Essays in Applied Microeconomics



Lukas Hensel

St Peter's College
University of Oxford

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

September 2019

Essays in Applied Microeconomics

Lukas Hensel

St Peter's College
University of Oxford

Abstract This thesis consists of three self-contained chapters on topics in applied microeconomics. In Chapter 1 "The Impact of Economic Shocks on Wage Dynamics in Indonesia" I study the impact of negative productivity shocks on agricultural wages in Indonesia. I study whether and when downward nominal wage rigidity prevents labor market adjustments to negative productivity shocks. I provide evidence that the cost to the economy and to agricultural households caused by downward wage rigidities are limited in magnitude. I use transitory labor demand shocks in agricultural labor markets induced by plausibly exogenous rainfall fluctuations to show that downward adjustment of nominal wages is a function of the cost of maintaining the nominal wage at the prevailing level. I also develop a theoretical framework with reference-dependent utility over past wages, in line with fairness norms, that is consistent with the observed wage rigidities.

In Chapter 2 "Income Shocks and Suicides: Causal Evidence From Indonesia" (coauthored with Cornelius Christian and Chris Roth) we explore the role of income shocks on the prevalence of suicides in Indonesia. We find that a cash transfer program reduces the yearly suicide rate by 0.36 per 100,000 people, corresponding to an 18 percent decrease. To come to this conclusion, we use a difference-in-differences approach, exploiting a cash transfer's nationwide roll-out, and corroborate the findings using a randomized experiment. Agricultural productivity shocks similar to those in Chapter 1 also affect the suicide rate. Encouragingly, the cash transfer program reduces the causal impact of negative agricultural productivity shocks, suggesting an important role for policy interventions. We only observe these effects for households engaged in agricultural production. Finally, we provide evidence for depression as a psychological mechanism through which income shocks affect the suicide rate.

In Chapter 3 "Political Activists as Free-Riders: Evidence from a Natural Field Experiment" (coauthored with Anselm Hager, Johannes Hermle, and Chris Roth) we explore how a citizen's decision to participate in political activism depends on the participation of others. We examine this core question of collective action in a natural field experiment in collaboration with a major European party during a recent national election. In a seemingly unrelated party survey, we randomly assign canvassers to true information about the canvassing intentions of their peers. We find that treated canvassers significantly reduce both their canvassing intentions and behavior when learning that their peers participate more in canvassing than previously believed.

Supervisors This thesis was supervised by Climent Quintana-Domeque (University of Exeter) and Simon Quinn (University of Oxford).

Total words This thesis contains 81,534 words. This word count has been calculated by counting the number of words on a representative page (381 words, page 2) and multiplying this with the number of pages, excluding the cover, table of contents, and references.

Statement of authorship Chapters 1 of my thesis "The Impact of Economic Shocks on Wage Dynamics in Indonesia" is single-authored and it is fully my own work.

Chapter 2 "Income Shocks and Suicides: Causal Evidence From Indonesia" is joint work with Christopher Roth and Cornelius Christian, both former DPhil students at the University of Oxford. Christopher Roth is currently a postdoctoral research fellow at the Institute on Behavior and Inequality (briq) and Cornelius Christian is an assistant professor in Economics at Brock University. All three of us developed the overall research design. I was equally involved in the development of the overall research design. I contributed most to the design of the analysis involving the mechanism analysis in section 4 and its interpretation in section 2. I also conducted much of the data handling and analysis (including but not limited to merging survey data across years, generating rainfall data, and generating the final result tables and figures). I drafted most of sections 2, 4, and 5. Additionally, I made major contributions to writing all other sections of the chapter. I had a particularly important role in successfully responding to referee comments received from the Review of Economics and Statistics in 2018.

Chapter 3 "Political Activists as Free-Riders: Evidence from a Natural Field Experiment" is joint work with Anselm Hager, Johannes Hermle, and Christopher Roth. Anselm Hager is an assistant professor in political economy at the graduate school of decision sciences in Konstanz. He received his PhD in Political Science in 2017 at Columbia University. Johannes Hermle is a PhD student at the Department of Economics at the University of California - Berkeley. All coauthors jointly designed the field experiment. Lukas implemented all fieldwork and data collection activities for this project. He also conducted all data analysis himself. Lukas conceived the conceptual framework in section 2 and he came up with the idea of using LASSO techniques to test for the heterogeneity tests in section 5. All other sections of the chapter were written in equal shares by all coauthors.

Acknowledgment of funding I acknowledge funding from the Economic and Social Research Council (ESRC), the University of Oxford Social Sciences Doctoral Training Centre (DTC), the German Academic Exchange Service (DAAD), and the German Academic Scholarship Foundation. Furthermore, I acknowledge financial support for conference presentations from St Peter's College, the Royal Economic Society, the European Economic Association, the Institute for the Future of Labor (IZA), the Centre for Experimental Research on Fairness, Inequality and Rationality (FAIR), and the Institute on Behavior and Inequality (briq). I am very grateful for the support I have received.

Ethical approval For Chapter three ethical approval has been sought and obtained from the University of Oxford CUREC committee.

Declaration The work in this thesis is based on research carried out by me at the University of Oxford between June 2015 and February 2019. No part of this thesis has been submitted for a doctoral degree at this, or any other, university.

Lukas Hensel (*September 2019)*

Acknowledgements

I am grateful to my Ph.D. advisors Climent Quintana-Domeque and Simon Quinn for their support and guidance in writing this thesis. Climent, thank you for providing invaluable guidance on all aspects of becoming an academic economist. You were always available to answer my questions and your feedback greatly improved my writing. I consider myself very fortunate to have had you as an advisor. Simon, thank you, for being so welcoming from the very beginning. Your feedback throughout my years at Oxford was crucial to develop my presentation skills. You also shaped much of my perspective on empirical research through your relentless questioning of the use and interpretation of statistical methods.

I would like to thank Kate Orkin for her trust and patience while working as a research assistant for her. Kate, your passion for development and academic rigor was an inspiration throughout my time at Oxford. You also taught me the ins and outs of fieldwork which was crucial for me to become an independent researcher. Thank you for being a great mentor, Kate. Moreover, I would like to thank Johannes Abeler for encouraging me to think harder about the framing and contribution of my research.

I am grateful to all my coauthors who have taught me much of what I know about conducting research: Christopher Roth, Anselm Hager, Johannes Hermle, Cornelius Christian, Marc Witte, Tegsay Tekleselassie, Girum Abebe, Stefano Caria, Stefan Dercon, Kate Orkin, and Robert Garlick. I would like to thank Christopher Roth for being an amazing coauthor, friend, and mentor. Thank you for introducing me to the world of academic publishing and for pushing me to connect to the world of academia outside of Oxford. Your energy, creativity, and motivation are a constant source of inspiration. I would like to thank Anselm Hager for his patience in explaining political science to his economist coauthors. Thank you also to Johannes Hermle for your creativity and writing skills. I would also like to thank Marc Witte for your great ideas and tireless work on our joint projects. Stefano Caria, thank you for questioning our research ideas until we find the very best. Finally, thank you to Robert Garlick for teaching me to use good coding practices and for constantly introducing me to new statistical methods.

Oxford's great academic community provided the ideal environment to write this thesis. I am deeply indebted to everyone who provided academic and personal support throughout my five years in Oxford. I would like to thank Michael Kölle, Giulio Schiania, Markus Parlasca, Ben Kett, Katherine Stapleton, Ashley Pople, Vanessa Schreiber, Madhav Vaidyanathan, Emma Riley, Mahreen Mahmud, Anouk Rigterink, Gabriel Ulyssea, Julien Labonne, and Chris Woodruff

for many inspiring conversations. I would like to thank Rose Page for providing crucial administrative support for many of my research endeavors. Thank you also Filip Bubenheimer and David Hagebölling broadening my horizon and being fun to hang out with.

Myrto Oikonomou, thank you for always being willing to listen and for all the fun we had together in Oxford. I thank Jemima Peppel for being the heart and soul of our first-year study group and a good friend. A very heartfelt thank you to Timo Reinelt and Alexander Haas who have accompanied me through much of my academic and personal life. Thank you both for your constant support, advise, and company throughout the process of writing this thesis. Without you, I would not be where I am today.

Part of this thesis was written during a research stay at the University of California - Berkeley. I would like to thank Supreet Kaur for hosting me and for providing feedback on my research. Moreover, this thesis has benefited enormously from the opportunity to attend and present my work at seminars, conferences, and summer schools in Europe, the US, and Africa. I would like to thank all organizers and participants of these events.

Last but not least, I thank my family and friends who accompanied me through this time. I thank my parents and brother for their love and unconditional support. Thank you also to Sangjung Ha. Sangjung, you have been a crucial source of love, support, and inspirations throughout the five years of writing this thesis. Thank you everything we experienced together and for encouraging me to go out of my comfort zone and explore life.

Lukas Hensel
Oxford, September 2019

Contents

Introduction **1**

1 The Impact of Economic Shocks on Wage Dynamics in Indonesia **5**
1.1 Introduction . . . 6
1.2 Modeling Rigid Wages with Reference Dependence . . . 8
1.3 Context and Data . . . 14
1.4 Empirical Specifications . . . 23
1.5 Main Results . . . 27
1.6 Conclusion . . . 40

Appendices **42**
1.A Mathematical Appendix . . . 42
1.B Additional Tables . . . 56
1.C Additional Figures . . . 65

2 Income Shocks and Suicides: Causal Evidence From Indonesia **69**
2.1 Introduction . . . 70
2.2 Conceptual Framework, Context, and Data . . . 72
2.3 Main Results . . . 82
2.4 Mechanism . . . 95
2.5 Interpreting Effect Sizes . . . 99
2.6 Conclusion . . . 101

Appendices **104**
2.A Effect Size Calculations . . . 106
2.B Additional Tables . . . 110
2.C Additional Figures . . . 137

3 Political Activists as Free-Riders: Evidence from a Natural Field Experiment **144**
3.1 Introduction . . . 145
3.2 Conceptual framework . . . 150
3.3 Experimental Design . . . 153
3.4 Belief Updating . . . 157
3.5 Results . . . 158
3.6 Conclusion . . . 172

Appendices **175**
3.A Mathematical Appendix 176
3.B Invitation email 179
3.C Survey instrument 180
3.D Deviations from pre-analysis plan 182
3.E Additional Figures 183
3.F Additional Tables 185

Conclusion **214**

References **215**

Introduction

This thesis is the result of my research in two broad areas: development economics and behavioral political economy. My research agenda is unified by the rigorous use of experimental and quasi-experimental methods to establish causal answers to the questions at hand. The first two chapters of this thesis explore the role of market and government programs in mitigating the consequences of economic shocks for poor individuals in Indonesia. In the third chapter, I explore the strategic interactions among political activists in electoral campaigning.

Indonesia is the ideal context to study the impact of economic shocks on households and individuals. With more than 260 million inhabitants it is the world's fourth largest country. Furthermore, 31% of the population still works in the agricultural sector which exposes them to large year-to-year fluctuations in income (The World Bank, 2019). Hence, my research is relevant to a large number of individuals even before extrapolating to different contexts. The reliance of the agricultural sector in Indonesia on rain-fed irrigation has econometric benefits. It allows me to exploit plausibly random variation in rainfall to identify the causal impact of (agricultural) productivity shocks for a large number of individuals. Finally, Indonesia has several high-quality data sources dating back several decades. This enables me to draw on several data sources to show the robustness of my results.

Reducing households' exposure to negative economic shocks is a policy priority. Negative income shocks can have severe short-term, and long-term consequences, in particular for already poor populations. Children born during times of economic stress have worse later life outcomes (Maccini and Yang, 2009; Adhvaryu et al., 2019). It can also have more immediate consequences, such as short-term effects on mental health and suicides (Haushofer and Shapiro, 2016; Baird et al., 2011; Carleton, 2017). In this thesis I explore two mechanisms through which the effects of negative agricultural productivity shocks on individuals and households in Indonesia can be mitigated.

In the first chapter of this dissertation, I explore whether formal wage work in the agricultural sector in Indonesia has the potential to dampen the negative income effects of agricultural productivity shocks. Wage work can serve as an insurance against negative income shocks for agricultural households (Kochar, 1995). However, wage jobs in non-agricultural sectors often require higher levels of education and, often, migration to different cities (Kleemans and Magruder, 2018). For individuals unwilling or unable to switch sectors or locations, wage-jobs in the agricultural sector can be the only alternative. The insurance function of such jobs, however, relies on wages and employment opportunities not being too correlated with agricultural productivity. I find that wage work in the agricultural sector provides at best partial insurance against negative income shocks. Wages do not significantly adjust to small negative shocks, however, wages do adjust downwards when the difference between the prevailing wage and the competitive wage becomes too large.

Chapter 1 of my dissertation contributes to the ongoing debate around the existence and consequences of downward nominal wage rigidity by providing a theoretical framework and supporting empirical evidence that can explain the contradictory empirical results of previous studies (Kaur, 2019; Franklin and Labonne, 2019). This chapter also speaks to the literature on the efficiency of labor market reactions to productivity shocks in the presence of multiple sources of labor market frictions. In particular, my results suggest that firing frictions might diminish the negative employment effects of downward nominal wage rigidity, thereby potentially shifting the burden from workers to firms.

In the second chapter of this dissertation my coauthors and I explore the role of a government-led conditional cash transfer program in Indonesia in fighting a particularly dire consequence of poverty – suicides. Suicide is a pressing public health concern, causing 800,000 deaths per year globally (WHO, 2014). The link between negative income shocks and suicides has been well established (Becker and Woessmann, 2018; Cutler et al., 2001; Carleton, 2017). However, there is little evidence of causal effect of government-sponsored poverty alleviation programs and their interaction with negative economic shocks on suicidal behavior.

Poverty alleviation programs provide, in theory, better protection against volatile economic shocks than formal wage work considered in Chapter 1. The payment of benefits is uncor-

related with agricultural productivity shocks and thus can work as a more effective insurance mechanism. To identify the causal effect of conditional cash transfer programs on suicides, my coauthors and I deploy a difference-in-differences strategy exploiting the nationwide roll-out of the Program Keluarga Harapan (PKH) conditional cash transfer program. We find that the cash transfer program reduced the suicide rate by, on average, 18% (0.36 suicides per 100,000 people). We also show that agricultural productivity shocks, proxied by rainfall, are negatively correlated with the suicide rate. However, the PKH program also mitigates the effect of agricultural productivity shocks on the suicide rate. This suggests that economic safety nets can play an important role in ensuring the poor against devastating negative income shocks.

Taken together, Chapter 1 and 2 emphasize the that formal and informal institutions such as government programs and, to a lesser extent, market rigidities can reduce the severity of negative income shocks which in turn can lead to less volatility in economic and mental well-being and, literally, save lives. While we have a relatively good understanding of the economic consequences of such programs, we are just starting to explore their psychological benefits. As more and more governments implement or consider conditional cash-transfer programs, comprehensive cost-benefit analysis of such programs, including their effect on suicides, provides crucial guidance to policymakers.

In the third Chapter of my dissertation my coauthors and I explore the role of strategic interactions between political activists in their door-to-door canvassing effort. We present the results of a natural field experiment conducted in cooperation with a major political party in a western European democracy. We examine whether and how party supporters' canvassing efforts depend on their beliefs about the canvassing efforts of their fellow party supporters. The experiment is set in the context of a large door-to-door canvassing campaign in the run-up to a nationwide general election. We focus on canvassing as it is an important form of political activism which directly influences political outcomes (Pons, 2014) and is present in many democracies around the world (Plasser and Plasser, 2002).

We find that on, on average, political activists' effort choices exhibit strategic substitutability. We estimate that party supporters who learn that their peers plan to exert more effort than they previously expected canvass 14.39 (38%) fewer doors. We further find that the effects are

driven by party supporters with weaker social ties to the party (as proxied by prior canvassing experience, whether the respondent is a party member, and party membership duration).

This chapter contributes to a growing body of literature investigating the motivation and behavior of political activists such as party supporters (Perez-Truglia and Cruces, 2017) and protesters (Cantoni et al., 2019; González, 2018; Enikolopov et al., 2019; Cantoni et al., 2016; Enikolopov et al., 2017; Passarelli and Tabellini, 2017; Acemoglu et al., 2018; Madestam et al., 2013). We make several contributions to existing literature. First, the experimental data we use is distinct in two main respects: we draw on a unique combination of survey and behavioral outcome data collected through a smartphone application. Second, we provide evidence from a natural field experiment on behalf of the party, in which participants are not aware of their participation in an experiment, alleviating concerns about experimenter demand effects (de Quidt et al., 2018). Finally, we investigate the participation in a two-month-long campaign compared to the turnout to a single political event (eg a protest), our setting also allows us to study responses along the intensive margin and dynamic effects over time.

Chapter 1

The Impact of Economic Shocks on Wage Dynamics in Indonesia

Abstract When does downward nominal wage rigidity prevent labor market adjustments to negative productivity shocks? I provide evidence that the cost to the economy caused by downward wage rigidities are limited in magnitude. In particular, I use transitory labor demand shocks in agricultural labor markets induced by plausibly exogenous rainfall fluctuations to show that downward adjustment of nominal wages is a function of the cost of maintaining the nominal wage at the prevailing level: nominal wages adjust downwards when these costs are high, while they remain fixed when the costs of keeping the wage level constant are relatively low. To shown this, I exploit the fact that past positive productivity shocks exogenously increase these costs by raising the prevailing nominal wage level. By providing convergent evidence from large repeated cross-sections and data on individual nominal wage changes, I rule out that my results are driven by compositional changes in my sample or selective migration. Finally, I present a theoretical framework with reference dependent utility over past wages that is consistent with the observed wage rigidities.

Acknowledgements I am grateful to Climent Quintana-Domeque for his guidance and encouragement in working on this chapter. I also thank Christopher Roth for valuable feedback throughout the process of writing this paper. For helpful discussions and feedback, I thank Simon Franklin, Douglas Gollin, Surpreet Kaur, Clement Imbert, Simon Quinn, Jan Priebe, Giulio Schinaia, Timo Reinelt, and Sasha Parameswaran.

1.1 Introduction

Labor market frictions can have important economic consequences. This is particularly true for developing countries, where labor typically comprises a larger share of factor inputs for production. Well-functioning labor markets in developing countries are important to mitigate the impact of economic shocks (Colmer, 2018) and for efficient allocation of labor resources (Bryan and Morten, 2019). The wage - as the price of labor - plays a particularly important role in the functioning of labor markets: partial or non-adjustment of wages to economic shocks can have vital consequences for profits, employment and production. In particular, downward rigidity of nominal wages is a potential source of large economic distortions, which may result in the misallocation of factor inputs. A long literature has examined the economic effects of downward nominal wage rigidity, ranging from employment effects (Card, 1990; Kaur, 2019) to aggravated business cycle volatility (Clarida et al., 1999; Galí, 2009; Benigno and Ricci, 2011). However, few studies have used credible causal identification strategies.

The results of credible empirical work testing for the existence of downward nominal wage rigidity are contradictory. For example, Kaur (2019) finds strong evidence of downward rigidity of nominal agricultural wages in response to agricultural productivity shocks in India. Franklin and Labonne (2019), on the other hand, find a large degree of downward flexibility of nominal wages following a typhoon in the Philippines. There are multiple potential explanation for these divergent results, including data-structure, identification strategy, and context. Understanding whether and why downward nominal wage rigidities exist is crucial for our understanding of the distributional and aggregate consequences of negative economic shocks.

I develop a theoretical framework and provide supporting empirical results that is consistent with the existence of both downward nominal wage rigidity and flexibility. Using plausibly exogenous variation in agricultural labor productivity, I show that agricultural wages in Indonesia are at least partially downwardly flexible. To explain the heterogeneous adjustment patterns of nominal wages to economic shocks, I introduce the effective productivity gap that measures the cost of not adjusting the wage downwards. In particular, I exploit differences in the prevailing nominal wage induced by past productivity shocks to show that wages *only* adjust downwards for *large* values of the effective productivity gap. I thus provide evidence that treating the existence

of downward nominal wage rigidity as binary is an oversimplification.

In contrast to much of the existing literature, I find that downward nominal wage rigidity does not necessarily cause negative employment effects. I argue that the negative effects can be partially offset by rigidities in the firing of workers. In particular, I provide evidence that the employment effects of positive labor demand shocks are persistent and that the non-downward adjustment of nominal wages to negative productivity shocks does not necessarily lead to a decrease in employment. This has potentially important implications for the distribution of the cost of downward nominal wage rigidity.

In my empirical analysis I overcome several challenges commonly faced by other studies in this literature. A common problem in many of the empirical studies on downward rigidity of nominal wages is the identification of exogenous labor demand shocks. To overcome this problem I exploit the fact that rice yields, the main agricultural crop in large parts of Indonesia, depend on monsoon rainfall (Naylor et al., 2007; Levine and Yang, 2014). I use quasi-random variation in rainfall as proxies for agricultural productivity. I validate this proxy by demonstrating that it predicts rice yields in Indonesia over the sample period from 1993 to 2010. Data structure is another problem faced by many studies. Studies using cross-sectional wage level data are generally subject to aggregation bias due to endogenous sample composition (Bils, 1985; Keane et al., 1988; Keane and Prasad, 1996). However, studies who use individual wage changes usually lack the power to study employment effects.[1] I avoid this problem by providing convergent evidence from both large repeated cross-sections and individual-level wage change data. This allows me to plausibly rule out that my results are driven by selective migration or labor force participation.

The most important contribution of this paper is to provide theory and evidence that reconciles the seemingly contradictory results on the existence of downward nominal wage rigidity. My results suggest that nominal wages are partially downwardly flexibility, which is in line with a large part of the existing literature on wage dynamics and economic shocks in developing countries. Jayachandran (2006), Franklin and Labonne (2019), and Gignoux and Menéndez (2015) all find some degree of elasticity of wages with respect to negative economic shocks. On the other hand, my results contrast with Kaur (2019)'s finding of fully downwardly rigid nominal

[1] Card (1990) and Franklin and Labonne (2019) are rare examples of studies using wage change data and are powered to study employment effects.

wages.[2] Importantly, my results indicate the costs of downwardly rigid wages to the economy are endogenously limited as firms will adjust their wages when the cost of not doing so becomes to high.

I also contribute to the literature on the efficiency of labor market reactions to economic shocks by showing that, in presence of several frictions, the distributional and welfare consequence of downward nominal wage rigidity are ambiguous. Flexible labor markets can be important mechanisms to mitigate the impact of negative economic shocks, especially when those shocks are predictable (Rosenzweig and Udry, 2014; Colmer, 2018; Gröger and Zylberberg, 2016). On the other hand, factors hindering labor mobility or job search can lead to large individual and aggregate economic inefficiencies (Bryan and Morten, 2019; Franklin, 2017). My results show that wage and employment rigidities have the potential to partially offset each other. Rigidities in the firing process in particular have the potential of mitigating the negative employment effects of downward nominal wage rigidity.

Finally, speaking to the to broader literature investigating the impact of labor market frictions in developing countries, my findings emphasizes the potential of wage setting as an potentially import source of frictions. This is an important addition to recent experimental work which has mostly focused information and, to a lesser extent, spatial frictions as important factors contributing to labor market inefficiencies (e.g. Abebe et al., 2018; Abel et al., 2019; Franklin, 2017; Bassi and Nansamba, 2019; Carranza et al., 2019).

I proceed as follows. In Section 1.2, I present a theoretical framework that introduces the effective productivity gap. In Section 1.3 I present the context and data of the this paper. In Section 1.4, I introduce the empirical specifications used and in Section 1.5 I present the results. Finally, I conclude in Section 1.6.

