejection of Intimate Partner Violence: Evidence of Global Cultural Diffusion." *American Sociological Review* 78 (2): 240–265.

Price, Joseph. 2008. "Parent-Child Quality Time: Does Birth Order Matter?" *The Journal of Human Resources* 43 (1): 240–265.

Quarm, Daisy. 2018. "Random Measurement Error as a Source of Discrepancies between the Reports of Wives and Husbands concerning Marital Power and Task Allocation." *Journal of Marriage and Family* 43 (3): 521–535.

Rahman, Lupin, and Vijayendra Rao. 2004. "The determinants of gender equity in India: Examining Dyson and Moore's thesis with new data." *Population and Development Review* 30 (2): 239–268.

Rasul, Imran. 2008a. "Household bargaining over fertility : Theory and evidence from Malaysia." *Journal of Development Economics* 86:215–241.

———. 2008 b. "Household bargaining over fertility: Theory and evidence from Malaysia." *Journal of Development Economics* 86 (2): 215–241.

Reading, Richard. 2008. "The impact of exposure to domestic violence on children and young people: A review of the literature." *Child Abuse & Neglect* 32:797–810.

Retherford, Robert D., and William H. Sewell. 1991. "Birth Order and Intelligence: Further Tests of the Confluence Model." *American Sociological Review* 56 (2): 141–158.

Rezaeian, Mohsen. 2013. "The association between natural disasters and violence: A systematic review of the literature and a call for more epidemiological studies." *Journal of Research in Medical Sciences* 18 (12): 1103–1107.

Rhodes, Jean, Christian Chan, Christina Paxson, Cecilia Elena Rouse, Mary Waters, and Elizabeth Fussell. 2010. "The Impact of Hurricane Katrina on the Mental and Physical Health of Low-Income Parents in New Orleans." *American Journal of Orthopsychiatry* 80 (2): 237–247.

Rodgers, Joseph Lee. 1984. "Confluence Effects: Not Here, Not Now!" *Developmental Psychology* 20 (2): 321–31.

Rodgers, Joseph Lee, Hobart Cleveland, Edwin *van den* Oord, and David C. Rowe. 2000. "Resolving the debate over birth order, family size, and intelligence." *The American Psychologist* 55 (6): 599–612.

Rosborough, Stephanie, Jennifer L. Chan, and Parveen Parmar. 2009. "Responding to Gender-based Violence in Disasters: Grappling With Research Methods to Clear the Way for Planning." *Disaster Medicine and Public Health Preparedness* 3 (1): 8–10.

Rosenzweig, Mark R. 1995. "Why are there returns to schooling?" *American Economic Review* 85 (2): 153–158.

Rowe, David C., Wendy J. Vesterdal, and Joseph L. Rodgers. 1998. "Herrnstein's syllogism: genetic and shared environmental influences on IQ, education, and income." *Intelligence* 26 (4): 405–423.

Rubonis, Anthony V., and Leonard Bickman. 1991. "Psychological Impairment in the Wake of Disaster: The Disaster-Psychopathology Relationship." *Psychological Bulletin* 109 (3): 384–399.

Runyan, Desmond, Corrine Wattam, Robin Ikeda, Hassan Fatma, and Laurie Ramiro. 2002. "Child abuse and neglect by parents and other caregivers." Chap. 3 in *World report on violence and health,* edited by Etienne G. Krug, Linda L. Dahlberg, Mercy James A., Anthony B. Zwi, and Lozano Rafael, 57–86. Geneva: World Health Organization.

Sacerdote, Bruce. 2011. "Peer Effects in Education: How Might They Work, How Big Are They and How Much Do We Know Thus Far?" Chap. 4 in *Handbook of the Economics of Education,* edited by Eric A. Hanushek, Stephen Machin, and Ludger Woessmann, 3:249–277. St. Louis, MO: Elsevier.

———. 2014. "Experimental and Quasi-Experimental Analysis of Peer Effects: Two Steps Forward?" *Annual Review of Economics* 6 (1): 253–272.

Safilios-Rothschild, Constantina. 1970. "The Study of Family Power Structure: A Review 1960-1969." *Journal of Marriage and Family* 32 (4): 539–552.

Schaner, Simone, and Smita Das. 2016. "Female Labor Force Participation in Asia: Indonesia Country Study." *ADB Economics Working Paper Series* 474.