1.2 Modeling Rigid Wages with Reference Dependence

In this section, I extend Kaur (2019)'s theoretical framework to develop a framework in which wage setting that is governed by two regimes: First, for a sufficiently large difference between the

[2] Kirchberger (2017)'s result of positive short term wage effects of earthquakes is a cautionary tale of the need to account unexpected of effects of economic shocks, such as the positive economic effects of increased reconstruction efforts.

reference wage and the lower market clearing wage it is optimal for firms to simultaneously offer the market clearing wage. Second, for a sufficiently small difference firms will optimally offer the reference wage above the market clearing wage.[3] I operationalize this pattern by introducing the effective productivity gap, a measure for the opportunity cost of not adjusting nominal wages to negative productivity shocks.

1.2.1 Set-up

The basic set-up of the model follows Kaur (2019) and assumes simultaneous wage setting and exogenous producer prices in a static setting. There is a continuum of workers i with a mass normalized to one and workers' payoff from not working is set to zero. The utility function of worker i is:

$$U_i = u\left(\frac{w}{p}\right) - \psi_i e\left(1 + \frac{1-\lambda}{\lambda}\mathbb{1}_{w<\bar{w}_{t-1}}\right) \tag{1.1}$$

$u(\cdot)$ is a continuous, increasing, twice-differentiable, and concave function. w is the contemporaneous nominal wage, p is the price level. Workers exhibit reference dependent utility, which causes them to work with less effort if they are paid below their reference wage.[4] $\bar{w}_{t-1}$ is last period's average nominal wage which is assumed to be the reference wage.[5] Note that I omit the t subscripts for contemporaneous variables due to the static nature of the model. The reference wage is the only intertemporal component of the model. $\mathbb{1}_{w<\bar{w}_{t-1}}$ is an indicator function, which is equal to one if the paid wage is below the reference wage. The individual disutility of labor ψ_i indexes all workers and is uniformly distributed on $[0, \bar{\psi}]$. e is the level of effort workers exert, which is discontinuous around the reference wage:

$$e = \begin{cases} 1, & \text{if } w \geq \bar{w}_{t-1} \\ \lambda, & \text{if } w < \bar{w}_{t-1} \end{cases} \tag{1.2}$$

If workers are paid at least the reference wage, they exert effort $e = 1$. Workers reduce their effort to $\lambda \in (0, 1]$ if they are paid less than the reference wage. $\lambda = 1$ means that workers exhibit no reference dependence. Note that workers decrease effort by exactly the amount

[3] Kaur (2019) only proves the existence of fully rigid wages.

[4] This specification can be interpreted as fairness norms and is based on research suggesting that workers reduce effort in response to wage which are perceived as unfair: Fehr et al. (2009) summarize the existing literature on fairness norms and labor markets and find that effort responds particularly strongly to nominal wage cuts below a 'fair' wage. Their review builds on theoretical models (Akerlof and Yellen, 1990), qualitative evidence (Bewley, 1995), and experimental findings (Fehr and Gächter, 2000b). However, fairness norms can also take different forms. For example, a norm for equity can be important for worker effort in response to wage setting (Abeler et al., 2010).

[5] Throughout this analysis I follow Kaur (2019) in focusing on pure strategy Nash Equilibria. This means that all workers will be paid the same wage each period so that reference wage also equals worker i's last period wage.

needed to offset the disutility induced by the reference dependence. This implies the following aggregate labor supply equation:

$$L^s = \frac{u(\frac{w}{p})}{\bar{\psi}} \tag{1.3}$$

There are J identical firms, where J is arbitrarily large. Firm j has the following profit function:

$$\pi_j = p\theta f(L_j e) - w_j L_j \tag{1.4}$$

$\theta > 0$ is the overall level of productivity, L_j is labor employed by firm j, $L_j e$ is effective labor, and w_j is the nominal wage set by firm j. $f(\cdot)$ is a continuous, increasing, twice differentiable, and concave production function. I assume that firms simultaneously set wages to maximize their profits. Furthermore, I assume that firms hire labor in descending order of their wage offer. If multiple firms offer the same wage, the order of hiring will be random between all firms offering the same wage. I further assume that workers with the lowest disutility of labor will get hired first.[6]

At a given wage w^* and constant effort by workers the first order condition of the firm (1.5) determines individuals' labor demand in either of those cases:

$$w^*(e, \theta, p) = pe\theta f'(eL^*), \text{ for } e \in \{\lambda, 1\} \tag{1.5}$$

Market clearing requires that labor demand equals labor supply:

$$JL^* = \frac{u(\frac{w}{p})}{\bar{\psi}} \tag{1.6}$$

Kaur (2019) shows that, in the case without reference dependence, equations 1.5 and 1.6 fully determine the labor market equilibrium.[7]

1.2.2 The Impact of Reference Dependence: Partial Downward Rigidity

Reference dependence introduces a discontinuity into to the marginal product of labor and, hence, firms' profit function. To see this note that the first order condition of the firms is discontinuous around last period's average wage $\bar{w}_{t-1}$. In this case, the first order condition for a given wage w is:

[6] As described by Kaur (2019), the specification of an allocation mechanism is necessary to prove the propositions.

[7] For convenience of the reader the proof is repeated in Appendix Section 1.A.1.

$$w = \begin{cases} p\theta f'(L), \text{ if } w \geq \bar{w}_{t-1} \\ p\theta\lambda f'(\lambda L), \text{ if } w < \bar{w}_{t-1} \end{cases} \tag{1.7}$$

For $w \geq \bar{w}_{t-1}$ the first order condition corresponds to the benchmark case of full effort $e = 1$. If $w < \bar{w}_{t-1}$ the first order condition reflects the lower productivity of each unit of labor at the lower level of effort, $e = \lambda$.

Let θ_R be implicitly defined as $w^*(1, \theta_R, p) = \bar{w}_{t-1}$, i.e. the productivity level at which the reference wage equals the market clearing wage at full effort. Further denote $L^*(e, \theta, p, w^*)$ as the labor demand at $w^*(e, \theta, p)$. The labor demand at the reference wage ($w = \bar{w}_{t-1}$) is determined by the first order condition (1.5) and is $\bar{L}(1, \theta, p, \bar{w}_{t-1})$.[8] $\pi^*(e, \theta, w^*, L^*)$ and $\bar{\pi}(1, \theta, \bar{w}_{t-1}, \bar{L})$ are the profits at the respective wage levels. Note that for simplicity of presentation, I only show the arguments of the function which are relevant to the analysis from here onwards.

Proposition 1. Asymmetric adjustment of wages to shocks

For λ sufficiently close to 1 there exists a θ'_R with $0 < \theta'_R < \theta_R$ such that

1. *For $0 < \theta \leq \theta'_R$, all firms j setting $w_j = w^*(\lambda, \theta, p) < \bar{w}_{t-1}$ is a pure strategy Nash Equilibrium. Furthermore, $w_j = \bar{w}_{t-1}$ for all j is the only other candidate pure strategy Nash Equilibrium and $\int_0^1 \pi^*(\lambda, \theta, w^*)dj > \int_0^1 \bar{\pi}(1, \theta, \bar{w}_{t-1})dj$.*
2. *For $\theta'_R < \theta \leq \theta_R$ all firms setting $w_j = \bar{w}_{t-1}$ is the unique pure strategy Nash Equilibrium.*
3. *For $\theta_R < \theta$ the wage will correspond to the benchmark case $w_j = w^*(1, \theta, p)$.*

Proof. See Appendix 1.A.2. ■

Proposition 1 distinguishes between three ranges of productivities. It states that as long as the decrease of workers' effort in response to being paid below the reference wage is sufficiently small, for low levels of productivity $\theta < \theta'_R$ firms would not deviate upwards from the benchmark wage at low effort ($w^*(\lambda, \theta, p)$) if all firms set this wage. It further states that total profits at the only other candidate pure strategy Nash Equilibrium are smaller than at the aforementioned equilibrium. Thus, while downward adjustment of wages is not necessarily the unique pure strategy Nash Equilibrium, it can be argued that it is the only plausible equilibrium, as long as

[8] I will omit the reference to the effort from now on, as effort is always 1 at the reference wage.

coordination costs are relatively low.[9]

At medium levels of productivity $\theta'_R < \theta \leq \theta_R$, the only pure strategy Nash Equilibrium wage will be all firms setting the reference wage and there will be excess labor supply. For high levels of productivity ($\theta \geq \theta_R$), the reference wage does not bind and wages will correspond to the benchmark case.

Proposition 1 implies that, for sufficiently large productivity shocks, wage are likely to adjust down to the competitive benchmark. This limits the potential magnitude of distortive effects of downward nominal wage rigidity.

1.2.3 The Impact of Past Productivity

To facilitate the analysis of past shocks, Proposition 2 introduces the *effective productivity gap* denoted by ρ. It is a measure of how costly it would be to maintain nominal wages at last period's level. It incorporates three factors which influence these costs: current productivity lowers the cost by increasing the current benchmark nominal wage level; last period's productivity increases the cost by raising last period's nominal wage level, and inflation decreases the cost by increasing current real wage levels. Note that throughout this analysis I normalize the past price level p_{t-1} to one such that inflation is equivalent to the contemporaneous price level p.

Proposition 2. The effective productivity gap

Define the effective productivity gap as $\rho(\theta_t, \theta_{t-1}, p) = \frac{\theta_{t-1}}{\theta_t} p^{-1}$. *Further define the reference wage implicitly by* $\bar{w}_{t-1} = \theta_{t-1} f'(L)$ *and equation* (1.6). *The following holds:*

1. $\rho \leq 1 \Leftrightarrow \theta \geq \theta_R$, *and there exists* ρ^* *such that* $\rho^* \geq \rho > 1 \Leftrightarrow \theta'_R \leq \theta < \theta_R$.

2. *An increase in* θ_{t-1} *widens the range of* θ_t *for which wages adjust downwards.*

Proof. See Appendix 1.A.3. ■

The advantage of using ρ is that one can directly analyze the impact of past shocks. Proposition 2 assumes that the last period's wage was set according to the the benchmark case.[10] Part one states that there is a one-to-one mapping from ρ to θ such that ρ can be used to

[9] In the context of Indonesian village labor markets, this is a reasonable assumption. While there are many potential employers in each village, there is a substantial amount of cooperation going on. According to the 1993 Indonesian Village Census (PODES) 60% of villages in districts included in this analysis have a farmers' cooperative. In Java which is the most populous Island, this share goes up to 80%. This indicates a large amount of cooperation between farmers, which supports the assumption.

[10] I show empirically that considering further lags of productivity does not alter the results.

determine the predicted wage level. The second part states that increases in past productivity make downward adjustment of current nominal wages more likely, by increasing the cost of maintaining last period's wage level. This prediction is the main driver of the interpretation of my empirical results.

1.2.4 The Effect on labor Supply

The model yields the following predictions for the employment relative to the benchmark case.

Proposition 3. The effect on labor supply

The following holds for total employed labor L^e relative to labor employed in the benchmark case $L^*(1, w^*)$:

1. $L^*(1, w^*) > L^e = \bar{L}(1, \bar{w}_{t-1})$ *if* $\rho \in (1, \rho^*)$
2. $L^*(1, w^*) = L^e = \bar{L}(1, \bar{w}_{t-1})$ *if* $\rho \leq 1$
3. *The prediction for* $L^e = L^*(\lambda, w^*(\lambda))$ *relative to* $L^*(1, w^*(1))$ *if* $\rho > \rho^*$ *is ambiguous.*

Proof. See Appendix 1.A.4. ■

The first part of Proposition 3 implies the prediction that if downward nominal wage rigidity binds and wages do not adjust, there is less labor demanded than in the benchmark case. The second part of the proposition states that if the reference wage is below the benchmark case, there is no difference in employment. The third part states that without further assumptions the model's prediction about total labor supply when wages adjust downwards is ambiguous. The intuition behind this prediction is that reduced effort decreases marginal productivity of effective labor, while it increases the amount of labor needed per unit of effective labor. This leads to an ambiguous direction of the net employment effect.

1.2.5 The Impact of Inflation

Note that $\frac{d\rho}{dp} = -\frac{\theta_{t-1}}{\theta_t} p^{-2} < 0$, that is the effective productivity gap decreases in inflation. This means that inflation has two effects in the model with reference dependence:[11] First, it decreases the range of productivity realizations at which wages adjust downwards, by lowering the cost of maintaining the nominal wage at the reference wage level. Thus, binding downward

[11] If there is no reference dependence, inflation is neutral. To see this note that employed labor does not change with the price: $JL^* = \frac{u(\theta f'(L^*)}{\psi}$. Furthermore, nominal wages change one to one with the price: $\frac{dw^*}{dp} = p$. This means that real profits are not affected.

nominal wage rigidity become more likely when the productivity difference between two periods is relatively large. Second, inflation decreases the contemporaneous real wage, which decreases the range of productivities for which the contemporaneous nominal wage is below the reference wage. Hence, high inflation makes binding nominal wage rigidity less likely for small productivity differences between two consecutive periods.

1.2.6 Discussion

The model discussed above is static, but the results translate to a dynamic setting in which firms would choose wages in two consecutive periods. In principle, firms would have an incentive to set lower wages in the first period to lower the reference wage in the next period. However, for a large number of firms J, the contribution of a single firm's wage to the average wage is negligible. This implies that firms set wages without considering the impact on profits in future periods. If firms engaged in an intertemporal relationship with workers, this would maintain the qualitative predictions of the model although its quantitative predictions could change (Kaur, 2019).

Given the assumption of low coordination costs, one could think of modeling collusion between firms. However, even if firms perfectly colluded and acted as a monopsonist, the general pattern of wage setting would not change. The monopsonist would still weigh the cost of keeping the wage at the reference wage against the cost of reduced effort, which would lead to the same switching behavior pattern in wage setting. My results are compatible with either model and I do not take a strong stance on the competitiveness of agricultural labor markets in Indonesia.

The choice of last period's average nominal wage as a reference point is in line with survey evidence from other contexts (Kahneman et al., 1986; Kaur, 2019). However, I only show that this model of reference dependence is consistent with my empirical findings.

1.3 Context and Data

1.3.1 The Indonesian Agricultural Sector

The agricultural sector in Indonesia is to a large extent governed by seasonal monsoon rainfall. The timing of the monsoon and the total quantity of rainfall are highly correlated with the 'El

Niño Southern Oscillation' (ENSO) phenomenon. In El Niño years Indonesia generally experiences a significant delay and shortfall of rainfall. This shortfall has important implications for agricultural production, in particular for rice which is the main crop planted in large parts of Indonesia.[12] Planting of rice usually takes place during the onset of monsoon (September to November). On unirrigated fields, rice can only be planted after at least 20cm of cumulative rainfall. Following Naylor et al. (2007), I exploit this threshold for the definition of rainfall shocks. In normal years the rice harvest usually takes place from February to April followed by a second (smaller) dry growing season. Harvesting for the dry season takes place from June to August. In years of low rainfall, rice production is delayed so that some of the production is pushed into the dry season. Rice production during the dry season is not affected by the El Niño cycle (Falcon et al., 2004). Due to its location around the equator, temperature in Indonesia is relatively constant within or across years and it does not affect agricultural yields (Kleemans and Magruder, 2018).

Based on the 1996 Indonesian village census about 25% of the agricultural wetland relies exclusively on rain fed irrigation. While this may appear to be little, much of the installed irrigation capacity is to some degree reliant on rainfall.[13]

Agricultural labor markets in Indonesia are reported to be active with most agricultural households hiring day labor for all stages of the agricultural production cycle (LaFave and Thomas, 2016). However, labor market patterns demonstrate a substantial amount of cyclicality.[14]

There are also geographic differences in the timing of the agricultural production cycle. Figure 1.1 displays the differential patterns of monsoon onset in Indonesia. The differences are particularly pronounced between the northern Sumatra and Java.[15] Section 1.3.2 describes how I account for this cyclicality in agricultural production and labor markets.

[12] According to the Global Rice Science Partnership (GRSP) 77% of all farmers engage in rice planting (`http://ricepedia.org/index.php/indonesia`, last accessed April 2016).

[13] `http://ricepedia.org/index.php/indonesia`, last accessed April 2016.

[14] Over the sample period of 1998 to 2007, nominal agricultural wages in August were on average 14% lower than agricultural wage in February of the same year. This is in line with the finding of LaFave and Thomas (2016) that wages in the harvesting season are higher than for other agricultural activities.

[15] In Section 1.5.3 I show that my results are robust to excluding Sumatra from the analysis.

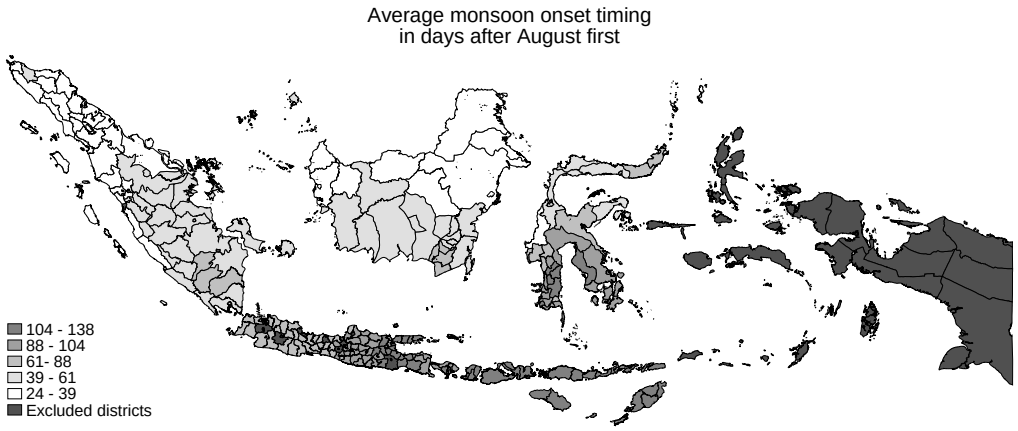


Figure 1.1: Differential Patterns of Monsoon Arrival

1.3.2 Data Description

Labor Market Data

This paper uses two different sets of Indonesian labor market data. The first is a nationwide repeated cross-section for the years 1998 until 2007. I combine National Socioeconomic Survey (Survei Sosial Ekonomi Nasional or SUSENAS) data from 1998 to 2004 with National Annual labor Market Survey (Survei Angkatan Kerja Nasional or SAKERNAS) data from 2005 to 2007. Both questionnaires contain very similar questions about last month's income and days worked last week. The most important reason for this particular choice is the timing of the surveys. As agricultural productivity during the dry season is not affected by rainfall, I only use observations that fall in the provincial monsoon season as defined by Maccini and Yang (2009).[16] I merge the cross-sectional survey data at the district level.[17] This leaves me with ten consecutive years of wage data collected in February.

In theory, using these two data sources in a repeated cross-section does not pose a problem. First, individuals in both surveys are sampled using the same household listing, which is roughly updated every five years for the population surveys. This means that they represent the same underlying population. Second, my identification strategy relies on comparing individual level

[16] This excludes SAKERNAS data from 1994 to 2004 as the month of data collection was August, which does not fall in the monsoon season for a large part of my sample. Later SAKERNAS data do not identify individuals at the district level, which makes them unsuitable for this analysis. SUSENAS data from 2005 onwards were collected in August and SUSENAS questionnaires before 1998 are missing the income questions, which is why they cannot be used in the analysis.

[17] Districts or Kabupaten are the second administrative unit in Indonesia. Whenever district boundaries changed after 1998, I collapse the observations using 1998 district boundaries.

wages within a given year and I do not pool data sources within a year. Hence, small differences in wording or ordering of questions between years should not affect the results.[18]

The surveys do not directly ask for wages, but they ask for income and working time of individuals who consider wage work as their main job. Hence, my sample does not include workers which only occasionally sell their labor on the market. Furthermore, I focus on wages of non-casual workers because of data constraints.[19] I calculate the daily wages of individuals as monthly income divided by 4.5 times days worked last week.[20] I restrict the sample to individuals working in agriculture and I drop the top and bottom percentile of the wage distribution in each year. In total, I observe wages for 7.6% of individuals working in agriculture, which contrasts with the findings of active labor markets (e.g. in LaFave and Thomas, 2016). This is largely explained by the informal structure of agricultural labor markets and known underreporting of labor market activity (Smith et al., 2002). While this affects the external validity of my results, it is not a threat to the internal consistency of my findings.

The second data source is the Indonesian Family Life Survey (IFLS). The IFLS is a panel of around 12,000 Indonesian households, which are representative of 83% of the Indonesian population in 1993. As of now, there are four rounds of data (1993 1997, 2000, and 2007). The timing of the questionnaires falls broadly within the monsoon season, but the surveys tend to be fielded at the beginning rather than at the harvest stage. The questionnaires for the years 1993, 1997, and 2000 include detailed questions about current and past primary and secondary employment, including income, working time, and migration data for at least the last five years. The wage recall questions were not included in the last round of data collection. Given the unreliable nature of recall questions of longer periods of time, I only use the first year of recall data.[21] I define past wages as average monthly wage income last year divided by 4.5 times the

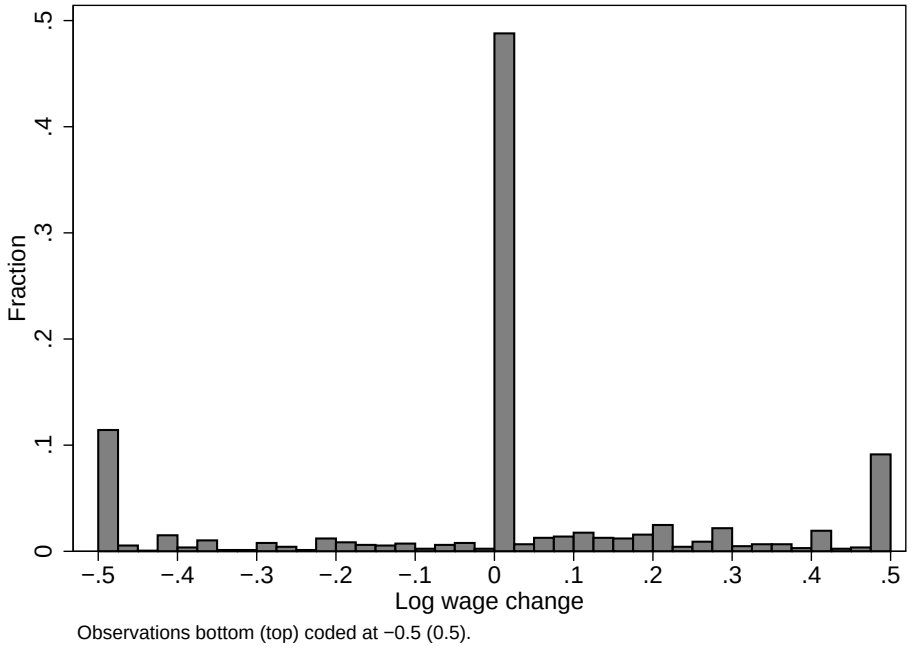
[18] In Section 1.5.3 I show that the results for just the SUSENAS sample are similar.

[19] I only explicitly observe income for casual workers for four of the ten years. In Section 1.5.3 I show that my results are robust to including these observations. Casual workers are defined as having more than one employer during the last month.

[20] Days worked is defined as days worked last week in all jobs. I observe days worked in the main job for the years 2005-2007, but I find only very small differences between the two measures (results available upon request).

[21] The evidence on quality of recall data in developing countries is mixed: Beegle et al. (2012) find that data quality for salient event such as harvest volume or fertilizer use is relatively high. However, they also find that data quality of labor use deteriorates with recall duration. Evidence on developed countries as summarized by Bound et al. (2001) also indicates that recall data on working hours is very noisily measured. Duncan and Hill (1985) find a significant deterioration of data quality of reported earnings with recall period. Notably, I found no study credibly assessing the data quality of recall data of more than two years in the past. All this led me to conclude, that using more than one year of recall data would jeopardize the reliability of my inference.

Figure 1.2: Distribution of Nominal Wage Changes in the IFLS



working hours in an average week last year.[22] To calculate the current wage I divide last month's wage income by hours worked in a normal week time 4.5. For individuals with a secondary job, I first check whether the first job is agricultural wage work. If that is the case I use the wage of the first job. If not, I check whether the secondary job is agricultural wage work. If that is true, I use the secondary wage as wage observation.

I use this wage data to construct first differences in log hourly wages for individuals working in the same job and who did not move to another district between the years. This restriction guarantees that the analysis is robust to sample selection issues, in particular, to selective migration and changes in labor force participation . I truncate wage changes by the top and bottom 1% as implausibly large changes might be caused by coding issues (Frankenberg and Karoly L, 1993).

Figure 1.2 shows the top and bottom coded distribution of log nominal wage changes observed in the IFLS data. Most notably, there is a large peak at zero. In total 49% of reported wage changes are zero with the remainder being split between of 22% of negative wage and about 29% of positive wage changes. The wage histogram does not immediately allow any conclusions about rigidities, rather one has to combine the histogram with economic shocks to

[22] Unfortunately, the IFLS surveys did not ask for days worked.

come to firm conclusions (Barattieri et al., 2014). I argue that this bunching at zero is at least partially driven by over-reporting of zero wage changes rather than a strong degree of nominal wages rigidity in both directions.[23] There could be other other sources of measurement error, but the converging empirical results suggest that it does not compromise the internal validity. In Section 1.4 I propose an empirical specification which is robust to measurement error in the magnitude of wage changes and only requires the sign to be accurate.

Naturally, each data source has advantages and disadvantages. By using two separate data sources I can overcome potential shortcomings of studies such as Kaur (2019) who relies only on cross-sectional and district aggregate data. Comparing results in both datasets allows me to tackle problems such as measurement error, selective migration, and endogenous sample composition provides me with a good measurement framework to rigorously answer the research question of interest.

To ensure that rainfall is a valid proxy for agricultural productivity, I restrict the sample of districts used in the analysis. First, I exclude the two most eastern, sparsely populated provinces Papua and Maluku from the cross-sectional data (they are not covered by the IFLS). The agricultural cycles in these provinces are very different from the rest of Indonesia (the monsoon season extends until July/August) and data quality in these provinces is inferior, in particular in the aftermath of the Asian financial crisis.[24] This only marginally affects the scope of my data as the two provinces are home to less than 2.5% of the Indonesian population (according to the 1995 intercensal survey). I also drop big cities defined as having more than 50,000 inhabitants in the 1930 census from both samples as rice yields are not responsive to rainfall in these locations (Levine and Yang, 2014; **?**).

This yields repeated cross-section data for individuals in 256 Indonesian districts in 26 provinces over ten consecutive years.[25] The IFLS data contain observations of wage changes in

[23] Recalling last year's income and working time correctly is likely to be more cognitively expensive to just use current income and working hours. Furthermore, as the questionnaire only asks for 'usual' income and hours, which could smooth negative or positive effect during the previous monsoon season. Both factors are likely to lead to serious over reporting of zero wage changes in the data. To see that this is not only true for negative wage cuts consider that inflation at the time of the IFLS surveys was between 9% and 12%. Thus, one would expect a larger positive side of the distribution (even without positive contemporaneous shocks).

[24] Budget constraints following the financial crisis led to smaller sample sizes during this period.

[25] Not all districts have observations in all years. For example, data were not collected in some districts in the province of Aceh in the early 2000s because of violent conflict.

162 districts in at least one of the following years: 1993, 1997, or 2000.[26]

Rainfall Data

I use the ERA-Interim/Land Reanalysis dataset which provides daily precipitation data from 1979 until 2010 for a grid of $0.25° \times 0.25°$.[27] I define rainfall at district level as rainfall at the grid point closest to the geometric center of the district. Reanalysis data is based on a mix of real weather observations (station and satellite data) and an atmospheric climate model. This allows the provision of uninterrupted observations for each grid point, which is particularly important as the definition of my rainfall shocks relies on daily precipitation data. A further advantage of reanalysis data is the homogeneous data quality across time and space, which alleviates the issue of endogenous weather station placement (Colmer, 2018).

I use the daily precipitation data to construct a measure of monsoon onset timing. I follow the literature in defining monsoon onset as the number of days after August first when cumulative rainfall of 20cm is reached. Onset delay is defined as deviation in days of monsoon onset from the district's long-term mean monsoon onset.[28]
This measure is similar to Kaur (2019)'s variable of rainfall in the first month of the monsoon season.[29] Appendix Table A1.1 shows summary statistics for both data sets.

In order to show that monsoon onset delay is a valid proxy for agricultural productivity, I show that monsoon onset delay affects rice yields in Indonesia. Rice is the most important crop for large parts of Indonesia and it is highly dependent on precipitation. I use provincial level rice yield data from 1993 to 2010 provided by the Indonesian Statistical Agency. I redefine the rainfall measures at province level, by averaging district level deviations at the province level.[30] The number of provinces increases over time, and I adjust the rainfall measure to in-

[26] Appendix Figure 1.C1 shows the spatial distribution of the included districts.

[27] As Indonesia is mostly stretched out along the equator, the grid size in kilometers is relatively constant.

[28] I do not standardize the the onset delay by district standard deviations, because the standard deviations depend mechanically on the geographic location of districts. Districts in the north-west are the first to receive monsoon rain, which means that early arrival is has a lower bound which mechanically lowers the variance. In Section 1.5.3 I provide some evidence that my results are not driven by different probabilities of experiencing rainfall shocks.

[29] I do not use the rainfall in the first month of the season as the onset delay dominates in the yields regression (results available upon request).

[30] This aggregation at province level means that the rice yield analysis is highly underpowered compared to my main analysis. I use much more fine grained spatial disaggregation to obtain my main results.

clude only districts within the new boundaries whenever province boundaries change. I further drop the provinces of Maluku and Papua as they are not used in the main analysis. This yields a province-year panel with a total of 28 provinces and 18 years.[31] Equation (1.8) shows the empirical specification I use to estimate the impact of monsoon onset timing on yields:

$$\ln(yield_{pt}) = \alpha_0 + \alpha_1 onset_{pt} + \delta_p + \delta_t + \varepsilon_{pt} \tag{1.8}$$

where $\ln(yield_{pt})$ is the natural logarithm of the rice yield in province p at time t, $onset_{pt}$ is the average deviation of monsoon onset with respect to the district long-term mean of province p at time t, and δ_p and δ_t are province and year fixed effects, respectively. To allow for intertemporal correlation of error terms, I cluster standard errors at province level. Given the small number of clusters these standard errors are likely to underestimate the true standard errors (Bertrand et al., 2004). Hence, I use the wild bootstrap procedure proposed by Cameron et al. (2008) to obtain p-values for the coefficient estimates.

Column 1 of Table 1.1 shows the results for the estimation. I find a statistically significant effect of monsoon onset on rice yields consistent with the idea of it being a proxy for agricultural productivity: an increase in the monsoon delay of one day is associated with a decrease of rice yields by on average 0.06 percent. To account for non-linearities in the rice yield function, I define non-linear positive and negative rainfall shocks.[32] Following Naylor et al. (2007) I define a negative shock as a delayed onset of thirty days or more and a positive shock as monsoon onset twenty days earlier than the local district mean. This asymmetric definition is justified by the fact that the distribution of rainfall shocks is right-skewed for my sample period (see Appendix Figure 1.C3). I then estimate the impact of the shocks on yields using equation (1.1):

$$\ln(yield_{pt}) = \alpha_0' + \alpha_2 Pos_{pt} + \alpha_3 Neg_{pt} + \delta_p + \delta_t + \varepsilon_{pt} \tag{1.9}$$

where Pos_{pt} (Neg_{pt}) is an indicator function which is one if province p in period t experienced a positive (negative) shock.

Column 2 of Table 1.1 demonstrates that these shock definitions capture the non-linearities of the rice yield curve. A positive monsoon onset shock increases rice yields by on average 1.3% and a negative shock decreases rice yields by on average 2.4%. These effects are statistically

[31] Four provinces have less than 18 observations.

[32] Appendix Figure 1.C2 displays a non-parametric estimation of the impact of monsoon onset on rice yields. The yield curve is flat for most of the distribution and only shows changes at the edges of the distribution, which suggests that a non-linear specification is appropriate to capture the effect of rainfall.

Table 1.1: Impact of Monsoon Onset on Rice Yields

		Impact on log province rice yields	
		(1)	(2)
	Rainfall measure		
1	Monsoon onset delay	-0.0006**	
		[0.02]	
2	positive shock		0.013*
			[0.076]
3	negative shock		-0.024**
			[0.042]
Observations		463	463
R-squared		0.96	0.97
Number of clusters		28	28
Province fixed effects		Yes	Yes
Year fixed effects		Yes	Yes

Notes: p-values obtained using wild bootstraps at province level with 1000 repetitions are in square brackets. Level of observation is Indonesian province. Sample is restricted to provinces used in the main analysis. Positive shocks are defined as monsoon onset of at least 20 days before the local mean. Negative shocks are defined as a monsoon onset of at least 30 days after the local mean. Monsoon onset is defined as reaching cumulative rainfall of 20cm after August first. *** p<0.01, ** p<0.05, * p<0.1.

significant at the 10% and 5% level, respectively.[33] This discrete definition allows the separation into three different productivity values θ: θ^H is the productivity induced by a positive shock, θ^M is the productivity in the absence of any shock, and θ^L is the productivity following a negative shock. In additional to the obvious ordering $\theta^H > \theta^M > \theta^L$, the results of the rice yields estimation allow an ordering by the severity of shocks. The coefficient on the negative shock is almost twice as large as the coefficient on the positive shock, which suggests that yields are more elastic to negative rainfall shocks.[34] The point estimates lead me to believe that negative shocks have a larger impact than positive shocks. This can be expressed as $\frac{\theta^H}{\theta^M} < \frac{\theta^M}{\theta^L} < \frac{\theta^H}{\theta^L}$.[35]

[33]Note that the difference in magnitudes is not driven by the asymmetric definition of shocks; with a symmetric cutoff of twenty days in either direction, the difference in effect sizes remains basically unchanged (results available upon request).

[34]This is plausible, since the negative impact on rice yields is likely to be underestimated, because crops are endogenously chosen in response to rainfall shocks (Ricepedia, `http://ricepedia.org/index.php/indonesia`, last accessed April 2016). On the most affected rice paddies will be substituted by other crops, which have lower yields than rice in normal years. Not surprisingly given the high level of aggregation of this analysis, I am not able to reject the hypothesis of $\alpha_2 = -\alpha_3$ when I cluster the standard errors on district level (the p-value is 0.34).

[35]Technically the estimated impact of shocks on rice yields implies $\left|\theta^H - \theta^M\right| < \left|\theta^L - \theta^M\right|$ which is a sufficient condition for the stated inequalities. I use the different representation to draw the analogy with the effective

1.4 Empirical Specifications

The empirical specification of this paper closely follows Kaur (2019). I start by considering the impact of contemporaneous shocks, before exploiting the dynamic structure of rainfall shocks. Lastly, I split negative shocks by their intensity to provide further tests of my results.

Testing whether nominal wages decrease with contemporaneous negative productivity shock is a straightforward approach to analyse nominal wage rigidity. Equation (1.10) describes how I implement this empirically:

$$w_{idt} = \beta_0 + \beta_1 Pos_{dt} + \beta_2 Neg_{dt} + \boldsymbol{\delta} X_{idt} + \delta_d + \delta_t + \varepsilon_{idt} \tag{1.10}$$

where w_{idt} is the log nominal wage of individual i in district d at time t, Pos_{dt} is a positive rainfall shock in district d and season t and Neg_{dt} is a negative rainfall shock in district d and season t, X_{idt} is a vector of predetermined individual characteristics (age, age squared, sex, a dummy for living in a rural area and an indicator for primary school completion) of individual i in district d at time t, and δ_d and δ_t are district and time fixed effects, respectively.
β_1 (β_2) can be interpreted as the percentage difference of wages in districts with a positive (negative) rainfall shock compared to wages in districts without any shock within a given agricultural season t. District fixed effects account for permanently different price or wage levels between districts and time fixed effects account for time varying levels of inflation.
In the benchmark model without reference dependence, wages fully adjust to the changes in productivity implied by the rainfall shock and therefore $\beta_1 > 0$ and $\beta_2 < 0$. With reference dependence, the prediction of the model depends on the severity of the shock. If the effective productivity gap ρ is always below the threshold ρ^*, wages are fully downwardly rigid and the prediction is $\beta_2 = 0$.

In a next step, I exploit the intertemporal patterns of rainfall shocks to further test the predictions of the model. Equation (1.11) demonstrates the specification I use to implement these tests:

$$\begin{aligned} w_{idt} =& \gamma_0 + \gamma_1 Pos_{dt} + \gamma_2 NonPos_{d,t-1} Neg_{dt} + \gamma_3 Pos_{d,t-1} Neg_{dt} + \gamma_4 Pos_{d,t-1} None_{dt} \\ &+ \boldsymbol{\delta} X_{idt} + \delta_d + \delta_t + \varepsilon_{idt} \end{aligned} \tag{1.11}$$

productivity gap defined in the model section.

where $NonPos_{d,t-1}$ is an indicator of no positive shock last year, and $None_{dt}$ indicates that there was no shock in period t. The omitted category in this equation is districts with no shock this year and no positive shock last year. The gamma-coefficients should be interpreted as comparing wages to this category. In the benchmark case of flexible wages, only the contemporaneous realizations of shocks matter with the predictions of $\gamma_1 > 0$, $\gamma_2 = \gamma_3 < 0$, and $\gamma_4 = 0$. Once reference dependence is introduced, Propositions 1 and 2 imply that the exact predictions depend on the size of the threshold effective productivity gap ρ^*.

For a high threshold ($\rho^* > \rho^{HL}$) nominal wages never adjust downwards and the predictions for the coefficients are $\gamma_2 = 0$ and $\gamma_3, \gamma_4 > 0$. For $\rho^{HL} > \rho^* > \rho^{ML}$ nominal wages only adjust to negative shocks when they are preceded by positive shocks. This changes the predictions to $\gamma_2 = 0$, $\gamma_3 < 0$ and $\gamma_4 > 0$. With $\rho^{HL} > \rho^{ML} > \rho^*$ nominal wages always adjust downwards to a negative productivity gap, but remain high following a lagged positive shock which is not followed by a negative shock. This translates into the following predictions for the coefficient in equation (1.11):[36] $\gamma_2 < 0, \gamma_3 < 0$, and $\gamma_4 > 0$. Lastly, consider the case in which inflation causes the effective productivity gap of the lagged positive shock to fall below 1 ($p \geq \frac{\theta_{t-1}}{\theta_t}$ such that $\rho^{MH} \leq 1$). In this case the downward rigidity does not bind for the lagged positive shock, so that the prediction changes to:[37] $\gamma_2 < 0, \gamma_3 < 0$, and $\gamma_4 = 0$.

To further strengthen the support for my proposed mechanism, I split negative shocks by intensity: I define strong negative shocks as observations of shocks which fall above the median onset delay for all districts which experience a negative shock. Weak negative shocks are defined as experiencing a negative shock and having monsoon onset delay below the median shock strength. Equation (1.12) is the empirical specification for this analysis:

$$\begin{aligned} w_{idt} = \xi_0 + \xi_1 Pos_{dt} + \xi_2 NonPos_{d,t-1} Neg^w_{dt} + \xi_3 NonPos_{d,t-1} Neg^s_{dt} + \xi_4 Pos_{d,t-1} Neg^w_{dt} \\ + \xi_5 Pos_{d,t-1} Neg^s_{dt} + \xi_6 Pos_{d,t-1} None_{dt} + \boldsymbol{\delta} X_{idt} + \delta_d + \delta_t + \varepsilon_{idt} \end{aligned} \tag{1.12}$$

where Neg^w_{dt} indicates a weak negative shock in district d at time t, and Neg^s_{dt} is a strong negative shock in district d at time t. This separation yields one clear prediction in terms of magnitude of the productivity gap: It is largest when a positive shock is followed by a strong

[36] It also implies $\beta_2 < 0$ in equation (1.10).

[37] Note that in this case one might also observe a noisily measured zero for γ_3, because the coefficient captures the difference between keeping the nominal wage at the reference level, while the control district increase their wages with inflation. This effect might be to small to show up in the data.

negative shock. This means that nominal wages are most likely to adjust downwards following this shock combination and thus that the coefficient ξ_5 is most likely to be negative.[38]

To rule out that selective migration and other factors drive the results of the cross-sectional estimation, I use first-difference wage data from the IFLS. This allows me to directly test for individual nominal wage changes, avoiding an endogenous composition of the labor force. Unfortunately, I only have three years of data, which limits identifying variation in rainfall shocks: Fewer than two percent of observations experience a contemporaneous positive onset shock and there is no observation that experiences a lagged negative shock (Appendix Figure 1.C4 shows the distribution of current and lagged district onset delay for the IFLS sample). This small sample size does not allow precise estimates of the impact of positive shocks, which is why the following specifications omit the positive shock variable and observations with positive shocks.[39] However, this does not limit the conclusions I can draw from the data, as I have sufficient observations for all the relevant shock categories to test the predictions of my model.

Parallel to the wage level analysis, equation (1.13) is a simple way to directly test for downward rigidity of nominal wages:

$$\mathbb{1}_{\Delta w_{idt} \geq 0} = \beta_0' + \beta_2' Neg_{dt} + \delta_t + \varepsilon_{idt} \tag{1.13}$$

where $\mathbb{1}_{\Delta w_{idt} \geq 0}$ is an indicator of non-negative wage changes. Using this variable instead of the actual wage changes is justified on several grounds.[40] First, by collapsing the wage data into a dummy variable I reduce noise from measurement error in the size of wage changes.[41] Second, the indicator function suffices to test for the existence of downward nominal wage rigidity which the primary question of interest. And third, by using this dummy instead of wage change, I make the analysis robust to relatively high levels of inflation.

Note that in this specification β_2' compares the probability of a negative nominal wage change in districts with a negative shock to districts without a negative shock. If I used wage changes instead of $\mathbb{1}_{\Delta w_{idt} \geq 0}$, the comparison group would experience average wage growth in line with inflation. This means an estimate of $\beta_2 < 0$ could also be caused by nominal wages staying at

[38] I cannot consider multiple, consecutive past positive shocks where the prediction would be even stronger due to small cell sizes.

[39] Including these observations does not change the results (results available upon request).

[40] I report the results for wage changes in Appendix Table A1.2 for the sake of completeness.

[41] I allow for arbitrary measurement error in the size of wage changes and only require that the sign is remembered correctly. Given the unusual high estimated magnitudes of a large part of non-zero wage changes observed in Figure 1.2, this specification is likely improve data quality significantly. Classical measurement error would speak against this specification, but the graphical evidence shows no evidence of such behavior.

the same level while wages in the control group increase.

Equation (1.13) also includes time fixed effects (δ_t), which capture the effect of general productivity growth in the same time span. These dummies are equivalent to fixed effects for the different IFLS waves and also control for the slightly different questionnaires and interview procedures. By basing my dependent variable on first-differences, I effectively control for district fixed effects and pre-determined individual controls.[42]
The predictions for this specification are the same as for (1.10): if nominal wages were fully downwardly rigid, there should not be any negative wage changes and the estimation should yield $\alpha_2' = 0$. If negative wage changes are predicted by productivity shocks ($\alpha_2' < 0$), this would be a clear sign of at least partial wage flexibility.

Next, I refine the test for different degrees of nominal wage rigidity by considering different shock combinations similar to the wage level analysis. Using the same indicator for non-negative wage changes as dependent variable, equation (1.14) allows to directly test for nominal wage rigidity:

$$\mathbb{1}_{\Delta w_{idt} \geq 0} = \gamma_0' + \gamma_2' None_{d,t-1} Neg_{dt} + \gamma_3' Pos_{d,t-1} Neg_{dt} + \gamma_4' Pos_{d,t-1} None_{dt} + \delta_t + \varepsilon_{idt} \tag{1.14}$$

The predictions for γ_2' are the same as in the wage level equation (1.11). The predictions for γ_3' and γ_4' change in the wage change specification. In the case of fully flexible wages, the model predicts $\gamma_4' < 0$ as wages adjust back downwards following a positive shock. In the case of full downward rigidity, the wages would not adjust back downwards and the prediction would be $\gamma_3' = \gamma_4' = 0$. As a last step, I again refine the analysis by dividing the negative shocks into strong and weak shocks. Equation (1.15) shows the corresponding specification:

$$\mathbb{1}_{\Delta w_{idt} \geq 0} = \xi_0' + \xi_1' Pos_{dt} + \xi_2' NonPos_{d,t-1} Neg^w_{dt} + \xi_3' NonPos_{d,t-1} Neg^s_{dt} + \xi_4' Pos_{d,t-1} Neg^w_{dt} + \xi_5' Pos_{d,t-1} Neg^s_{dt} + \xi_6' Pos_{d,t-1} None_{dt} + \boldsymbol{\delta} X_{idt} + \delta_d + \delta_t + \varepsilon_{idt} \tag{1.15}$$

The prediction for the coefficients are similar to those of equation (1.15): ξ_5' is the coefficient on the shock combination with the highest effective productivity gap and it is therefore

[42] Including district fixed effects would be equivalent to including differential district time trends. Similarly, I do not control for time constant individual characteristics, as this would amount to including differential time trends by each characteristic. Given the small sample size in the IFLS data, I do not have the power to credibly estimate a model with differential time trends.

the most likely to be negative. Also, as for equation (1.14) full rigidity implies $\xi_i' = 0$ for $i \in \{1, 2, 3, 4, 5, 6\}$.

To test the labor supply predictions of Proposition 3 I estimate equation (1.16):

$$\begin{aligned} l_{idt} =& \varphi_0 + \varphi_1 Pos_{dt} + \varphi_2 NonPos_{d,t-1} Neg_{dt} + \varphi_3 Pos_{d,t-1} Neg_{dt} + \varphi_4 Pos_{d,t-1} None_{dt} \\ &+ \delta X_{idt} + \delta_d + \delta_t + \varepsilon_{idt} \end{aligned} \tag{1.16}$$

where l_{idt} is the individual labor supply in agricultural wage work measured as days worked last week. As I aim to analyze movements in aggregate employment I only use the SUSENAS and SAKERNAS data for this estimation. The predictions of Proposition 3 depend on the threshold productivity gap. If the downward rigidities bind, the model predicts a decline in employment relative to the benchmark case. In particular, if $\rho^{HL} > \rho^* > \rho^{ML}$ the implications of Proposition 3 are: $\varphi_3 < \varphi_2 < 0$. Furthermore, the prediction of part three of the proposition implies that when rigidities do not bind, employment should be at the level of the control districts.

1.5 Main Results

The empirical results strongly support my theory that the effective productivity gap determines the existence of downward nominal wage rigidity. Before presenting the estimation results, I briefly discuss my choice of standard errors. For my preferred specification I cluster standard errors at the province year level. This allows for spatial correlation of error terms within a given year, which is appropriate given the spatial correlation of rainfall.[43] Clustering the standard error at the province year level yields 232 clusters for the wage level analysis, but it only yields 40 distinct clusters in the IFLS data. To ensure correct statistical inference using the IFLS data, I use wild-bootstraps for statistical inference. In the following result tables, I distinguish between clustered standard errors and wild-bootstrapped p-values, showing them in parenthesis and square brackets, respectively.

Table 1.2 shows the results of the basic test of whether nominal wages adjust to contemporaneous negative productivity shocks. Column 1 shows the results of the wage level specification. Agricultural wages in districts which experience a positive shock are about 5.3% higher than in

[43] The results are robust to clustering at the district level (Appendix Table A1.3).