Schooler, Carmi. 1972. "Birth order effects: Not here, not now!" *Psychological Bulletin* 78 (3): 161–175.

Schumacher, Julie A., Scott F. Coffey, Fran H. Norris, Melissa Tracy, Kahni Clements, and Sandro Galea. 2010. "Intimate Partner Violence and Hurricane Katrina: Predictors and Associated Mental Health Outcomes." *Violence and Victims* 25 (5): 588–603.

Sekhri, Sheetal, and Adam Storeygard. 2014. "Dowry deaths: Response to weather variability in India." *Journal of Development Economics* 111:212–223.

Sety, M., K. James, and J. Breckenridge. 2014. "Understanding the Risk of Domestic Violence During and Post Natural Disasters: Literature Review." Chap. 5 in *Issues of Gender and Sexual Orientation in Humanitarian Emergencies,* edited by Larry W. Roeder, 99–111. Cham: Springer International Publishing Switzerland.

Shah, Manisha, and Bryce Millett Steinberg. 2017. "Drought of Opportunities: Contemporaneous and Long Term Impacts of Rainfall Shocks on Human Capital." *Journal of Political Economy* 125 (2): 527–561.

Sonne, Susan C., Sudie E. Back, Claudia Diaz Zuniga, Carrie L. Randall, and Kathleen T. Brady. 2009. "Gender Differences in Individuals with Comorbid Alcohol Dependence and Post-traumatic Stress Disorder." *American Journal on Addictions* 12 (5): 412–423.

SOS Children's Villages International. 2016. *International Annual Report 2015.* Vienna, Austria. Accessed September 4, 2018. `https://issuu.com/soschildrensvillagesinternational/docs/sos%7B%5C_%7Dchildrens%7B%5C_%7Dvillages%7B%5C_%7Dinternationa`.

Statistics Indonesia (Badan Pusat Statistik - BPS) National Population and Family Planning Board (BKKBN) Kementerian Kesehatan (Kemenkes—MOH) and ICF International. 2012. *Indonesia Demographic and Health Survey 2012.* Jakarta, Indonesia: BPS, BKKBN, Kemenkes, ICF International.

Story, William T., and Sarah A. Burgard. 2012. "Couples' reports of household decision-making and the utilization of maternal health services in Bangladesh." *Social Science & Medicine* 75 (12): 2403–2411.

Strauss, John, Kathleen Beegle, Bondan Sikoki, Agus Dwiyanto, Yulia Herawati, and Firman Witoelar. 2004. *The Third Wave of the Indonesia Family Life Survey (IFLS): Overview and Field Report.* WR-144/1-NIA/NICHD.

Strauss, John, Firman Witoelar, and Bondan Sikoki. 2016. *The Fifth Wave of the Indonesia Family Life Survey (IFLS5): Overview and Field Report.* WR-1143/1-NIA/NICHD.

Strauss, John, Firman Witoelar, Bondan Sikoki, and Anna Marie Wattie. 2009. *The Fourth Wave of the Indonesian Family Life Survey (IFLS4): Overview and Field Report.* WR-675/1-NIA-NICHID.

Strobl, Eric. 2012. "The economic growth impact of natural disasters in developing countries: Evidence from hurricane strikes in the Central American and Caribbean regions." *Journal of Development Economics* 97 (1): 130–141.

Sulloway, Frank J. 2007. "Birth Order and Intelligence." *Science* 316 (5832): 1711–1712.

Tenikue, Michel, and Bertrand Verheyden. 2010. "Birth order and schooling: Theory and evidence from twelve Sub-Saharan countries." *Journal of African Economies* 19 (4): 459–495.

The United Nations Office for the Coordination of Humanitarian Affairs (OCHA). 2014. *Indonesia: Mt. Kelud Eruption.* Online report. Accessed September 4, 2018. `https://reliefweb.int/map/indonesia/indonesia-mt-kelud-eruption-14-february-2014`.

The World Bank Group. 2018. *The World Bank In Indonesia.* The World Bank. Accessed September 4, 2018. `https://www.worldbank.org/en/country/indonesia/overview`.

Thomas, Duncan, Elizabeth Frankenberg, and James P. Smith. 2001. "Lost but Not Forgotten: Attrition and Follow-up in the Indonesia Family Life Survey." *The Journal of Human Resources* 36 (3): 556–592.