Table 1.2: Basic Results

	Source: SUSENAS & SAKERNAS (1998-2007)	Source: Indonesian Family Life Survey (1993-2000)
Dependent variable	log daily wages	$\mathbb{1}_{\Delta w \geq 0}$
	(1)	(2)
Rainfall shock this year		
1 Positive	0.053***	
	(0.0179)	
2 Negative	-0.036	-0.085**
	(0.0242)	[0.026]
Mean of dependent variable	8.95	0.78
Observations	86,605	1,654
R-squared	0.563	0.004
Number of clusters	232	40
District FE	Yes	No
Time FE	Yes	Yes

Notes: Standard errors clustered at province year level displayed in parenthesis. p-values obtained using wild bootstraps at province-year level with 1000 repetitions displayed in square brackets. Sample restricted to agricultural wage workers. Positive shocks are defined as a monsoon onset of at at least 20 days before the district mean. Negative shocks are defined as a monsoon onset of at least 30 days after the local mean. Monsoon onset is defined as reaching cumulative rainfall of 20cm after August first. SAKERNAS & SUSENAS data consists of National Socioeconomic Survey data from 1998 to 2004 and of National labor Force Survey data collected in February 2005 to 2007. The positive shock category is omitted for the IFLS sample due to insufficient sample size. Sample restricted to individuals working in the same job in both years. Dependent variable is a dummy indicating a non-negative wage change. *** p<0.01, ** p<0.05, * p<0.1.

districts without a shock. This effect is significant at the 1% level. This indicates that agricultural wages are responsive to productivity shocks in general. Row 2 shows a decrease in wages by 3.6% in response to a negative shock. However, the effect is not statistically significant.

Column 2 shows that the probability of a decrease in the individual nominal wage in districts with a negative shock increases by 8.5 percentage points. This is equivalent to an increase in the probability of a negative wage change by almost forty percent. This effect is significant at the 5% level.[44,45]

[44] The large difference in R-squared between the cross-sectional and the panel data can be attributed the different measurement error and different nature of wage levels and changes. A large part of the variance of wage levels can be explained by time and district fixed effects. Wage changes on the other hand are very idiosyncratic and depend on unobservable and time specific local and individual characteristics.

[45] If I exclude observations with positive shocks from the wage level analysis the impact of the negative coefficient becomes significant, too (results available upon request).

The results of this basic estimation are a first sign that wages are not fully downwardly rigid. In particular, the observed decline of the probability of wage changes is clear evidence against full nominal wage rigidity. Effects on wage levels are less clear, but the relatively large point estimate could be a sign of heterogeneous downward adjustment.

The results in Table 1.3 shows my preferred specification and distinguishes impact of various magnitudes of the effective productivity gap. Column 1 and 2 show the results using nominal daily wage levels with and without individual controls. Including time-constant individual controls does not change the results significantly and suggests that observable changes in the underlying sample of wage earners do not drive the results.

Similar to the basic specification there is a significant wage increase of 5% to 6% with contemporaneous positive shocks. However, downward adjustment of nominal wages to contemporaneous negative shocks depends on the previous season's rainfall. Row 2 shows the impact of a negative shock following a season with no positive shock. The coefficient on this variable is not statistically different from zero. Row 3 reports the impact of a positive shock which is followed by a negative shock. The estimates imply that wages in districts that experience this shock combination are 7% to 9% lower than wages in control districts. The effect is significant at the 1% level. This result is in line with the theoretical prediction outlined in section 2 and is the main result of this paper.

The estimated impacts of a contemporaneous negative shock which is not preceded by a positive shock which is insignificant (with a t-statistic of less than 1) and has a fairly small point estimate (row 2). This absence of an observable downward adjustment is also in line with the theory of partial nominal wage rigidity. The coefficient on the lagged positive shock in row 4 is not statistically different from zero. This lack of an impact can be explained by consistently high inflation during the sample period, which causes the effective productivity gap to fall below one: During the years 1999 to 2007 the average national inflation was 8.4%, which is higher than the average wage increase following a positive shock.[46] This means that the nominal wage level in control districts would increase to the level in districts exposed to a positive shock and

[46] Excluding 1998 because inflation was around 80% in 1998 due to the one-off effects of the Asian Financial Crisis.

Table 1.3: Main Results

			Source: SUSENAS & SAKERNAS (1998-2007)		Source: Indonesian Family Life Survey (1993-2000)
Dependent variable			log daily wages		$\mathbb{1}_{\Delta w \geq 0}$
			(1)	(2)	(3)
	Shock last year	Shock this year			
1	any	positive	0.051***	0.064***	
			(0.019)	(0.015)	
2	non-positive	negative	-0.021	-0.020	-0.066
			(0.027)	(0.027)	[0.22]
3	positive	negative	-0.090***	-0.070***	-0.096**
			(0.027)	(0.026)	[0.04]
4	positive	none	0.010	0.022	0.056
			(0.022)	(0.019)	[0.28]
Mean of dependent variable			8.95	8.95	0.78
F-test p-value: row 2= row 3			0.01***	0.03**	0.24
Observations			86,605	86,605	1,654
R-squared			0.563	0.617	0.005
Number of clusters			232	232	40
Individuals Controls			No	Yes	-
District FE			Yes	Yes	-
Time FE			Yes	Yes	Yes

Notes: Standard errors clustered at province year level displayed in parenthesis. p-values obtained using wild bootstraps at province-year level with 1000 repetitions displayed in square brackets. Sample restricted to agricultural wage workers. Positive shocks are defined as a monsoon onset of at at least 20 days before the district mean. Negative shocks are defined as a monsoon onset of at least 30 days after the local mean. Monsoon onset is defined as reaching cumulative rainfall of 20cm after August first. SAKERNAS & SUSENAS data consist of National Socioeconomic Survey data from 1998 to 2004 and of National labor Force Survey data collected in February 2005 to 2007. The positive shock category is omitted for the IFLS sample due to insufficient sample size. Sample restricted to individuals working in the same job in both years. Dependent variable is a dummy indicating a non-negative wage change. Individual controls include age, age squared, sex and dummies for living in a rural area and having at least primary education. *** p<0.01, ** p<0.05, * p<0.1.

that it would be impossible to detect persistent nominal wage increases.[47]

I further test whether the coefficient in row 2 and 3 are statistically different from each other using standard F-tests. I can reject the hypothesis that the impact of a negative shock is the same regardless of the preceding shock for the cross-sectional data. For the IFLS data I less

[47]The high average levels inflation could also cause the negative point estimate in row 2, even when nominal wages do not adjust downwards.

powered and cannot reject equality at conventional levels. This leads to the conclusion that past shocks indeed matter for current wage responses to negative rainfall shocks.[48]

These results have two implications for the presence of downward nominal wage rigidity: First, I can reject the hypothesis of full nominal wage rigidity as wages clearly adjust downwards in some cases. Second, I find evidence that downward nominal wage rigidity is present when the effective productivity gap is not too high. Taken together, this is a clear indication of partial nominal wage rigidity.

One major innovation of this analysis relative to the existing literature is that it validates cross-sectional results with results obtained using data on wage changes. This allows to overcome potential problems of an endogenous labor force composition. By focusing on individuals which stayed in the labor force and did not migrate or switch jobs, I can credibly estimate the impact of productivity shocks on individual wages. Column 3 of Table 1.3 shows the results of equation (1.14). The results clearly indicate that negative wage changes are predicted by a negative rainfall shocks which are preceded by a positive shock. The point estimate implies a 9.6 percentage point increase in the probability of having a negative wage change. This is equivalent to an increase of almost 44% of the baseline probability of having a negative wage change and amounts to an average 32% probability of reporting a negative wage change in districts that experience this shock combination. Given the described wage reporting issues and relatively high inflation (between 9% and 12% for the IFLS years), this result supports the wage level result that wages adjust downwards in response to negative shocks which are preceded by positive shocks.

The coefficient on negative shock which are not preceded by a positive shock is also relatively large, but not statistically significant. The difference between the two becomes more pronounced once I include province time trends (see Section 1.5.3). While I cannot reject that they are equal to each other's point estimate, their relative size is in line with the predictions of the theory and the cross-sectional results.[49] Further note that the model predicts a zero impact of the lagged positive shock in wage change specification which is in line with the insignificant impact found in row 4. The results of the wage-change analysis presented above clearly support

[48] The results on the negative shock coefficient holds when I drop observations with a contemporaneous positive shock. In this case the coefficient on the lagged positive shock is positive and significant, too.

[49] I do not directly test for statistical difference of the coefficients, because I use wild bootstraps of the coefficients which to not yield estimates for variance-covariance matrix of the estimated coefficients.

Table 1.4: Impacts by Shock Strength

			Source: SUSENAS & SAKERNAS (1998-2007)		Source: Indonesian Family Life Survey (1993-2000)
Dependent variable			log daily wages		$\mathbb{1}_{\Delta w \geq 0}$
			(1)	(2)	(3)
	Shock last year	Shock this year			
1	any	positive	0.053***	0.065***	
			(0.0192)	(0.0149)	
2	non-positive	weak negative	-0.036	-0.032	-0.077
			(0.030)	(0.030)	[0.17]
3	non-positive	strong negative	0.018	0.001	-0.053
			(0.030)	(0.029)	[0.34]
4	positive	weak negative	0.018	0.017	-0.048
			(0.049)	(0.045)	[0.65]
5	positive	strong negative	-0.12***	-0.094***	-0.102*
			(0.027)	(0.022)	[0.05]
6	positive	none	0.012	0.024	0.056
			(0.022)	(0.019)	[0.31]
Mean of dependent variable			8.95	8.95	0.78
Observations			86,605	86,605	1,654
R-squared			0.563	0.617	0.006
Number of clusters			232	232	40
Individuals Controls			No	Yes	-
District FE			Yes	Yes	-
Time FE			Yes	Yes	Yes

Notes: Standard errors clustered at province year level displayed in parenthesis. p-values obtained using wild bootstraps at province-year level with 1000 repetitions displayed in square brackets. Sample restricted to agricultural wage workers. Positive shocks are defined as a monsoon onset of at at least 20 days before the district mean. Negative shocks are defined as a monsoon onset of at least 30 days after the local mean. Monsoon onset is defined as reaching cumulative rainfall of 20cm after August first. Weak negative shocks are below median shock strength. Strong negative shocks are at or above median shock strength. SAKERNAS & SUSENAS data consists of National Socioeconomic Survey data from 1998 to 2004 and of National labor Force Survey data collected in February 2005 to 2007. The positive shock category is omitted for the IFLS sample due to insufficient sample size. Sample restricted to individuals working in the same job in both years. Dependent variable is a dummy indicating a non-negative wage change. Individual controls include age, age squared, sex and dummies for living in a rural area and having at least primary education. *** p<0.01, ** p<0.05, * p<0.1.

the findings of the wage level analysis and provide evidence that the results are not driven by compositional changes.

To provide further evidence for my proposed mechanism, I compare the impact of strong and weak negative rainfall shocks. Table 1.4 reports the results of these estimations. The results support my theory: In all specifications the only significant negative impact is that of a positive shock being followed by a strong negative shock. This is in line with Proposition 2, because this shock combination has unambiguously the largest productivity gap.

Overall, my empirical findings suggest that the existence of downward nominal wage rigidity depends on observable factors: the evidence supports the idea that wages are more likely to adjust downwards when the cost of maintaining the wage at last period's nominal wage is high. My results contradict full nominal wage rigidity in agricultural labor markets as found by Kaur (2019), but at the same time my findings support the existence of downward nominal wage rigidity when these costs are sufficiently low.

1.5.1 Labor Supply Results

To test further predictions of the model, I estimate the impact of rainfall shocks on labor supply. Table 1.5 shows the results of this analysis using the SAKERNAS and SUSENAS data to estimate equation (1.16). The sample consists of all individuals working in agriculture in a given year in the sample period.[50] The dependent variable is days worked in agricultural wage work during the last week. Row 1 shows a significant positive effect of contemporaneous positive shocks on labor supply of 0.16 worker days or 37% of the mean. This positive impact is in line with the predictions of the model. The absolute size of the effect is in line with findings in the literature (Kaur, 2019), but the relative size is larger.[51] This is probably because I only observe relatively formal agricultural work, which explains the lower average in days worked.

Proposition 3 describes the labor supply predictions of the model in reaction to other shock combinations: When the rigidities bind and wages adjust downwards the labor supply predictions of a negative productivity shock are ambiguous. Hence, the insignificant impact of the negative shock which is preceded by a positive shock is not surprising and does not contradict the model.

[50] This includes self-employed workers in agriculture and is supposed to reflect the whole pool of potential agricultural wage laborers. Using the whole population instead of individuals working in agriculture does not change the results (results available upon request).

[51] My estimate should be more precise than those of Kaur (2019), because my sample is almost double her size.

If nominal wage rigidities bind and wages do not adjust downwards, the model predicts a de-

Table 1.5: labor Supply Results

			Impact on worker-days in agriculture	
			(1)	(2)
	Last year's shock	This year's shock		
1	any	positive	0.164***	0.164***
			(0.049)	(0.049)
2	non-positive	negative	-0.076	-0.076
			(0.068)	(0.068)
3	positive	negative	-0.006	-0.011
			(0.143)	(0.143)
4	positive	none	0.141**	0.139**
			(0.068)	(0.068)
Mean of dependent variable			0.43	0.43
Observations			1,190,023	1,190,023
R-squared			0.079	0.086
Number of clusters			232	232
Individuals Controls			No	Yes
District FE			Yes	Yes
Time FE			Yes	Yes

Notes: Data source is National Socioeconomic Survey from 1998 to 2004 and of National labor Force Survey in February 2005 to 2007. Dependent variable is days worked as agricultural wage workers. Sample restricted to individuals working in agricultural. Individual observations weighted using cross-sectional sample weights. Positive shocks are defined as a monsoon onset of at at least 20 days before the district mean. Negative shocks are defined as a monsoon onset of at least 30 days after the local mean. Monsoon onset is defined as reaching cumulative rainfall of 20cm after August first. Standard errors clustered at province year level displayed in parenthesis. Individual controls include age, age squared, sex and dummies for living in a rural area and having at least primary education. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

crease in labor supply. Given the results presented so far, I would expect this to happen when a non-positive shock is followed by a negative shock. However, the estimates in row 2 show no significant effect. The point estimates have the expected sign, but they also decline once district specific time trends are included (see Section 1.5.3). This contradicts the predictions of my basic model.

In addition, note that the results in Section 1.5 suggest that the wage rigidity does not bind for lagged positive shocks. Proposition 3 implies that in this case a lagged positive shock should

leave current labor supply unaffected. However, I find a significant positive employment effect of lagged positive shocks of 0.14 worker days, or 32% of the mean without a corresponding increase in wages (row 4). This again contradicts the predictions of my model.

My labor supply results suggest that there are factors preventing firms from reducing employment when the wage does not adjust downwards. One such factor could be firing frictions that make the layoffs of workers expensive. The Indonesia Jobs Report of the World Bank (The World Bank, 2010) asserts that employers are legally obliged to make costly severance payments to laid-off of workers. While compliance with this law is relatively low, it could still be part of the explanation of why employment is rigid, especially given that this analysis focuses on relatively formal agricultural workers.[52]

In Appendix Section 1.A.5 I analyze the theoretical implications of including firing frictions in the profit function of the firms in the following form:

$$\pi_j = p\theta f(eL_j) - w_j L_j - \phi(L_{j,t-1} - L_j)\mathbb{1}_{L_{j,t-1}>L_j} \tag{1.17}$$

where $\phi > 0$ is the cost of laying-off one worker and $\mathbb{1}_{L_{j,t-1}>L_j}$ is a dummy indicating a decrease in employment and $L_{j,t-1}$ is last period's employment. This adds a kink to the profit function, which modifies the predictions of the model: When nominal wage rigidities bind, which implies a decrease in labor demand in the basic model, firms now fire fewer workers. This is because the marginal benefit of fired workers is equal to their marginal productivity plus the firing cost ϕ. In Appendix Section 1.A.5 I show that for ϕ large enough the model can be reconciled with the empirical labor supply results.

The fact that positive productivity shocks have persistent employment effects, even in the presence of downwardly rigid nominal wages means that workers bear less of its cost than in otherwise frictionless markets. For example it could imply that firms' bear more of the cost whenever moderate negative productivity shocks occur, thus benefiting existing workers.[53]

[52] I do not take a strong stance of what exactly causes these frictions. Alternative explanations include hiring frictions or implicit contracts between firms and workers.

[53] However, workers might bear the cost through other channels. For example, firms could anticipate the rigid reaction to workers' wages which could lead to reduced hiring rates and lower overall levels of employment.

1.5.2 The Effect of Inflation

The model also has predictions for the effect of inflation, however testing this empirically is difficult due to the limited time span observed. The IFLS sample, in particular, displays little variation in inflation (both spatially and temporarily), so that it is impossible to estimate the impact of inflation. The SUSENAS and SAKERNAS sample displays more variation in inflation, but the results are still noisily estimated.[54] Appendix Table A1.4 shows the estimated impact of inflation, but the results remain largely inconclusive. Only the impact of the non-positive negative shock combination varies significantly with inflation (rows 3 and 4): Wages adjust downwards at low levels of inflation, but they do not adjust at higher levels of inflation. This observation is in line with the prediction that wage cuts are more likely to occur at higher values of effective productivity gap.

The aggregate economic consequences of nominal wage rigidity in the Indonesian context are ambiguous. First, average inflation is relatively high which implies that the rigidity does not bind very often. Second, even when it does bind it only prevents wages from falling when the cost of maintaining the current nominal wage are not too high. And third, observed employment consequences of downward nominal wage rigidity are not very strong. These findings imply that most of the burden of negative productivity shocks is borne by the firms rather than by the workers; the joint presence of both downward nominal wage rigidity and employment frictions seems to work as a partial informal insurance mechanism for workers. Only when the rigidities do not bind and wages adjust downwards is the burden of negative economic shocks shared between firms and workers.

However, the nature of my analysis does not allow for an explicit welfare analysis. Empirical evidence on the micro-foundations of downward nominal wage rigidity and other labor market frictions is needed to draw conclusions about welfare effects.[55]

[54] I average inflation data from major Indonesian cities provided by the Indonesian Statistical Agency on province level to obtain province specific inflation measures. To ensure that inflation is not driven by rainfall shocks, I use average inflation in all other provinces as inflation measure for each province. I exclude the year 1998, because of its exceptional high level of inflation.

[55] Theoretically, the 'theory of the second best' implies that it is at least possible for firing frictions to partially offset the inefficiencies in the labor market adjustment to economic shocks caused by downward nominal wage rigidity (Lipsey, R. G. Lancaster, 1956).

1.5.3 Ruling out Alternative Explanations

Could the observed results be caused by factors other than wage and employment rigidities? First, in-migration following positive shocks could cause the differential impact of contemporaneous negative shocks increasing the total labor supply.[56,57] This could in theory explain the impacts of past positive shocks on the wage reaction to contemporaneous negative shocks. However, this does not explain the insignificant effect of negative shocks which were not preceded by a positive shock. In theory it could be that the effects of negative productivity shocks are too weak to be detectable in the data. However, given the strong reaction of wages to positive shocks it is implausible that negative shocks (which have a stronger effect on productivity than positive shocks) should be too weak to influence wages. Furthermore, when I separate negative shocks by strength the point estimates for strong negative shocks are smaller than those for weak negative shocks (Table 1.4). This combined with the large sample size leave me confident that the null effect is not caused by lack of power.[58] Permanent in-migration also implies downward pressure on wages following a lagged positive shock without a shock in the next period, but again I do not find such an effect on wages. I cannot directly rule out that the inflow of workers offsets positive demand spillovers caused by in-migration (Santangelo (2019) finds such spillover effects in India). However, this is unlikely to be the case in Indonesia as Bazzi (2017) shows that wages decline with migration inflows induced by favorable rainfall.

Second, if rainfall shocks had a permanent impact (for example by filling water depots or moisturizing the soil), this could explain the persistent increase of employment following a positive shock. However, in this case wages would also remain high following a positive shock which is not the case. Similarly, if positive shocks allowed credit-constrained farmers to invest in

[56] By using the IFLS data, I have already ruled out that selective migration affects the sample composition directly.

[57] It is well-known that rainfall induces migration in Indonesia and that this migration occurs with some time lag (Bazzi, 2017). Furthermore, Rosenzweig and Udry (2014) show that the impact of rainfall shocks can be aggravated by migration if shocks are unexpected. This in would in particular be true if rainfall shock were positively serially correlated, as workers would expect good year to follow on another good year, which should increase the expected value of migrating to a given location. However, in Appendix Table A1.5 I show that monsoon onset delay is negatively serially correlated in Indonesia as a whole. This means that if intertemporal optimization takes place, negative shocks should be expected to follow on positive shocks. I workers adapted to this pattern, I would expect the serial correlation to weaken the impact of in-migration. Furthermore, I show that serial correlation disappears once time and district fixed effects are included. This means that past shocks in a given district have no predictive power about future shocks, once the local means and national trends are accounted for.

[58] The more noisily measured result using wage change data is not surprising due to the smaller sample size and described measurement issues.

capital this could explain the persistent increase in employment following a positive shock. For this to happen labor and capital would have to be complements which implies an increase in the marginal product of labor. This would cause the wage to increase following a lagged positive shock. However, this contradicts my empirical findings. Permanent shocks and migration combined could explain the persistent employment effect and the influence of past positive shocks on wage setting. However, it does, again, not explain why wages do not fall significantly with just a negative shock.

Third, it could be that districts are on differential wage growth paths which were not accounted for in the data. It is possible that the different shock combinations spuriously picked up these differential trends. Appendix Table A1.6 shows that the cross-sectional wage rigidity results are robust to including province and district specific linear time trends (columns 1 and 2). The positive employment effect of lagged shocks is not fully robust to including specific time trends, but it is only marginally insignificant when district time trends are included (column 4). Once I include district trends in the panel specification the estimated coefficients even increase further. Clearly, the inclusion of district trends decreases power and increases standard errors which is why these results are noisily measured.[59] Thus, I can rule out that my results are driven by differential trends.

Fourth, the definition of rainfall shocks in absolute terms causes some districts to receive either no shocks or only one kind of shock. If for example district which only ever received a positive shock during the period of 1979 to 2010 had systematically higher wages than districts which received a negative shock, this could lead to spurious estimates of negative effects of negative shocks. A similar argument can be made for district which only receive negative shocks. If districts with just negative shocks had systematically lower wages than districts without negative shocks, this difference could be picked up by the regression coefficient. Districts that do not receive any shock in this period are always in the control group. For this to bias the results, wages would have to move in the opposite direction of wages districts which experience a shock, but given the spatial correlation of monsoon onset this is relatively unlikely. Appendix Figure 1.C5 shows that the spatial distribution of districts without shocks closely follows the timing of monsoon onset. Those districts are clustered in the north west of Indonesia which means that

[59] If I cluster standard errors at the province year level, the results are significant (results available on request).

lack of shock coincides with a different agricultural production cycle. Wages in these district are ex-ante less likely to be affected by monsoon onset as the data collection falls into the second cropping season. Hence, excluding these shocks is similar to testing for the impact by differential agricultural production cycle. Appendix Table A1.7 shows that excluding districts which experienced either no shocks or only one type of shock does not affect point estimates and results are only slightly more nosily measured.

Fifth, measurement error could contaminate the results. If measurement error was classical, my findings would be attenuated towards zero which would strengthen the validity of my finding of partial downward flexibility. The presence of classical measurement error would theoretically weaken the finding of no impact of the negative shock which is not preceded by a positive shock. However, it does not account for the differential adjustment to contemporaneous negative shocks. If wages are systematically misreported, this could in theory bias my results in either direction. However, for this to explain the findings, misreporting of wages would have to be correlated with intertemporal shock patterns (for example underreporting of wages has to be more likely if a negative shock follows on a positive shock). Furthermore, the misreporting would have to occur in both data sources in a similar fashion.

Sixth, further lags of positive shocks could potentially influence the results by defining the wage level in the previous year. Including dummies for lagged positive shocks does not change my results for wage level (columns 1 and 3 of Appendix Table A1.8). The IFLS results are a bit more noisily measured, and the coefficient on the negative shock which is preceded by positive shock turns marginally insignificant (column 5 of Appendix Table A1.8), but the general interpretation of the results is still similar. The results are also robust to defining shocks by using a symmetric cutoff of twenty days for positive and negative shocks (even columns of Appendix Table A1.8). The cross-sectional results do not change at all, while the point estimates of both contemporaneous negative shocks increase further.

Appendix Table A1.9 provides further evidence that my results are not driven by selecting a specific sample or time period. Column 1 shows that the results are robust to including wages of casual workers for the years 2000 and 2005 to 2007 when they are reported. Column 2 shows the results excluding observations in February 1998 when the consequences of the Asian

Financial Crisis coincided with a major drought which caused a spike in inflation and had severe economic consequences in Indonesia (Frankenberg and Smith, 2003; Smith et al., 2002). The results confirm that my findings are not driven by one-off effects of the crisis. The results also hold when I restrict the cross-sectional sample to districts found in the IFLS data (column 3). Column 4 shows the results excluding the island of Sumatra from the regression. The idea behind this is that monsoon onset is relatively early for most of the island which could influence the results. The estimated impact of the positive negative shock combination declines, but it still remains significant. Column 5 shows the results of the main estimation only using years in which the SUSENAS are the data source.[60] The positive-negative coefficient decreases slightly and becomes marginally insignificant (with a p-value of 0.12) and the effect of the lagged positive shock turns marginally significant. Using hourly instead of daily wages show similar results (column 6). The coefficient on the negative shock following a positive shock this time has a p-value of 0.11, and the lagged positive coefficient becomes significant. The fact that the coefficient on lagged positive shocks is significant in some specifications can be interpreted as further evidence supporting the idea of downwardly rigid nominal wages.

1.6 Conclusion

In this paper I test for the presence of downward nominal wage rigidity in Indonesian agricultural labor markets. I provided evidence from two separate data sources that nominal wages exhibit differential adjustment patterns to negative labor demand shocks: when negative shocks are not preceded by a positive shock they do not adjust downwards. However, when negative shocks follow on a positive labor demand shock, nominal wages adjust downwards to the reduction in productivity. Additionally, I show that this effect is driven by differences in the cost of maintaining last period's nominal wage. Splitting negative productivity shocks by their severity, I provide further evidence for this mechanism: wages only adjust downwards when strong negative shocks are preceded by positive shocks, but not when weak negative shocks follow positive shocks. I further validate the results of a wage level analysis with an equivalent analysis using individual level wage change data. This enabled me to credibly deal with potential threats to identification such as selective migration or aggregation bias.

Presenting a model of agricultural labor markets, I formalize the idea that the cost of downward adjustment of wages determines whether nominal wages adjust to negative productivity

[60] Around 90% of my sample consists of SUSENAS data, because the SAKERNAS surveys were fairly small.

shocks. I show that the cost of paying workers' less than their reference wage by decreasing the nominal wage can be outweighed by the cost of keeping nominal wages constant, which leads firms to adjust their nominal wages downwards if it becomes to costly to not do so. This limits the potential economic cost of nominal wage rigidities to the economy.

In addition to analyzing wage-setting in response to productivity shocks, this paper also analyzes the labor supply response to labor productivity shocks. I found persistent employment effects following positive productivity shocks and my results show no observable negative employment effect of downward nominal wage rigidity. This is suggestive empirical evidence that, in the presence of other frictions, downward nominal wage rigidities can redistribute surplus from firms to workers. This paper cannot draw any firm conclusions about the distributional consequences of the nominal wage rigidities. I think that exploring the distributional consequence of labor market frictions and, in particular, downward nominal wage rigidity is a fruitful avenue of future research.

Appendices of Chapter 1

Description of Appendix to Chapter 1

Section 1.A lays out the proofs for the propositions described in the main text. Section 1.B contains additional results tabls. Table A1.1 provides summary statistics of all relevant variables. Table A1.2 contains the main results on wage differences using the IFLS data. Table A1.3 clusters the standard error at the district level. Table A1.4 displays interaction effects with inflation. Table A1.5 tests for serial correlation in rainfall. Table A1.6 accounts for district and province specific time trends. Table A1.7 shows robustness of the results to dropping districts without rainfall shocks. Table A1.8 show robustness to different definition of rainfall shocks. Table A1.9 shows robustness to using different samples of agricultural workers. Section 1.C displays additional figures for Chapter 1. Figure 1.C1 show which Indonesian districts are included in the sample. Figure 1.C2 shows the non-parametric estimates of the impact of rainfall on rice yields in Indonesia. Figure 1.C3 shows the spatial distribution of average monsoon onset. Figure 1.C4 shows the spatial distribution of average monsoon onset for the IFLS sample. Finally, Figure 1.C5 shows the spatial distribution of districts with different sets of rainfall shocks.

1.A Mathematical Appendix

1.A.1 Benchmark: Solution without Reference Dependence

Poposition 4 restates Kaur (2019)'s benchmark solution of the model without refence dependence to provide context for my theoretical results.

Proposition 4. Market clearing without reference dependence
For a constant level of worker effort $e \in \{\lambda, 1\}$*, equations* (1.5) *and* (1.6) *determine labor supply and the wage in the unique pure strategy Nash Equilibrium for all values of* θ*.*

Proof. See Kaur (2019) or Appendix section 1.A.1. ■

Proposition 4 states that there is a unique pure strategy Nash Equilibrium in the benchmark case of no reference dependence ($e = 1$) and the case of permanently low effort ($e = \lambda$). It further states that this equilibrium is fully described by the labor market clearing condition (1.6) and the firm's first order condition (1.5).

Proof of Proposition 4: Market Clearing without Reference Dependence

This proof is fully based on Kaur (2019) with the simple extension to a constant $e \in \{\lambda, 1\}$. To see that this holds, simply note that the labor supply equation (1.3) does not change with e and the first order condition is monotonous in e. Hence, all of the following holds for $e \in \{\lambda, 1\}$.

I first show that deviating from w^* is not profitable if all firms set $w = w^*$ and choose L^* according to the first order condition (1.5):

1. Suppose firm j deviates to $w_j = w^* - \epsilon$. This deviation will decrease the available labor supply to

$$L^{Avail} = \max\left\{\frac{1}{\bar{\psi}}u\left(\frac{w^* - \epsilon}{p}\right) - (J-1)L^*, 0\right\} \tag{1.18}$$

$$= \max\left\{\frac{1}{\bar{\psi}}u\left(\frac{w^* - \epsilon}{p}\right) - \frac{J-1}{J\bar{\psi}}u\left(\frac{w^*}{p}\right), 0\right\} \tag{1.19}$$

   The first order condition (1.5) and the above equation imply that at $w_j = w^* - \epsilon$: $L^{Avail} < L^* < L_j^{FOC}$. The deviation is profitable iff $\pi_j(e, w_j, L_j^{Avail}) > \pi^*(e, w^*, L^*)$.

   (a) If the available labor supply is zero ($L_j^{Avail} = 0$), then trivially profits are zero ($\pi_j(e, w_j, L^{Avail}) = 0$), which means that the profits from keeping the wage at w^* are weakly larger.

(b) If $L_j^{Avail} > 0$ the deviation will not be profitable if the following is greater than zero:

$$\pi^*(e, w^*, L^*) - \pi_j(e, w_j, L^{Avail}) = p\theta\left[f(eL^*) - f(eL^{Avail})\right] - w^*L^* + (w^* - \epsilon)L_j^{Avail} \tag{1.20}$$

$$= p\theta\left[f(eL^*) - f(eL^{Avail})\right] - w^*L^* \tag{1.21}$$

$$+ (w^* - \epsilon)\left(\frac{1}{\bar{\psi}}u\left(\frac{w^* - \epsilon}{p}\right) - (J-1)L^*\right) \tag{1.22}$$

$$= p\theta\left[f(eL^*) - f(eL^{Avail})\right] - \frac{\epsilon}{\bar{\psi}J}u\left(\frac{w^*}{p}\right) \tag{1.23}$$

$$+ \underbrace{\frac{w^* - \varepsilon}{\bar{\psi}J}\left(Ju\left(\frac{w^* - \epsilon}{p}\right) - u\left(\frac{w^*}{p}\right)\right)}_{<0, \text{ by (1.19)}} \tag{1.24}$$

$$> p\theta\left[f(eL^*) - f(eL^{Avail})\right] - \frac{\epsilon}{\bar{\psi}J}u\left(\frac{w^*}{p}\right) \tag{1.25}$$

$$= P(J) - \frac{\epsilon}{\bar{\psi}J}u\left(\frac{w^*}{p}\right) \tag{1.26}$$

Now note that:

$$\frac{\partial P(J)}{\partial J} = \frac{1}{J^2\bar{\psi}}u\left(\frac{w^*}{p}\right)\theta e\left[f'(eL^*) - f'(eL^{Avail})\right] > 0 \tag{1.27}$$

by concavity of $f(\cdot)$. Now define implicitly $\hat{J}$ by:

$$P(1) = \frac{\epsilon}{\bar{\psi}\hat{J}}u\left(\frac{w^*}{p}\right) \tag{1.28}$$

It remains to show that $P(J) - \frac{\epsilon}{\bar{\psi}J}u\left(\frac{w^*}{p}\right) > 0$ for any sufficiently large J. To see that this holds for any $J > \hat{J}$, note that for any positive number X:

$$P(\hat{J} + X) > P(\hat{J}) \qquad \text{by (1.27)} \tag{1.29}$$

$$> P(1) \qquad \text{by (1.27)} \tag{1.30}$$

$$= \frac{\epsilon}{\bar{\psi}\hat{J}}u\left(\frac{w^*}{p}\right) \tag{1.31}$$

$$> \frac{\epsilon}{\bar{\psi}(\hat{J} + X)}u\left(\frac{w^*}{p}\right) \tag{1.32}$$

Hence, this deviation is not profitable for J sufficiently large.

2. Suppose firm j deviates to $w_j > w^*$: From the first order condition (1.7) follows that labor demand decreases to $L_j^{FOC} < L^*$. However, the firm could have hired L_j^{FOC} at lower wages w^* without the raise, so that this deviation cannot be profitable.

It remains to proof that no other wage can be a pure Strategy Nash Equilibrium. I prove this by showing that the market clearing condition has to hold.

1. Consider the case of excess labor supply: $JL^* < \frac{1}{\bar{\psi}J} u\left(\frac{w^*}{p}\right)$. If firm j sets $w = w^* - \epsilon$ it can still hire the same amount of labor with a strictly lower wage bill. To proof this denote the slack in the labor market as: $\delta = JL^* - \frac{1}{\bar{\psi}J} u\left(\frac{w^*}{p}\right)$. The firms can still hire L^* workers iff:

$$L^* \leq L^{Avail} \tag{1.33}$$

$$\Rightarrow L^* \leq \frac{1}{\bar{\psi}} u\left(\frac{w^* - \epsilon}{p}\right) - \frac{J-1}{J\bar{\psi}} u\left(\frac{w^*}{p}\right) \tag{1.34}$$

$$\Rightarrow \frac{1}{\bar{\psi}J} u\left(\frac{w^*}{p}\right) - \delta \leq \frac{1}{\bar{\psi}} u\left(\frac{w^* - \epsilon}{p}\right) \tag{1.35}$$

   Hence, there exists a ϵ such that an upward deviation is profitable. This implies that there cannot be excess labor supply.

2. Now consider the case of excess labor demand: $JL^* > \frac{1}{\bar{\psi}J} u\left(\frac{w^*}{p}\right)$. This means that each firm hires strictly less labor than what would be optimal by the first order condition (1.5). If firm j raise its wage minimally to $w_j = w^* + \epsilon$, the allocation mechanism implies that it will be able to satisfy its labor demand and increase profits. To see this denote, $L_j^{FOC}(w_j)$ as the labor demand at w_j as determined by the first order condition (1.5). The deviation will be profitable iff:

$$p\theta f(L_j^{FOC}(w^* + \epsilon)) - (w^* + \epsilon) L_j^{FOC} > p\theta f\left(\frac{1}{\bar{\psi}J} u\left(\frac{w^*}{p}\right)\right) - w^* \frac{1}{\bar{\psi}J} u\left(\frac{w^*}{p}\right) \tag{1.36}$$

   For infinitesimally small wage increases:

$$\lim_{\epsilon \to 0} p\theta f(L_j^{FOC}(w^* + \epsilon)) - (w^* + \epsilon) L_j^{FOC} \tag{1.37}$$

$$= p\theta f(L_j^{FOC}(w^*)) - w^* L_j^{FOC} \qquad \text{by continuity of } f(\cdot) \tag{1.38}$$

$$> p\theta f\left(\frac{1}{\bar{\psi}J} u\left(\frac{w^*}{p}\right)\right) - w^* \frac{1}{\bar{\psi}J} u\left(\frac{w^*}{p}\right) \qquad \text{as } L_j^{FOC} \text{maximizes profits at } w^* \tag{1.39}$$

This implies that there cannot be excess labor demand. ■

1.A.2 Proof of Proposition 1: Asymmetric Adjustment of Wages to Shocks

Proof of part 1 of the proposition:

This proof uses arguments similar to Kaur (2019) but extends her results to include the possibility of wage adjustment to negative shocks. Define $\theta'_R > 0$ as the level of productivity at which the market clearing profit at low effort $e = \lambda$ is equal to the profit at the reference wage:

$$\pi^*(\lambda, \theta'_R, w^*, L^*(\theta'_R)) = \bar{\pi}(1, \theta'_R, \bar{w}_{t-1}, \bar{L}(\theta'_R)) \tag{1.40}$$

First I show that θ'_R exists for λ sufficiently close to one by using a mean value theorem. Note that if all firms set the same wage for an arbitrary $0 < \theta' < \theta_R$:

$$\pi^*(1, \theta', w^*, L^*(\theta') > \bar{\pi}(1, \theta', \bar{w}, \bar{L}(\theta')), \quad \text{as } \pi^*(1, \theta', w^*, L^*(\theta')) \text{ and } \theta' < \theta_R \tag{1.41}$$

$$\text{profits at the optimal wage must be higher} \tag{1.42}$$

$$\text{than profits at the higher reference wage.} \tag{1.43}$$

$$\tag{1.44}$$

Now if workers lower their effort because $w^* < \bar{w}_{t-1}$, firms' profits decrease. The magnitude of this loss depends on the size of λ. Note that:

$$\lim_{\lambda \to 1} \pi^*(\lambda, \theta', w^*(\lambda), L^*(\lambda)) = \pi^*(1, \theta', w^*(1), L^*(1)) > \bar{\pi}(1, \theta', \bar{w}, \bar{L}(\theta')) \tag{1.45}$$

and

$$\lim_{\lambda \to 0} \pi(\lambda, \theta', w^*(\lambda), L^*(\lambda)) = \pi(0, \theta', 0, 0) = 0 \tag{1.46}$$

Hence, by continuity of the production and, therefore, continuity of the profit function there is a value $\lambda^*(\theta') < 1$ such that $\pi^*(\lambda^*(\theta'), \theta', w^*, L^*(\theta')) = \bar{\pi}(1, \theta', \bar{w}_{t-1}, \bar{L}(\theta'))$. Further note that the profit function is monotone in λ:

$$\frac{\partial \pi(\lambda, \theta', w^*, L^*)}{\partial \lambda} = p\theta' \lambda f'(\lambda L^*) > 0 \tag{1.47}$$

Hence, there is exactly one $\lambda^*(\theta')$. Further note that by implicit differentiation of (1.40)

$$\frac{\partial \lambda^*}{\partial \theta'} = \frac{\frac{\partial \bar{\pi}}{\partial \theta} - \frac{\partial \pi^*}{\partial \theta}}{\frac{\partial \pi(\lambda^*, \theta', w^*, L^*)}{\partial \lambda^*}} \tag{1.48}$$

$$= \frac{pf(\bar{L}) - pf(\lambda L^*)}{p\theta' \lambda f'(\lambda L^*)} \tag{1.49}$$

$$= \frac{pf(\bar{L}) - pf(\lambda \frac{u(w^*/p)}{J\bar{\psi}})}{p\theta' \lambda f'(\lambda L^*)} \tag{1.50}$$

$$= \frac{T(J)}{p\theta' \lambda f'(\lambda L^*)} \tag{1.51}$$

The sign of this derivative is determined by $T(J)$ as the the denominator is always greater zero by monotonicity of the production function. Further note that $\frac{u(w^*/p)}{J\bar{\psi}})$ gets arbitrary small for large but finite values of J. Hence, $T(J) > 0$ for J large enough (a more formal version can be found in the proof of Proposition 4). As $\lambda^*(\theta')$ is continuous, this result jointly with $\lim_{\theta' \to \theta_R} \lambda^* = 1$ implies that for every λ close enough to one there exists a $0 < \theta'_R < \theta_R$ such that equation (1.40) holds.

I prove the first part of the proposition by showing that for $0 < \theta < \theta'_R$ it is not profitable to deviate from if all firms set $w = w^*(\lambda, \theta)$:

1. Suppose firm j deviates to $\bar{w}_{t-1} > w_j > w^*(\lambda, \theta)$. From the first order condition (1.7) is follows that labor demand decreases to $L_j^{FOC} < L^*$. However, the firm could have hired L_j^{FOC} at lower wages before, so that this deviation cannot be profitable.

2. Suppose a firm deviates to $w = w^*(\lambda, \theta) - \epsilon$. As effort is constant at $e = \lambda$ for $w = w^*(\lambda, \theta) - \epsilon$, Proposition 4 implies that this deviation is not profitable.

3. Suppose a firm deviates to $w = \bar{w}_{t-1}$. The firm would make a profit of $\bar{\pi} = \pi(1, \theta, \bar{w}_{t-1}, \bar{L}(\bar{w}_{t-1}))$.[61] This deviation is profitable if $\bar{\pi} \geq \pi^* = \pi^*(\lambda, \theta, w^*, L^*)$. By definition at θ'_R: $\bar{\pi} = \pi^*$. I proceed by showing that the profit difference $G(\theta) = \bar{\pi} - \pi^*(\lambda)$ is increasing for $\theta \in (0, \theta_R)$. As $G(\theta'_R) = 0$ this implies that for $\theta < \theta'_R$ an upward deviation is not profitable.

   Consider the total derivative of the profit functions with respect to productivity:

$$\frac{d\bar{\pi}}{d\theta} = pf(\bar{L}) - \underbrace{p\theta f'(\bar{L})\bar{L}' - \bar{w}_{t-1}\bar{L}'}_{=0, \text{ by FOC}} = pf(\bar{L}) \tag{1.52}$$

[61] The firm gets to hire labor first, so that it will be able to satisfy its labor demand.

and

$$\frac{d\pi^*}{d\theta} = \frac{\partial \pi^*}{\partial \theta} + \frac{\partial \pi^*}{\partial L^*}\frac{dL^*}{d\theta} + \frac{\partial \pi^*}{\partial w^*}\frac{dw^*}{d\theta} \tag{1.53}$$

$$= pf(\lambda L^*) - \underbrace{p\theta\lambda f'(\lambda L^*)\frac{dL^*}{d\theta} - w^*\frac{dL^*}{d\theta}}_{=0, \text{ by FOC}} - L^*\frac{dw^*}{d\theta} \tag{1.54}$$

$$= pf(\lambda L^*) - L^*\frac{-p\lambda f'(L^*)}{p\theta\lambda^2 f''(\lambda L^*)\frac{u'(w^*/p)}{\bar{\psi}Jp} - 1} \qquad \text{by implicit function theorem} \tag{1.55}$$

and equations 1.3 and 1.7. (1.56)

$$= pf(\lambda L^*) + L^* w^* \theta^{-1} \underbrace{\frac{1}{(p\theta\lambda^2 f''(\lambda L^*)\frac{u'(w^*/p)}{\bar{\psi}Jp} - 1)}}_{\equiv\chi<0, \text{ by concavity of } f(\cdot) \text{ and } u'(\cdot)>0.} \tag{1.57}$$

This yields:

$$\frac{dG(\theta)}{d\theta} = \frac{d\bar{\pi}}{d\theta} - \frac{d\pi^*}{d\theta} \tag{1.58}$$

$$= pf'(\bar{L}) - pf'(\lambda L^*) - L^* w^* \theta^{-1}\chi \tag{1.59}$$

$$> pf'(\bar{L}) - pf'(\lambda L^*), \qquad \text{as } \chi < 0 \tag{1.60}$$

$$= pf'(\bar{L}) - pf'(\lambda\frac{u(w^*/p)}{J\bar{\psi}}) \tag{1.61}$$

$$= F(J) \tag{1.62}$$

Further note that:

$$\frac{\partial F(J)}{\partial J} = -pf'(\lambda L^*)\frac{-\lambda u(w^*/p)}{J^2\bar{\psi}} > 0 \tag{1.63}$$

by concavity of $f(\cdot)$.

The deviation will not be profitable for any J such that $F(J) > 0$. Note that $F(J)$ is increasing in θ_t (as $\bar{L}$ is increasing in θ). Hence, there is a $\hat{\theta}$ such that $\theta > \hat{\theta} \Rightarrow F(J) > 0$. In this case $F(J) > 0$ holds trivially. For each θ such that $\theta > \hat{\theta}$ and $0 < \theta < \theta'_R$ there exists a $\tilde{J}$ such that $F(\tilde{J}) = 0$. It suffices to show that $F(J) > 0$ holds for the largest $J > \tilde{J}' = \max_{\theta \in (0,\theta'_R)} \tilde{J}$. The following shows that this holds for any $J > \tilde{J}'$ and a positive number X:

$$F(\tilde{J}' + X) > F(\tilde{J}'), \qquad \text{as } \frac{\partial F(J)}{\partial J} > 0 \tag{1.64}$$

$$= 0, \qquad \text{as } F(\tilde{J}) = 0 \tag{1.65}$$

Hence, for J sufficiently large, the difference in profits is increasing in the productivity space. Note that this also implies:

$$\pi^*(w^*, \theta) < \bar{\pi}(\bar{w}_{t-1}, \theta) \text{ for } \theta \in (\theta'_R, \theta_R) \tag{1.66}$$

.

4. Suppose a firm deviates to $w > \bar{w}_{t-1}$. From the first order condition (1.7) follows that $L_j^{FOC} < \bar{L}$. However, the firm could have hired $\bar{L}$ workers at lower wages before, so that this deviation cannot be more profitable than deviating to $\bar{w}_{t-1}$. The above shows that deviating to $\bar{w}_{t-1}$ is not profitable.

To prove the second part of the statement, I first show that the reference wage is the only other candidate pure strategy Nash Equilibrium. The following proves that regardless of where firms set their wages, it always profitable to deviate to either $\bar{w}_{t-1}$ or to $w^*(\lambda)$:

1. Suppose all firms set $w = w' > \bar{w}_{t-1}$. At this wage there is an incentive to lower wage by Proposition 4. If workers did not decrease effort, the wage would fall to $w^*(1)$. However, as for all $w = w' > \bar{w}_{t-1}$ lowering the wage at a constant effort is profitable, so that wages will be cut to at least $\bar{w}_{t-1} > w^*(1)$.

2. Suppose all firms set $w = w' \in (w^*(\lambda), \bar{w}_{t-1})$. Again, there will be incentive to lower the wage by Proposition 4. This time the wage will be decreased to $w = w^*(\lambda)$ as workers do not change their effort for this wage range.

3. Suppose all firms set $w = w'' < w^*(\lambda)$. At this point there is excess labor demand and by Proposition 4 there are incentives to raise wages until the labor market clears . As the effort is constant for $w = w' \in (w^*(\lambda), \bar{w}_{t-1})$ it is profitable to increase the wage to $w = w^*(\lambda)$.