Twyman, Jennifer, Pilar Useche, and Carmen Diana Deere. 2015. "Gendered Perceptions of Land Ownership and Agricultural Decision-making in Ecuador: Who Are the Farm Managers?" *Land Economics* 91 (3): 479–500.

Uddin, Jalal, Mohammad Habibullah, and Md Nasim-Us Sabah. 2016. "Correlates of unmet need for contraception in Bangladesh: does couples ' concordance in household decision making matter?" *Contraception* 94 (1): 18–26.

UN Women (United Nations Entity for Gender Equality and the Empowerment of Women). 2011. *Violence against Women Prevalence Data: Surveys by Country.* Online report. Accessed September 4, 2018. http://www.endvawnow.org/uploads/browser/files/vaw%7B%5C_%7Dprevalence%7B%5C_%7Dmatrix%7B%5C_%7D15april%7B%5C_%7D2011.pdf.

Usano, Robert J., Carol S. Fullerton, Lars Weisaeth, and Beverley Raphael. 2017. *Textbook of disaster psychiatry.* Cambridge: Cambridge University Press.

Van Aalst, Maarten K. 2006. "The impacts of climate change on the risk of natural disasters." *Disasters* 30 (1): 5–18.

Wahlstrom, Margareta, and Debarati Guha-Sapir. 2015. *The human cost of weather-related disasters 1995-2015.* Geneva, Switzerland: UNISDR.

Wang, Shing-yi. 2014. "Property Rights and Intra-Household Bargaining." *Journal of Development Economics* 107:192–201.

Warsini, Sri, Petra Buettner, Jane Mills, Caryn West, and Kim Usher. 2014. "The Psychosocial Impact of the Environmental Damage Caused by the MT Merapi Eruption on Survivors in Indonesia." *EcoHealth* 11 (4): 491–501.

Watson, Robert T., and Daniel L. Albritton, eds. 2001. *Climate change 2001: Synthesis report: Third assessment report of the Intergovernmental Panel on Climate Change.* Cambridge: Cambridge University Press.

Westoff, Charles F. 2010. "Desired Number of Children: 2000–2008." *DHS Comparative Reports* 25.

Wichman, Aaron L., Joseph Lee Rodgers, and Robert C. Maccallum. 2006. "A Multilevel Approach to the Relationship Between Birth Order and Intelligence." *Personality and Social Psychology Bulletin* 32 (1): 117–127.

World Health Organization (WHO). 2002. *Gender and Health in Disasters.* Information Sheet. Accessed September 4, 2018. http://www.who.int/gender/gwhgendernd2.pdf.

———. 2008. *Country policies and strategies for combating GBV.* Factsheet. Accessed September 4, 2018. www.searo.who.int/entity/gender/data/indonesia.pdf?ua=1.

Wunderman, Richard, ed. 2014a. "Global Volcanism Program, 2014. Report on Kelut (Indonesia)." In *Bulletin of the Global Volcanism Network,* 39:02. Washington, DC: Smithsonian Institution.

Wunderman, Richard, ed. 2014b. "Global Volcanism Program, 2014. Report on Merapi (Indonesia)." In *Bulletin of the Global Volcanism Network,* 39:10. Washington, DC: Smithsonian Institution.

York, Travis T., Charles Gibson, and Susan Rankin. 2015. "Defining and Measuring Academic Success." *Practical Assessment, Research & Evaluation* 20 (5): 1–20.

Zajonc, Robert B. 1976. "Family configuration and intelligence." *Science* 192 (4236): 227–236.

———. 2001. "The family Dynamics of Intellectual Development." *American Psychologist* 56 (6/7): 490–496.

Zajonc, Robert B., and Gregory B. Markus. 1975. "Birth order and intellectual development." *Psychological Review* 82 (1): 74–88.

Zajonc, Robert B., Hazel Markus, and Gregory B. Markus. 1979. "The Birth Order Puzzle." *Journal of Personality and Social Psychology* 37 (8): 1325–1341.

Zajonc, Robert B., and Patricia R. Mullally. 1997. "Birth order: Reconciling conflicting effects." *American Psychologist* 52 (7): 685–699.

Zajonc, Robert B., and Frank J. Sulloway. 2007. "The confluence model: birth order as a within-family or between-family dynamic?" *Personality & Social Psychology Bulletin* 33 (9): 1187–94.