Hence, $\bar{w}_{t-1}$ is the only other candidate pure strategy Nash Equilibrium. Whether it is a Nash Equilibrium depends on the shape of the profit function. It remains to show that the sum of profit at $w^*(\lambda)$ is greater than at the reference wage:

$$\int_0^1 \pi_j^*(\lambda, w^*)dj = \pi^*(\lambda, w^*) \tag{1.67}$$

$$> \bar{\pi} \qquad \text{as } \theta < \theta'_R \tag{1.68}$$

$$= \int_0^1 \bar{\pi}_j dj \tag{1.69}$$

■

Proof of part 2 of Proposition 1:

The result of this proof is similar to Kaur (2019), but I use a different argument to demonstrate that a downward deviation is not profitable. I first show that $w = \bar{w}_{t-1}$ is as pure strategy Nash Equilibrium by demonstrating that it is not profitable to deviate from $\bar{w}_{t-1}$, as long as J is sufficiently large. The argument is similar to part one of the proof.

1. Suppose firm j sets $w_j > \bar{w}_{t-1}$. Them it follows from the first order condition (1.7) that $L_j^{FOC} < \bar{L}$. However, the firm could have hired $\bar{L}$ workers at lower wages before, so that this deviation cannot be profitable.

2. Suppose that that firm j sets $w_j < \bar{w}_{t-1}$. Recall equation (1.19) for the available labor supply at wage w_j. In this case available labor supply can be positive, but it is always going to be smaller than the labor supply if all firms set $w = w_j$:[62]

$$L^*(w_j) - L_j^{avail} = \frac{u(\frac{w_j}{p})}{J\bar{\psi}} - \frac{u(\frac{w_j}{p})}{\bar{\psi}} + (J-1)\bar{L} \tag{1.70}$$

$$= \frac{u(\frac{w_j}{p})}{J\bar{\psi}}(1-J) + (J-1)\bar{L} \tag{1.71}$$

$$= (J-1)(\bar{L} - \frac{u(\frac{w_j}{p})}{J\bar{\psi}}) \tag{1.72}$$

$$> 0, \qquad \text{For } w_j < \bar{w}_{t-1} \text{ by equation (1.6)} \tag{1.73}$$

$$\tag{1.74}$$

Now consider profits at w_j:

$$\pi(w_j) = p\theta f(\lambda L_j^{avail}) - w_j L_j^{avail} \tag{1.75}$$

$$< p\theta f(\lambda L^*(\lambda, w_j)) - w_j L^*(\lambda, w_j), \quad \text{by equation (1.74)} \tag{1.76}$$

$$\leq p\theta f(\lambda L^*(\lambda, w^*)) - w^* L^*(\lambda, w^*), \quad \text{as shown in the proof of Proposition 4} \tag{1.77}$$

$$= \pi^*(w^*, \theta) \tag{1.78}$$

$$< \bar{\pi}(\bar{w}_{t-1}, \theta), \qquad \text{as } \pi^*(w^*, \theta) < \bar{\pi}(\bar{w}_{t-1}, \theta) \text{ for } \theta \in (\theta_R', \theta_R) \text{ by (1.66)} \tag{1.79}$$

$$\tag{1.80}$$

[62] I showed the case for $L_j^{avail} = 0$ in the proof of part one of the proposition.

Hence, it can never be profitable to deviate from $\bar{w}_{t-1}$ for $\theta \in (\theta'_R, \theta_R)$.
It remains to show that deviating to $\bar{w}_{t-1}$ is always profitable for $\bar{w}_{t-1}$ to be the unique pure Strategy Nash Equilibrium:

1. Suppose all firms set $w = w' > \bar{w}_{t-1}$. Then there will be excess labor supply: $JL' < \frac{1}{\bar{\psi}} u\left(\frac{w}{p}\right)$. For $w' > \bar{w}_{t-1}$ this is because labor demand decreases in w, labor supply increases in w, and there is excess labor supply at $\bar{w}_{t-1}$. As shown in the proof of Proposition 4, there are incentives to decrease the wage to $\bar{w}_{t-1}$, at which the discontinuity kicks in and no deviation is profitable.

2. Suppose all firms set $w = w' \in (w^*(\lambda), \bar{w}_{t-1})$. Again, there will be excess labor supply: $JL(\lambda, w') < \frac{1}{\bar{\psi}} u\left(\frac{w}{p}\right)$. As shown in the proof of proposition one, there will incentives to decrease the wage to $w^*(\lambda)$.

3. Suppose all firms set $w = w'' < w^*(\lambda)$. At this point there is excess labor demand and there are incentives to raise wages until the labor market clears at $w^*(\lambda)$ (by Proposition 4).

4. In a last step I show that it is profitable to deviate from $w^*(\lambda)$ to $\bar{w}_{t-1}$. The definition of θ'_R and equation (1.66) imply for $\theta \in (\theta'_R < \theta < \theta_R)$:

$$\pi^*(\lambda, w^*) < \bar{\pi}(\bar{w}_{t-1}) \tag{1.81}$$

   Thus, a deviation from $w^*(\lambda)$ to $\bar{w}_{t-1}$ is profitable and $w = \bar{w}_{t-1}$ is the unique pure strategy Nash Equilibrium.

■

Proof of part 3 of Proposition 1

Note that the benchmark wage is feasible for this range of productivities. Further, note that I showed in part 1 of this proof that profits decrease in workers effort. As workers only ever decreases the effort, profits cannot increase for any $w < w^*(1)$. By Proposition 4 the wage determined by (1.5) and (1.6) is the only pure strategy Nash Equilibrium for this range of productivities. ■

1.A.3 Proof of Proposition 2: The Impact of Past Shocks

For the first part of the proposition, recall that θ_R is defined by $p\theta_R f'(L^*(\bar{w}_{t-1}, p)) = \bar{w}_{t-1}$. Further note that $\theta_{t-1} f'(L^*(\bar{w}_{t-1}, 1)) = \bar{w}_{t-1}$. Combining the definition yields:

$$\theta_{t-1} f'(L*(\bar{w}_{t-1})) = p\theta_R f'(\bar{L}) \;\Leftrightarrow \tag{1.82}$$

$$\theta_{t-1} p^{-1} = \theta_R \frac{f'(L^*(\bar{w}_{t-1}, p))}{f'(L^*(\bar{w}_{t-1}, 1))} \tag{1.83}$$

Implicit differentiation of the first order condition (1.7) and the labor market clearing condition (1.6) yields:

$$\frac{\partial L^*(\bar{w}_{t-1}, p)}{\partial p} = -\frac{0}{J - f''(L^*)\theta u'(\bar{w}_{t-1}/p)\bar{\psi}^{-1}} = 0 \tag{1.84}$$

This implies that:

$$\theta_{t-1} p^{-1} = \theta_R \tag{1.85}$$

$$\rho = \frac{\theta_R}{\theta_t} \qquad\qquad | \cdot \theta_t^{-1} \tag{1.86}$$

Hence, $\rho \geq 1$ implies $\theta_R \geq \theta_t$ and vice versa.

Now define $\rho^* = \frac{\theta_R}{\theta'_R}$. Combining this with equation (1.86) yields

$$1 < \rho < \rho^* \tag{1.87}$$

$$1 < \frac{\theta_R}{\theta_t} < \frac{\theta_R}{\theta'_R} \;\Leftrightarrow \tag{1.88}$$

$$\theta'_R < \theta_t < \theta_R \tag{1.89}$$

and

$$\rho > \rho^* \tag{1.90}$$

$$\frac{\theta_R}{\theta_t} > \frac{\theta_R}{\theta'_R} \;\Leftrightarrow \tag{1.91}$$

$$\theta'_R > \theta_t \tag{1.92}$$

For the second part, define $\theta_{t|\rho=\rho^*} = \frac{\theta'_R \theta_{t-1}}{\theta_R} p^{-1}$ as the productivity below which wages adjust downwards. Now note that $\frac{d\theta_{t|\rho=\rho^*}}{d\theta_{t-1}} = \frac{\theta'_R}{\theta_R} p^{-1} > 0$. Hence, there is wider range of $\theta_t < \theta'_R$ for which wages adjust downwards. ■

1.A.4 Proof of Proposition 3: The Effect on labor Supply

I prove each part of the proposition in turn:

1. To proof the first part note that 1 and 2 imply that $w = \bar{w}_{t-1} > w^*(1, \theta)$ is the unique pure Strategy Nash Equilibrium. This means that $L^e = \bar{L}(1, \bar{w}_{t-1})$ is determined by the first order condition (1.5). As $\bar{w}_{t-1} > w^*(1, \theta)$ together with the first order conditions imply that: $L^e = \bar{L}(1, \bar{w}_{t-1}) < L^*(1, w^*(1))$.

2. Proposition 2 implies that in the case the outcomes correspond to the benchmark case. This implies $L^e = L^*(1, w^*(1))$

3. For the third part note that Proposition 1 and 2 imply that for $\rho \geq \rho^*$ wages will adjust to $w^*(\lambda)$. with L^e defined by: $w^*(\lambda) = p\theta\lambda f'(\lambda L^e)$. Employed labor in the benchmark case is defined by: $w^*(1) = p\theta \ f'(L^*(1))$. Without further assumptions on the shape of the production function the relative size of employed labor is unclear. Kaur (2019) assumes $\lambda f'(\lambda L) < f'(L)$ which would implies a decrease of labor supply relative to the benchmark case.

1.A.5 Extension of the Model with Firing Frictions

I introduce firing frictions in the profit function of firm j as:

$$\pi_j = p\theta f(eL_j) - w_j L_j - \phi(L_{j,t-1} - L_j)\mathbb{1}_{L_{j,t-1} > L_j} \tag{1.93}$$

where $L_{j,t-1}$ is last periods employment level of firm j. I assume that employment is set jointly with $\bar{w}_{t-1}$ according to the benchmark case.[63] $\mathbb{1}_{L_{j,t-1} > L_j}$ is an indicator function which is one when employment this period is lower than employment last period and ϕ is the unit cost of laying workers off. This extension introduces a discontinuity into the first order condition of firms at constant effort e:

$$w = \begin{cases} p\theta e f'(eL) + \phi, \text{ if } L_{j,t-1} > L_j \\ p\theta e f'(eL), \text{ if } L_{j,t-1} < L_j \end{cases} \tag{1.94}$$

$\theta_t > \theta_{t-1}$ implies that for $L_{j,t-1} > L_j$ more workers are going to be hired than in the model with firing frictions if effort remains constant. I proceed in two steps:

[63] In a dynamic version of this model, firms would anticipate the potential costs of lay-off as higher less workers. However, this would not change the qualitative message of the extension.

First, I consider the case in which the wage rigidity does not bind but the employment restriction binds. This is the case if $p\theta_t > \theta_{t-1} > \theta_t$. The wage rigidity does not bind because $\rho < 1$. To see that the employment restriction binds, note the definition of past labor supply:

$$L_{t-1} = \frac{1}{\bar{\psi} J} u(\theta_{t-1} f'(L_{t-1})) \tag{1.95}$$

Further note the definition of labor supply in the benchmark case:

$$L^* = \frac{1}{\bar{\psi} J} u(\theta_t f'(L^*)) \tag{1.96}$$

This implies that: $L^* < L_{t-1}$ and, thus, that the employment friction binds. This changes the first labor supply in the current period to:

$$L^{*\prime} = \frac{1}{\bar{\psi} J} u(\theta_t f'(L^{*\prime}) + \frac{\phi}{p}) \tag{1.97}$$

This directly implies $L^{*\prime} > L^*$ by monotonicity of $u(\cdot)$. Hence, in this case employment is going to be higher than in the benchmark case.

Second, I consider the case of binding wage rigidities: $\rho^{*\prime} > \rho > 1$.[64] Recall the implicit definition of $\bar{L}(\bar{w}_{t-1})$ as $\bar{w}_{t-1} = p\theta f(\bar{L}(\bar{w}_{t-1}))$. Note that for $\rho^{*\prime} > \rho > 1$ last period will have more employed workers than this period if wages do not adjust downwards:

$$\rho > 1 \tag{1.98}$$

$$\Rightarrow > \theta_{t-1} > \theta_t p \tag{1.99}$$

Note that $\theta_{t-1} > \theta_t p$ implies $L_{t-1} > \bar{L}(\bar{w}_{t-1})$ by the first order condition (1.5) and the assumed wage setting in the preceding period. This implies that L_j is determined by the first part of (1.94). Comparing the employment in the benchmark case to the employment level now yields:

$$L_j \geq L^*(1, w^*) \text{ if } \bar{w}_{t-1} + \phi \geq w^* \tag{1.100}$$

$$L_j < L^*(1, w^*) \text{ if } \bar{w}_{t-1} + \phi < w^* \tag{1.101}$$

The above implies that the level of employment relative to the reference wage depends on the cost of firing. This shows that employment does not have to decreases if nominal wage do not adjust downwards.

[64] Note that the threshold at which switching is profitable changes once employment frictions are introduced. Using similar arguments as in the proof of Proposition 1, it can be show that $\rho^{*\prime}$ exists.

1.B Additional Tables

Table A1.1: Summary Statistics

		N	mean	sd	min	max
SUSENAS and SAKERNAS data						
District level variables						
Monsoon onset delay		86,610	-0.952	21.94	-64.55	76.51
Inflation		86,610	18.06	25.49	2.351	78.96
Inflation without 1998		73,639	7.451	3.526	2.351	13.95
Rainfall shocks dummies						
Shock last year	Shock this year					
	Positive	86,610	0.171	0.376	0	1
	Negative	86,610	0.128	0.334	0	1
Non-Positive	Negative	86,610	0.093	0.290	0	1
Positive	Negative	86,610	0.036	0.185	0	1
Positive	None	86,610	0.128	0.334	0	1
Individual level variables						
Log nominal agricultural wage		86,610	8.949	0.811	6.359	11.61
Working days in agriculture		1,190,023	0.430	1.557	0	7
Rural		86,610	0.868	0.338	0	1
Male		86,610	0.645	0.478	0	1
Age		86,610	37.24	13.38	10	99
Primary		86,610	0.473	0.499	0	1
IFLS data						
District level variables						
Monsoon onset delay		1,688	9.928	19.79	-32.43	70.91
Inflation		1,688	9.797	0.814	9.111	11.15
Rainfall shocks dummies						
Shock last year	Shock this year					
	Positive	1,688	0.020	0.141	0	1
	Negative	1,688	0.161	0.368	0	1
Non-Positive	Negative	1,688	0.073	0.260	0	1
Positive	Negative	1,688	0.088	0.284	0	1
Positive	None	1,688	0.198	0.399	0	1
Individual level variables						
Log difference nominal agricultural wage		1,688	-0.022	0.482	-1.792	1.593
Dummy for non-negative wage change		1,688	0.774	0.419	0	1
Rural		1,688	0.806	0.396	0	1
Male		1,688	0.575	0.494	0	1
Age		1,688	41.24	14.27	15	84
Primary		1,688	0.723	0.448	0	1

Notes: Positive shocks are defined as a monsoon onset of at at least 20 days before the district mean. Negative shocks are defined as a monsoon onset of at least 30 days after the local mean. Monsoon onset is defined as reaching cumulative rainfall of 20cm after August first. SAKERNAS & SUSENAS data consists of National Socioeconomic Survey data from 1998 to 2004 and of National labor Force Survey data collected in February 2005 to 2007.

Table A1.2: Impacts on Wage Difference

			Source: Indonesian Family Life Survey (1993-2000)		
Dependent variable			log wages differences		
			(1)	(2)	(3)
	Shock last year	Shock this year			
1	any	negative	-0.0366** [0.03]		
2	non-positive	negative		-0.037** [0.04]	
3	non-positive	weak negative			0.039 [0.94]
4	non-positive	strong negative			-0.14** [0.04]
4	positive	negative		-0.033 [0.32]	
5	positive	weak negative			0.004 [1.00]
6	positive	strong negative			-0.038 [0.23]
6	positive	none		.021 [0.71]	0.022 [0.73]
Observations			1,654	1,654	1,654
R-squared			0.005	0.006	0.008
Number of clusters			232	232	40
Time FE			Yes	Yes	Yes

Notes: p-values obtained using wild bootstraps at province-year level with 1000 repetitions displayed in square brackets. Positive shocks are defined as a monsoon onset of at at least 20 days before the district mean. Negative shocks are defined as a monsoon onset of at least 30 days after the local mean. Monsoon onset is defined as reaching cumulative rainfall of 20cm after August first. Weak negative shocks are below median shock strength. Strong negative shocks are above median shock strength. 3. The positive shock category is omitted due to insufficient sample size. Sample restricted to individuals working in the same agricultural job in both years. *** p<0.01, ** p<0.05, * p<0.1.

Table A1.3: Robustness Checks 1: Standard Errors

			Source: SUSENAS & SAKERNAS (1998-2007)[a]		Source: Indonesian Family Life Survey (1993-2000)[b]
Dependent variable			log daily wages	worker days in agriculture	$\mathbb{1}_{\Delta w \geq 0}$
			(1)	(2)	(3)
	Shock last year	Shock this year			
1	any	positive	0.064***	0.164***	
			(0.024)	(0.050)	
2	non-positive	negative	-0.020	-0.076**	-0.066
			(0.020)	(0.039)	(0.041)
3	positive	negative	-0.070*	-0.011	-0.096**
			(0.040)	(0.078)	(0.048)
4	positive	none	0.022	0.139**	0.056
			(0.022)	(0.060)	(0.066)
Observations			86,605	1,190,023	1,654
R-squared			0.617	0.086	0.005
Number of clusters			248	248	144
Clustered at district level			Yes	Yes	Yes
Individuals Controls			Yes	Yes	-
District FE			Yes	Yes	-
Time FE			Yes	Yes	Yes

Notes: Standard errors clustered at district level displayed in parenthesis. Sample restricted to agricultural wage workers Positive shocks are defined as a monsoon onset of at at least 20 days before the district mean. Negative shocks are defined as a monsoon onset of at least 30 days after the local mean. Monsoon onset is defined as reaching cumulative rainfall of 20cm after August first. SAKERNAS & SUSENAS data consists of National Socioeconomic Survey data from 1998 to 2004 and of National labor Force Survey data collected in February 2005 to 2007. The positive shock category is omitted for the IFLS sample due to insufficient sample size. Sample restricted to individuals working in the same job in both years. Dependent variable is a dummy indicating a non-negative wage change. Individual controls include age, age squared, sex and dummies for living in a rural area and having at least primary education. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A1.4: The Effects of Inflation

			Source: SUSENAS & SAKERNAS (1999-2007)	
Dependent variable			log daily wage	worker-days in agriculture
			(1)	(2)
	Last year's shock	This year's shock		
1	any	positive	0.095***	0.234**
			(0.025)	(0.104)
2	interaction with inflation		-0.006	-0.008
			(0.004)	(0.011)
3	non-positive	negative	-0.195*	-0.146
			(0.101)	(0.147)
4	interaction with inflation		0.0257*	0.008
			(0.0154)	(0.0254)
5	positive	negative	-0.133	0.442
			(0.275)	(0.298)
6	interaction with inflation		0.001	-0.022
			(0.036)	(0.033)
4	positive	none	-0.105	0.211
			(0.227)	(0.399)
6	interaction with inflation		0.013	-0.0036
			(0.024)	(0.038)
Observations			73,634	1,065,212
R-squared			0.575	0.084
Number of clusters			232	232
Individuals Controls			Yes	Yes
District FE			Yes	Yes
Time FE			Yes	Yes

Notes: All wage observations are nominal wages. Data source is National Socioeconomic Survey from 1999 to 2004 and of National labor Force Survey in February 2005 to 2007. Sample restricted to individuals working in agricultural. Top and bottom one percent of wage observations dropped are dropped in each sample. Individual observations weighted using cross-sectional sample weights. Inflation measure is national average of inflation excluding inflation in the own province. Positive shocks are defined as a monsoon onset of at at least 20 days before the district mean. Negative shocks are defined as a monsoon onset of at least 30 days after the local mean. Monsoon onset is defined as reaching cumulative rainfall of 20cm after August first. Standard errors clustered at province year level displayed in parenthesis. Individual controls include age, age squared, sex and dummies for living in a rural area and having at least primary education. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A1.5: Serial Correlation of Monsoon Onset

	AR(1) coefficient	
	(1)	(2)
Monsoon onset delay	-0.176***	-0.0045
	(0.012)	(0.012)
District fixed effects	No	Yes
Time fixed effects	No	Yes

Notes: Level of analysis is Indonesian districts. Sample is restricted to districts used in the main analysis. Monsoon onset delay is defined as the deviation in days from the district average monsoon onset. Monsoon onset is defined as reaching cumulative rainfall of 20cm after August first. Standard errors clustered at province year level displayed in parenthesis. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A1.6: Robustness Checks 2: Specific Time Trends

			Source: SUSENAS & SAKERNAS (1998-2007)				Source: Indonesian Family Life Survey (1993-2000)	
	Dependent variable		log daily wage		worker-days in agriculture		$\mathbb{1}_{\Delta w \geq 0}$	
			(1)	(2)	(3)	(4)	(5)	(6)
	Shock last year	Shock this year						
1	any	positive	0.053***	0.028*	0.108***	0.084*		
			(0.016)	(0.015)	(0.036)	(0.033)		
2	non-positive	negative	-0.017	0.007	-0.026	0.021	-0.032	-0.075
			(0.024)	(0.026)	(0.043)	(0.027)	[0.64]	[0.28]
3	positive	negative	-0.067**	-0.064**	0.056	0.030	-0.085	-0.114
			(0.027)	(0.032)	(0.060)	(0.046)	[0.29]	[0.24]
4	positive	none	0.020	0.0043	0.097	0.081	0.065	0.059
			(0.018)	(0.017)	(0.062)	(0.041)	[0.14]	[0.22]
Observations			86,605	86,605	1,190,023	1,190,023	1,654	1,654
R-squared			0.618	0.624	0.097	0.111	0.026	0.115
Number of clusters			232	232	232	232	40	40
Province specific time trends			Yes	No	Yes	No	Yes	No
District specific time trends			No	Yes	No	Yes	No	Yes

Notes: Standard errors clustered at province year level displayed in parenthesis. p-values obtained using wild bootstraps at province-year level with 1000 repetitions displayed in square brackets. Positive shocks are defined as a monsoon onset of at at least 20 days before the district mean. Negative shocks are defined as a monsoon onset of at least 30 days after the local mean. Monsoon onset is defined as reaching cumulative rainfall of 20cm after August first. SAKERNAS & SUSENAS data consists of National Socioeconomic Survey data from 1998 to 2004 and of National labor Force Survey data collected in February 2005 to 2007. The positive shock category is omitted for the IFLS sample due to insufficient sample size. Sample restricted to individuals working in the same job in both years. Dependent variable is a dummy indicating a non-negative wage change. Individual controls include age, age squared, sex and dummies for living in a rural area and having at least primary education. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A1.7: Robustness Checks 3: Absolute Shock Definition

			Source: SUSENAS & SAKERNAS (1998-2007)				Source: Indonesian Family Life Survey (1993-2000)	
	Dependent variable		log daily wage		worker-days in agriculture		$\mathbb{1}_{\Delta w \geq 0}$	
			(1)	(2)	(3)	(4)	(5)	(6)
	Shock last year	Shock this year						
1	any	positive	0.063***	0.053**	0.130***	0.102**		
			(0.0153)	(0.021)	(0.048)	(0.049)		
2	non-positive	negative	-0.018	-0.013	-0.0812	-0.102	-0.063	-0.053
			(0.027)	(0.028)	(0.071)	(0.078)	[0.27]	[0.55]
3	positive	negative	-0.068***	-0.071**	-0.042	-0.108	-0.093*	-0.083
			(0.026)	(0.029)	(0.150)	(0.160)	[0.08]	[0.14]
4	positive	none	0.019	0.020	0.122*	0.152**	0.060	0.113
			(0.020)	(0.025)	(0.069)	(0.065)	[0.30]	[0.13]
Observations			82,012	67,880	1,103,523	927,880	1,530	1,410
R-squared			0.608	0.563	0.085	0.075	0.005	0.005
Number of clusters			232	232	232	232	40	40
Excluded districts with no shock category			Yes	No	Yes	No	Yes	No
Excluded district with missing shock category			No	Yes	No	Yes	No	Yes

Notes: Districts with no shock category are districts which did not experience any shock in the year 1979-2010. Districts with missing shock category are districts which did not experience either a positive, a negative or either shock Standard errors clustered at province year level displayed in parenthesis. p-values obtained using wild bootstraps at province-year level with 1000 repetitions displayed in square brackets. Positive shocks are defined as a monsoon onset of at at least 20 days before the district mean. Negative shocks are defined as a monsoon onset of at least 30 days after the local mean. Monsoon onset is defined as reaching cumulative rainfall of 20cm after August first. SAKERNAS & SUSENAS data consists of National Socioeconomic Survey data from 1998 to 2004 and of National labor Force Survey data collected in February 2005 to 2007. The positive shock category is omitted for the IFLS sample due to insufficient sample size. Sample restricted to individuals working in the same job in both years. Dependent variable is a dummy indicating a non-negative wage change. Individual controls include age, age squared, sex and dummies for living in a rural area and having at least primary education. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A1.8: Robustness Checks 4: Shock Controls and Definition

			Source: SUSENAS & SAKERNAS (1998-2007)				Source: Indonesian Family Life Survey (1993-2000)	
	Dependent variable		log daily wage		worker-days in agriculture		$\mathbb{1}_{\Delta w \geq 0}$	
			(1)	(2)	(3)	(4)	(5)	(6)
	Shock last year	Shock this year						
1	any	positive	0.058***	0.063***	0.187***	0.166***		
			(0.017)	(0.015)	(0.051)	(0.048)		
2	non-positive	negative	-0.023	-0.026	-0.090	-0.028	-0.049*	-0.099
			(0.027)	(0.027)	(0.067)	(0.057)	[0.05]	[0.16]
3	positive	negative	-0.083***	-0.082***	-0.004	0.030	-0.075	-0.141**
			(0.026)	(0.031)	(0.143)	(0.113)	[0.12]	[0.02]
4	positive	none	0.022	0.023	0.155**	0.139**	0.046	0.039
			(0.018)	(0.020)	(0.069)	(0.069)	[0.43]	[0.46]
Observations			86,605	86,605	1,190,023	1,190,023	1,654	1,654
R-squared			0.617	0.617	0.087	0.086	0.011	0.006
Number of clusters			232	232	232	232	40	40
Shock history controls			Yes	No	Yes	No	Yes	No
Symmetric shock definitions			No	Yes	No	Yes	No	Yes

Notes: Shock history control include dummies for lagged positive shock in periods $t-2$ to $t-5$. 2. Symmetric positive (negative) shocks are defined as monsoon onset of 20 days before (after) the district mean. Monsoon onset is defined as reaching cumulative rainfall of 20cm after August first. Standard errors clustered at province year level displayed in parenthesis. p-values obtained using wild bootstraps at province-year level with 1000 repetitions displayed in square brackets. SAKERNAS & SUSENAS data consists of National Socioeconomic Survey data from 1998 to 2004 and of National labor Force Survey data collected in February 2005 to 2007. The positive shock category is omitted for the IFLS sample due to insufficient sample size. Sample restricted to individuals working in the same job in both years. Dependent variable is a dummy indicating a non-negative wage change. Individual controls include age, age squared, sex and dummies for living in a rural area and having at least primary education. *** p<0.01, ** p<0.05, * p<0.1.

Table A1.9: Robustness Checks 5: Cross-Sectional Sample Selection

Dependent variable			log daily wage					log hourly wage
			Include casual workers[a]	Drop observations in 1998	Restrict to IFLS districts[b]	Exclude Sumatra	Restrict to SUSENAS data	
			(1)	(2)	(3)	(4)	(5)	(6)
	Shock last year	Shock this year						
1	any	positive	0.059***	0.063***	0.061***	0.089***	0.071***	0.097***
			(0.015)	(0.012)	(0.019)	(0.020)	(0.013)	(0.017)
2	non-positive	negative	-0.019	-0.044	-0.012	-0.020	-0.009	-0.019
			(0.022)	(0.039)	(0.030)	(0.026)	(0.031)	(0.034)
3	positive	negative	-0.083***	-0.136**	-0.062**	-0.051*	-0.044	-0.058
			(0.026)	(0.057)	(0.025)	(0.030)	(0.028)	(0.036)
4	positive	none	0.026	0.019	0.020	0.060**	0.028*	0.038**
			(0.025)	(0.022)	(0.022)	(0.029)	(0.017)	(0.018)
Observations			111,445	73,634	61,425	55,467	81,863	86,662
R-squared			0.619	0.575	0.610	0.515	0.541	0.552
Number of clusters			232	208	145	155	162	232

Notes: Data source is repeated cross-section consists of National Socioeconomic Survey data from 1998 to 2004 and of National labor Force Survey data collected in February 2005 to 2007. Sample restricted to agricultural wage workers. Top and bottom one percent of wage observations dropped are dropped in each sample. Individual observations weighted using cross-sectional sample weights. Positive shocks are defined as a monsoon onset of at at least 20 days before the district mean. Negative shocks are defined as a monsoon onset of at least 30 days after the local mean. Monsoon onset is defined as reaching cumulative rainfall of 20cm after August first. Standard errors clustered at province-year level in parenthesis. All specification include time and district fixed effects and the individual controls age, age squared, sex and dummies for living in a rural area and having at least primary education. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

[a] This includes wage observations of casual wage workers for the years, 2000 and 2005 to 2007.

[b] This excludes districts with no wage observation in the IFLS survey.

1.C Additional Figures

Figure 1.C1: Spatial Distribution of Districts

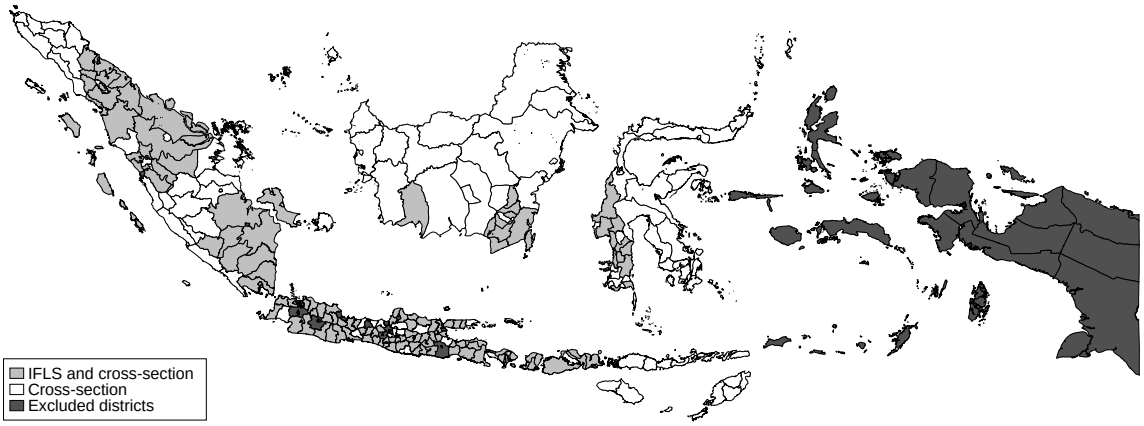
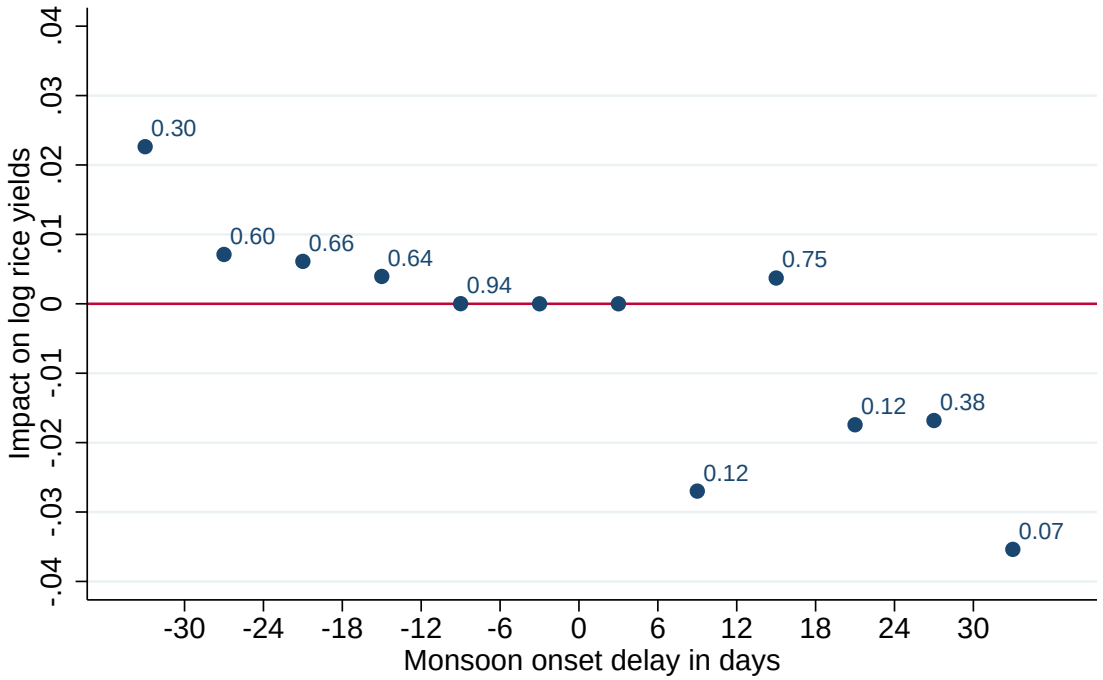
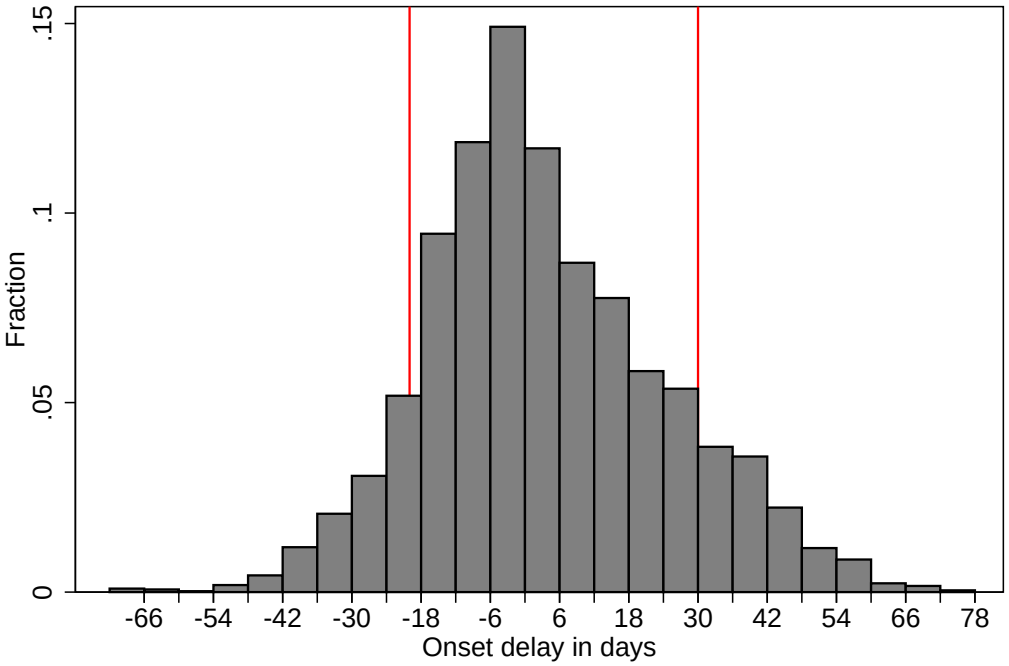


Figure 1.C2: Non-Parametric Impact of Monsoon Onset on Rice Yields



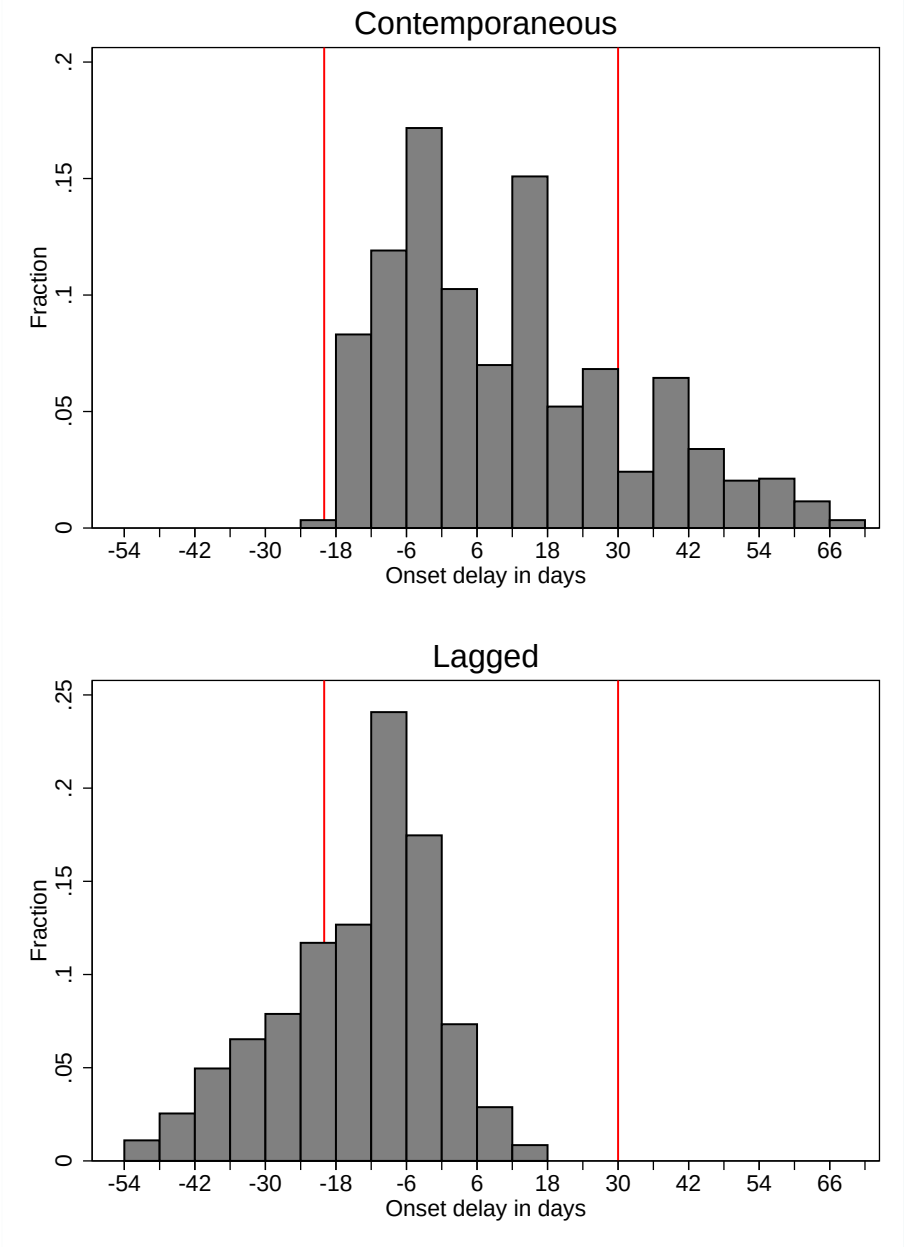
Impact of monsoon onset delay in 6 day intervalls. First (last) category defined as onset deviation of less (more) than -30 (30) days. Onset deviations between -6 and 6 days are the baseline category. Point estimate labels indicate wild bootstraped p-values with 1000 replications.

Figure 1.C3: Distribution of District Monsoon Onset Delay



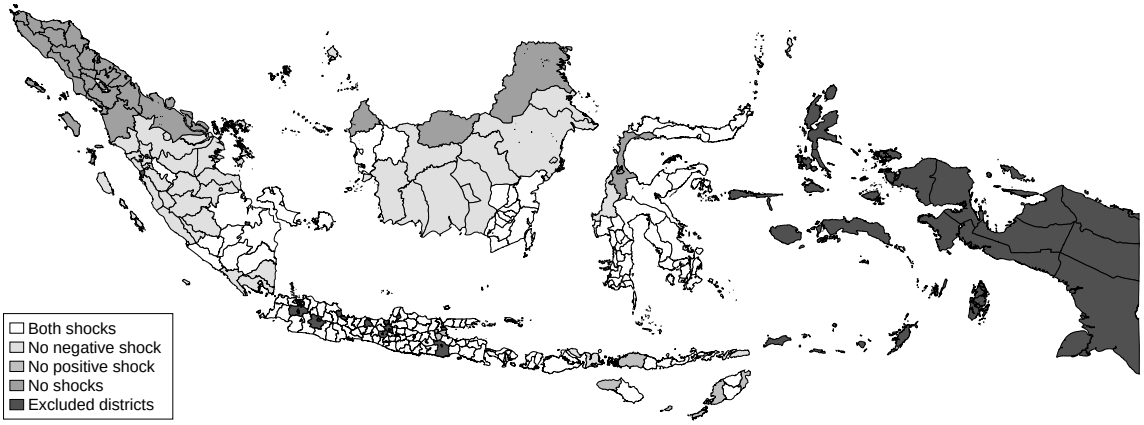
Restricted to years with rice yield observation and to provinces used in the main analysis. Vertical lines indicate cutotts for positive and negative shocks.

Figure 1.C4: Distribution of Monsoon Onset Delay for IFLS data



Vertical lines indicate cutotts for positive and negative shocks.

Figure 1.C5: Spatial Distribution of District without Rainfall Shocks



Chapter 2

Income Shocks and Suicides: Causal Evidence From Indonesia

with **Cornelius Christian**, Brock University and **Christopher Roth**, Institute on Inequality and Beahvior

Abstract We examine how income shocks affect the suicide rate in Indonesia. We use a difference-in-differences approach, exploiting a cash transfer program's nationwide roll-out, and corroborate the findings using a randomized experiment. Our estimates show that the cash transfers reduce the yearly suicide rate by 0.36 per 100,000 people, corresponding to an 18 percent decrease. Moreover, a different type of income shock, variability in agricultural productivity, also affects the suicide rate. The cash transfer program reduces the causal impact of the agricultural productivity shocks, suggesting an important role for policy interventions. Finally, we provide evidence for depression as a psychological mechanism.

2.1 Introduction

Suicide is a pressing public health concern, causing 800,000 deaths per year globally (WHO, 2014). Suicidal behavior is one of the leading causes of death for people with severe mental illness (Hawton et al., 2013). While the effects of economic conditions on mental health and depression are well-established (Haushofer and Shapiro, 2016; Baird et al., 2011), little is known about the causal effect of economic well-being on suicidal behavior. More specifically, no causal evidence exists that examines whether positive income shocks, such as poverty alleviation programs, can decrease suicides.

It is difficult to credibly quantify the impact of improved economic well-being and, in particular, poverty alleviation programs on suicide rates. First, it is challenging to find suicide data at sufficiently low levels of geographic disaggregation to ensure statistical power. Second, the timing and geographic placement of large government programs is usually endogenous, and even if small-scale programs demonstrate causality, external validity concerns predominate (Deaton, 2010; Allcott, 2015).

Our data and empirical setting address these difficulties. We focus on Indonesia, the world's fourth most populous country. We leverage unique Indonesian village-level census data from 2000 to 2014, and use several identification strategies to examine the causal effect of income shocks: The first is a conditional cash transfer program, Program Keluarga Harapan (PKH), which provides households with yearly cash transfers worth about 10% of their pre-treatment annual consumption over six years. Using a difference-in-differences specification, we estimate the effect of the program's nationwide roll-out. The program targeted subdistricts with high levels of poverty rates, sufficient supply-side health and educational institutions, and was launched in 2007 when it covered 13% of Indonesian subdistricts in seven provinces. The program was widely expanded, reaching 57% of subdistricts across all 33 provinces in 2013. We use two stages of the roll-out as identifying variation: the first stage of subdistricts who received the treatment from 2007 to 2011 and the second stage of subdistricts who received the treatment in 2012 and 2013. We also analyze a randomized controlled trial of the same program. Second, we exploit plausibly exogenous agricultural productivity shocks, using rainfall variation, and examine whether the effects of the cash transfer vary systematically with these agricultural productivity

shocks.

We establish several novel facts regarding the relationship between economic well-being and suicides. Our first finding is that both the cash transfer's roll-out and the randomized cash transfer cause a large reduction of suicides. Our main specification suggests that an average per-capita transfer of 22.45 USD targeted at the poorest 10 percent of households causes a decrease in the suicide rate by approximately 0.36 suicides per 100,000 population per year. This corresponds to a reduction of the suicide rate by approximately 18 percent relative to the mean suicide rate (in 2011 and 2014) of 2 suicides per 100,000 in control subdistricts. Second, we show that agricultural productivity shocks, proxied by rainfall, significantly affect the incidence of suicide. A one standard deviation increase in rainfall increases yearly per-capita consumption by 21.6 USD and decreases the number of suicides per 100,000 inhabitants by approximately 0.08, corresponding to a reduction of 6 percent relative to the average suicide rate in Indonesia between 2000 and 2014. For the subsamples affected by the income shocks, our calculations suggest that per-dollar effects on suicides identified using rainfall are significantly smaller than those from cash transfers. Third, we establish that cash transfers lower suicides most strongly in subdistricts experiencing negative agricultural productivity shocks. This is consistent with social welfare programs mitigating the adverse effects of negative economic shocks.

Finally, we provide evidence that income shocks may ignite suicidal behavior by affecting people's mental health. The medical literature suggests that stress (Mann et al., 1999) and mental illness, such as depression, are major causes for suicides (Boldrini and Mann, 2015). We use panel data from the Indonesian Family Life Survey to show that agricultural productivity shocks causally affect farmers' mental health. A one standard deviation increase in subdistrict level rainfall increases consumption by approximately seven percent, and reduces depression by 0.12 of a standard deviation. Moreover, the relationship between rainfall and depression is absent for non-farmers, indicating that the effects of rainfall operate through an economic channel.

We contribute to the literature on how economic and social circumstances affect suicides (Becker and Woessmann, 2018; Daly et al., 2013; Cutler et al., 2001; Stevenson and Wolfers, 2006; Ludwig et al., 2009; Campaniello et al., 2017). Concurrent work by Carleton (2017) shows that across Indian states, for temperatures above 20°C, a 1°C rise causes roughly 70 suicides

per day, particularly during the agricultural growing season.[1] Our evidence supports the view that the positive effects of the cash transfer program on recipients' mental health (Haushofer and Shapiro, 2016) outweigh negative spillovers of such programs (Baird et al., 2013).

Our paper also contributes to the literature on poverty, income shocks, and mental health, the latter of which is usually measured through self-reported scales (Adhvaryu et al., 2019; Gardner and Oswald, 2007; Cesarini et al., 2015; Friedman and Thomas, 2009; Das et al., 2007; Persson and Rossin-Slater, 2018; Kuhn et al., 2011; Baicker et al., 2013; Kling et al., 2007; Devoto et al., 2012). In particular, previous papers examine the effects of randomized cash transfers on mental health through survey questions (Paxson and Schady, 2007; Baird et al., 2013). An exception is Haushofer and Shapiro (2016) who show that large, unconditional cash transfers can reduce cortisol levels, consistent with self-reported reductions in stress.

Our study advances the literature on poverty and mental health in three ways. First, we provide the first causal evidence on whether positive income shocks, and in particular government poverty alleviation programs, reduce suicides. Second, unlike previous evidence on income shocks and mental health which mostly relies on small-scale experiments, we use both a large-scale nationwide roll-out of a conditional cash transfer program and a randomized evaluation of the same program. Finally, we provide the first evidence on the interaction between social welfare programs and agricultural productivity shocks on mental health, underlining the potential of government programs to alleviate negative consequences of adverse economic shocks.

We proceed as follows. In Section 2.2, we outline a conceptual framework, describe the data, and the cash transfer program. In Section 2.3, we present our identification strategies, and describe our results using both the cash transfer and agricultural productivity shocks. In Section 2.4, we examine mechanisms underlying our estimated treatment effects by employing microdata on depression. In Section 2.5, we discuss effect sizes and we conclude in Section 2.6.

2.2 Conceptual Framework, Context, and Data

In this Section we describe the conceptual framework, the main features of the conditional cash transfer program, the suicide data and the construction of the subdistrict panel. Finally, we

[1] There is also correlational evidence highlighting that recessions increase suicides (Reeves et al., 2012; Chang et al., 2013).

report some basic descriptive statistics about Indonesian suicide rates.

2.2.1 Conceptual Framework

Related literature What are the theoretical links between economic well-being and suicides? Previous psychiatric disorders are the most important factor in explaining death by suicide; 90% of those who committed suicide had such a disorder (Cavanagh et al., 2003). Hawton et al. (2013) claim that specific mental health problems, such as depression, are mainly responsible for the decision to commit suicide. Poverty and negative economic shocks are also associated with mental health disorders and depression in particular (Haushofer and Fehr, 2014; McInerney et al., 2013; Schilbach et al., 2016). Negative life experiences, such as loss of income and job loss, have also been identified as risk factors for suicide.[2]

Stress is a likely mechanism at play, as suicide risk is correlated with abnormal cortisol concentrations, and a maladaptive cortisol response to stress (O'Connor and Nock, 2014). This is in line with psychological models of suicide which emphasize that pre-existing medical conditions coupled with stress lead to suicidal behavior (Mann et al., 1999).[3]

The main object of interest in this paper is to understand how economic shocks affect the suicide rate. Positive economic shocks could affect the suicide rate directly by mitigating the consequences of negative life experiences or through improvements in mental health. In Section 2.4, we provide suggestive evidence in support of mental health as a channel by analyzing the relationship between economic shocks and depression.

Functional Form Our data and setting allow us to provide evidence on the functional form of the relationship between the suicide rate and income. Models of reference-dependent preferences predict declining effects of permanent income shocks over time as individuals' reference points

[2] In addition, there is a literature emphasizing that suicides can be contagious, i.e. that the social environment plays an important role (Hedström et al., 2008). There are also literatures that examine how genes (Roy, 1992), social isolation (Appleby et al., 1999), and personality traits (Blüml et al., 2013) affect people's tendency to commit suicide.

[3] Joiner (2005) puts forward an interpersonal theory of suicide, which posits that the coexistence of high feelings of burdensomeness, low levels of belongingness, and the belief that these conditions are hopeless to change, lead to the development of suicidal desires. Finally, social isolation and lack of social support consistently predict suicide risk (Appleby et al., 1999).

adapt to the income level.[4] Indeed, previous evidence suggests that the effects of permanent improvements in economic circumstances on self-reported mental well-being decline over time as individuals adapt to their new economic situation (e.g. Galiani et al., 2018; Frederick and Loewenstein, 1999). We test the predictions of models of reference-dependence by analyzing the dynamic treatment effects over the program duration of six years.

If the marginal impact of income on the suicide rate decreases in income, then we expect a convex relationship between the suicide rate and income. In this case, the impact of a positive income shock should be larger if the recipients of the shock are poorer to begin with. If the relationship is linear there should be no heterogeneity by initial poverty levels. We test the implications of different functional forms by analyzing treatment heterogeneity of the cash transfer program by the extent of initial poverty, as proxied by agricultural productivity shocks. We also use micro-data to test whether the effects of agricultural productivity shocks on depression depend on initial poverty as proxied by pre-treatment per-capita household consumption.

Types of Income Shocks Different types of income shocks may have different quantitative impacts on the suicide rate. The effect of an income shock likely depends on its effect on expected lifetime income and the uncertainty about lifetime income, that is its current consumption value. Therefore, predictable and regular cash transfers should have larger per-dollar impacts on the suicide rate than unpredictable agricultural income shocks, a prediction which we test in this paper.

2.2.2 The Conditional Cash Transfer Program

We use the Program Keluarga Harapan (PKH) conditional cash transfer program to analyze the impact on suicide rates (The World Bank, 2011; Banerjee et al., 2017; Cahyadi et al., 2018). A pilot version of the PKH was introduced in 2007 for 600,000 households and the program was then expanded to cover 5.2 million households in 2014 with a target of 10 million recipients in 2018 (The World Bank, 2017). The introduction of the PKH program was part of a wider effort to reform the Indonesian social security system. In 2005 the Indonesian government removed universal fuel subsidies. To alleviate the immediate inflationary shock for poor

[4] Models in which the reference point fully adapts after one period predict that there is a fall in suicides only in the first year of a permanent income shock and no effects thereafter.

and near-poor households, the Indonesian government introduced an unconditional cash transfer program covering 19 million households from 2005 to 2006. Furthermore, existing large-scale social assistance programs, such as rice subsidies for poor households (Beras untuk Rakyat Miskin, or Raskin) and subsidized health insurance for the poor (Asuransi Kesehatan Miskin, Askeskin), were expanded. The PKH program was introduced in 2007 after the unconditional cash transfer ended to provide more targeted assistance to the poorest households (The World Bank, 2012).[5]

PKH was designed to improve poor households' health and education through a cash transfer, conditional on their participation in health and education services (The World Bank, 2011). The intervention's size was substantial: households received between 39 and 220 US dollars per year. The average received amount constituted about 10% of pre-PKH yearly household expenditure (80.82 US dollars at 2005 prices) between 2007 and 2014 (The World Bank, 2011, 2017). Households were part of the PKH program for up to six years.

The total cost of the PKH program from 2007 to 2014 was around 716 million USD at 2005 prices. In total there were 7.6 million household-years of cash transfer such that the average expenditure per household-year was 94 USD (The World Bank, 2017). According to the 2011 village census, the average household size in Indonesia in 2011 was 3.6 individuals which implies an expenditure of 26.11 USD per treated individual and year. Subtracting administrative overhead households received 22.45 USD per capita on average.

Geographic Roll-Out of Program In Section 2.3.1 we exploit two stages of the roll-out of the PKH program as identifying variation for the difference-in-differences estimation: the first stage of subdistricts who received the treatment from 2007 to 2011 and the second stage of subdistricts who received the treatment in 2012 and 2013. The program was first implemented as a pilot program in 2007 in seven provinces: West Java, East Java, North Sulawesi, Gorontalo, East Nusa Tenggara (NTT), West Sumatra and DKI Jakarta. These provinces are quite diverse in terms of their poverty levels, and other economic and geographic characteristics. Because the program's focus is on poverty alleviation, upper income quintile districts were initially excluded from PKH eligibility, based on an index considering poverty rates, malnutrition and schooling

[5] PKH's coverage is significantly below of that of the other social benefit programs, but due to its targeted nature the World Bank considers it one of the most effective Indonesian social assistance programs (The World Bank, 2017).

records.[6] Most of the 2007 roll-out of the program was randomized among selected eligible districts.

From 2008 to 2010 the program maintained its pilot status but was further expanded in the following provinces: Nanggroe Aceh Darussalam, North Sumatra, Banten, South Kalimanten, West Nusa Tenggara, and the Yogyakarta Special Region. Since 2010 the Secretariat of the National Team for the Acceleration of Poverty Reduction (TNP2K), at the Office of the Vice-President, has been promoting the nationwide expansion of PKH leading to all Indonesian provinces being covered by 2012. Of the subdistricts included in our analysis, the PKH program covered about 13% of all Indonesian subdistricts when it started in 2007. By 2013, 57% of all Indonesian subdistricts in our sample were in receipt of the program.

At the macro level, targets for overall recipient numbers and total expenditure were set at the national level which determined the overall speed of the roll-out. At the micro level, target subdistricts were determined in cooperation with provincial and district governments who made recommendations. The final decisions took into account three main factors: subdistrict poverty levels, existence of the necessary supply-side institutions (e.g. educational institutions and health centers), and the willingness of local partners to cooperate.

Randomized Experiment In Section 2.3.2 we use a subsample of subdistricts in which the treatment status was randomly assigned and which formed the basis for the World Bank's evaluation of the program (The World Bank, 2011). A total of 736 subdistricts were included in the sample, with 438 subdistricts randomized to the treatment group (Cahyadi et al., 2018). Of these, we observe treatment assignment for the 360 subdistricts that were randomly chosen for data collection by the World Bank (180 treatment, 180 control). Political pressures and a consequent unexpected program expansion in East Java resulted in deviations of the realized allocation from the intended one.[7] To deal with this contamination, we use the orig-

[6] Districts receiving the rural community-driven development project were eligible for the community cash transfer program (PNPM Generasi), and therefore were ineligible for the PKH program during the pilot operation.

[7] In particular, 37 out of the 360 subdistricts that were supposed to be part of the control group received PKH funds before 2011. Moreover, for a very few subdistricts, the program started in 2008 or 2009 rather than in 2007. Bias might result from this contamination, since it is possible that unobserved factors within the contaminated subdistricts also affected household responses. The contamination increased further leading to 30 percent of control subdistricts receiving the program by 2014.

inal treatment assignment to measure the conditional cash transfer program's impact on suicide.

Beneficiary Selection At the subdistrict level, the cash transfer program was offered to a list of eligible households that satisfied both certain demographic as well as certain economic requirements. A 2005 census from a national unconditional cash transfer program was initially used to construct the list of eligible households per village. Approximately 30-40 percent of beneficiaries from the unconditional cash transfer program were not included in the list of eligible households.[8] Based on this list of households, demographic data was used to identify eligible households that fulfilled one of the following program criteria: (i) households with pregnant and/or lactating women; (ii) households with children aged 0-15 years; (iii) households with children aged 16-18 years who have not yet completed 9 years of basic education.

However, only the subset of eligible households with the lowest predicted consumption were included in the program. In the end, approximately 10% of households received the program. The classification was based on proxy-means tests of all households on the list of eligibles to identify program beneficiaries. The proxy-means tests consisted of 29 variables, including housing characteristics, education levels, sources of fuel, employment information and access to health and education services.

2.2.3 Data

We use the censuses of all Indonesian villages (PODES) from 2000, 2003, 2005, 2011, and 2014 to examine the PKH program's effect on suicide rates. The PODES data covers all 80,000 villages in Indonesia. In the village census, village heads report village characteristics, such as population size, the presence of health and educational institutions, and the percentage of farmers.

Outcome Definitions The census contains data on suicides at the village level. In 2000, 2003 and 2005 the village head was asked whether any suicide occurred in their village in the previous year. In 2011 village heads were asked about the number of suicides committed in the last year. Finally, in 2014 the PODES survey asked for the number of suicides and suicide attempts in the

[8] Statistics Indonesia also conducted additional interviews in targeted subdistricts to identify newly poor households, in an attempt to minimize exclusion errors.

village during the last year. Village level population data is available from 2000 to 2011, but not for 2014. To obtain population data for 2014, we extrapolate population size using a linear trend from the years 2005 and 2011 to 2014 at the subdistrict level.

We use the PODES census data to construct the main outcome measure of interest: the number of suicides per 100,000 inhabitants at the subdistrict level. For the years 2000, 2003 and 2005 (all prior to the cash transfer program), we use the number of villages with at least one suicide per 100,000 inhabitants as a proxy for the actual suicide rate. For the PODES 2011 data we directly use the number of suicides per 100,000 inhabitants at the subdistrict level. Lastly, we define the suicide rate in 2014 as the number of suicides and suicide attempts per 100,000 inhabitants.

We also construct a measure of the suicide rate for which the definition does not vary across years (except for 2014 when the question included suicide attempts). Specifically, we extrapolate the expected suicide rate from the subdistrict mean of village level occurrences of at least one suicide. This requires two assumptions: first, we assume that suicides are Poisson-distributed. Second, we assume that suicides in all villages of the same subdistrict are independent and have the same Poisson parameter λ. Under these assumptions we can use the mean incidence of at least one suicide ($\bar{s}$) to calculate the expected number of suicides $E_v(s)$ in a given village as $E_v(s) = -\ln(1-\bar{s})$.[9] To calculate the expected number of suicides at the subdistrict level we multiply $E_v(s)$ by the number of villages in a subdistrict. We validate this measure with data on the actual number of suicides in the years 2011 and 2014 and find a correlation of $\rho = 0.95$. Moreover, in Section 2.3.4 we show that our results are robust to using different outcome measures.

According to the WHO Health Data repository the age-standardized suicide rate in Indonesia stood at around 3 per 100,000 in 2015. This seems relatively well-aligned with our data, where we find mean raw suicide rates of approximately 2 per 100,000 at the subdistrict level in 2011

[9] This is derived from the cumulative distribution function of the Poisson Distribution: $CDF(k) = e^{-\lambda}\sum_{i=0}^{\lfloor k \rfloor}\frac{\lambda^i}{i!}$. We observe the fraction of villages with zero reported suicides, which gives a subdistrict specific estimate of $CDF(0)$. Expressed as fraction of villages with at least one suicide we get: $CDF(0) = 1-\bar{s}$. Using the functional form of the Poisson distribution we obtain $E_v(s) = \lambda = -\ln(1-\bar{s})$. There is one subdistrict in 2014 where all villages have at least one suicide or suicide attempt. For this observation we use the average subdistrict λ over the preceding four census waves.

and 2014.[10]

Indonesia has low suicide rates from an international perspective. In a WHO world-wide ranking of all nations by suicide rates, Indonesia ranked 173 out of 183 nations. This low baseline suicide rate could imply stronger social norms against suicides in Indonesia. These norms might not only affect the level of the suicide rate, but also the elasticity of suicides with respect to economic shocks. However, it is theoretically unclear how these norms would affect the relationship between the suicide rate and economic circumstances. On the one hand, decisions to commit suicide might be more marginal than in other countries, that is a small improvement in economic circumstances could prevent more suicides. On the other hand, it is possible that because of the strong stigma non-economic factors play a more important role for suicides in Indonesia.

Subdistrict Panel Construction Since our key identifying variation is at the subdistrict level, we aggregate our village panel at the subdistrict level and collapse our observations at the subdistrict boundaries from 2000. We use these subdistrict boundaries as 2000 is the first year for which suicide data is available.[11] For practical reasons we construct our panel at the subdistrict level as the number of administrative units in Indonesia substantially increased over time.[12] In the experimental sample there are 314 subdistricts according to 2000 boundary definitions. In our analysis we only make use of 310 subdistricts for which we can construct a cross-walk between 2000 and 2014. For the roll-out of the program and the analysis of agricultural productivity shocks (presented in Section 3.4) we employ the universe of Indonesian subdistricts for which we could construct a consistent panel between 2000 and 2014. We were able to construct such a panel for 3138 out of all 3928 subdistricts according to 2000 boundary definitions. The panel's construction was based on a subdistrict-level crosswalk for the time period of 2000 to

[10] The slight discrepancy between those numbers can be explained by the fact that the WHO estimates are model-based and only partially take into account micro-data on suicides in Indonesia.

[11] We use 2000 borders to provide a consistent interpretation of coefficients for all our subdistrict level specifications. Furthermore, using subdistrict border definitions from after 2000 would complicate both the presentation of pre-trends and the analysis of interactions between agricultural productivity shocks and cash transfers (presented in Section 2.3.1 and 2.3.6). We show that our results are robust to using the 2006 subdistrict borders (Tables A2.8 and A2.14).

[12] Decentralization reforms beginning from 1998 significantly increased the proliferation of administrative units. For example, the number of districts increased from 302 in 1999 to more than 500 in 2014 (Bazzi and Gudgeon, 2019). The number of subdistricts increased from about 3000 in the late 1990s to approximately 7000 in 2014. Constructing a cross-walk at the village level is particularly challenging and would necessarily result in a large number of incorrect matches over time. This would substantially increase measurement error of outcomes.

2014. Owing to the subdistrict-splits, there are 1485 cases in which a subdistrict split from 2000 to 2014. If only a part of the 2000 subdistrict received the cash transfer in a given year, we define the treatment indicator as the fraction of new subdistricts receiving the PKH program. When we drop the observations with partially treated origin subdistricts our estimated treatment effects barely change (see Tables A2.6 and A2.12).

Descriptives and Correlates of Suicide Rates Table 2.1 displays descriptive statistics at the subdistrict level from the PODES 2005 data, before the cash transfer was implemented. On average, a subdistrict consists of 17.2 villages and has a population of approximately 56,000. Most subdistricts are relatively rural, with on average 74% of villages classified as rural and 56% of the population working as farmers. Thus, agricultural productivity shocks are likely to affect large parts of the population. The RCT sample is more rural than the average, but otherwise is similar to non-RCT subdistricts.

Table 2.1: Summary statistics

	Mean	SD	Median	Min.	Max.	Obs.
Panel A: Population data						
Suicide rate	1.39	2.14	0.65	0.00	64.52	3138
Any suicide	0.53	0.50	1.00	0.00	1.00	3138
Educ. inst./100k pop.	152.03	49.66	146.20	11.84	781.56	3138
Health inst./100k pop.	181.98	59.06	174.19	17.88	788.13	3138
% vill. with paved road	0.72	0.32	0.83	0.00	1.00	3138
% vill. with lighting	0.78	0.32	1.00	0.00	1.00	3138
% rural villages	0.74	0.40	1.00	0.00	1.00	3138
Number of villages	17.22	12.75	14.00	2.00	137.00	3138
Population size	56021	40427	47484	2371	415394	3138
Number of families	14090	9851	12182	540	90682	3138
Percentage farmers	56.84	30.97	67.17	0.00	100.00	3138
Panel B: Early adopters (2007-2011)						
Suicide rate	1.31	1.86	0.55	0.00	14.31	713
Any suicide	0.54	0.50	1.00	0.00	1.00	713
Educ. inst./100k pop.	142.78	49.19	139.42	58.82	475.15	713
Health inst./100k pop.	173.71	53.78	170.60	17.88	480.05	713
% vill. with paved road	0.77	0.30	0.92	0.00	1.00	713
% vill. with lighting	0.86	0.27	1.00	0.00	1.00	713

% rural villages	0.73	0.41	1.00	0.00	1.00	713
Number of villages	14.03	7.36	13.00	3.00	77.00	713
Population size	63903	43382	53035	5594	314932	713
Number of families	15996	10030	13546	1523	68967	713
Percentage farmers	51.89	31.32	62.09	0.00	98.06	713
Panel C: RCT data						
Suicide rate	1.41	1.98	0.63	0.00	14.50	310
Any suicide	0.52	0.50	1.00	0.00	1.00	310
Educ. inst./100k pop.	161.63	49.83	151.46	69.72	475.15	310
Health inst./100k pop.	181.80	43.36	176.57	82.88	504.15	310
% vill. with paved road	0.70	0.29	0.79	0.00	1.00	310
% vill. with lighting	0.78	0.32	0.98	0.00	1.00	310
% rural villages	0.90	0.27	1.00	0.00	1.00	310
Number of villages	16.74	6.85	16.00	3.00	46.00	310
Population size	81363	48614	66604	10116	260321	310
Number of families	15558	9121	13361	2345	68967	310
Percentage farmers	60.00	23.85	64.00	0.00	98.00	310

Table 2.1 displays summary statistics based on the Indonesia Village census 2005 . Observations are weighted using the 2005 population except for number of villages, population size, and number of families. Panel A shows summary statistics for the population of subdistricts in Indonesia. Panel B shows summary statistics for subdistricts that received

We exploit baseline data from the 2005 census to characterize the correlates of suicides. We find a strong economic gradient in suicide rates. Subdistricts in districts with a 10% larger share of the households below the poverty line have, on average, a 0.142 higher suicide rate per 100,000 people (see Column 1 of Table A2.1 in the Online Appendix). The same pattern is apparent when we consider per capita expenditure at the district level. The share of farmers at the subdistrict level is strongly positively correlated with suicide rates. The share of farmers at the subdistrict level remains significantly correlated with suicide rates after controlling for local crime rates, health, education, and social institutions.[13] Table A2.1 also reveals that crime per capita is weakly positively related to suicide incidence, but that social organizations per capita and educational institutions per capita are not correlated with suicide rates.[14]

Finally, we observe an increase in the occurrence of any suicide in subdistricts over time.

[13]Since the share of farmer variable is available at the subdistrict level, this may explain why it remains statistically significant, while per capita expenditure and fraction poor which are measured at the district level, become statistically insignificant.

[14]We find that health institutions per capita are positively correlated with suicide rates, consistent with the government targeting health care provision to more needy subdistricts.

The incidence of at least one suicide at the subdistrict level increased from 21 percent in 2000 to 45 percent in 2005 and 52 percent in 2011 and 2014.[15,16]

2.3 Main Results

In this Section, we first present evidence from the difference-in-differences approach using the nationwide roll-out of the cash transfer program. We then show results of a randomized controlled trial of the same program. Thereafter, we examine the dynamics of treatment effects, and assess the robustness of our findings. Finally, we study how agricultural productivity shocks affect suicide rates, and examine how they interact with the roll-out of the cash transfer program.

2.3.1 Nationwide Program Roll-Out

We provide evidence that the PKH conditional cash transfer program substantially decreased the suicide rate using a difference-in-differences approach exploiting the nationwide roll-out of the program. For our main specification we use suicide data from the census of villages from 2005, 2011, and 2014, but we also employ data from 2000 and 2003 to assess robustness and pre-trends. Our dependent variable is the number of suicides per 100,000 individuals (y_{st}) in subdistrict s and at time t.

To estimate treatment effects we include subdistrict fixed effects (α_s), time fixed effects (ϕ_t), and a treatment indicator, Treat_{st}, taking value one when a subdistrict started receiving the program.[17] This means that we effectively exploit differences in the timing of the receipt of the cash transfer program at the subdistrict level to estimate treatment effects. In particular, we use two stages of the roll-out as identifying variation to estimate treatment effects: the first stage of subdistricts who received the treatment from 2007 to 2011 and the second stage of subdistricts who received the treatment in 2012 and 2013. We use all subdistricts that had not received the cash transfer program in a given census wave as controls. Even though the identifying variation

[15] The increase in reported suicides between 2000 and 2005 could be the result of shifting norms around suicides, potentially affecting the willingness to report suicides by the village chiefs. However, such a shift in norms is unlikely to affect the direction and magnitude of our treatment effects estimates. For our preferred specification, we rely on data from 2005 to 2014 directly avoiding the influence of potential under-reporting in 2000 and 2003. Furthermore, even when we include the 2000 and 2003 census waves, the treatment effect point estimate hardly changes suggesting a limited of role of this under-reporting. Lastly, a shift of norms could only effect the direction of the treatment effect if it occurred differentially in subdistricts with and without the cash transfer program and in subdistricts with and without agricultural productivity shocks.

[16] Comparing other measures of suicides over time is complicated by changing survey questions and changes in the number of villages per subdistrict.

[17] For subdistricts that split up over time, the treatment variable indicates the fraction of subdistricts (based on the 2000 boundary definitions) that receive the treatment.

is at the subdistrict level, we cluster standard errors at the district level, as the roll-out of the program was correlated at the district level. We estimate all of our main specifications with OLS and employ population weights from 2005.[18] Our specification of interest is:

$$y_{st} = \delta_1 \text{Treat}_{st} + \alpha_s + \phi_t + \varepsilon_{st} \tag{2.1}$$

Our main coefficient of interest is δ_1, which provides us with the treatment effect for the subdistricts that had started receiving the program at the time of the data collection. Column 1 of Table 2.2 shows that receiving the cash transfer program of on average 22.45 USD per year reduces the number of suicides per 100,000 inhabitants by 0.36. This corresponds to a reduction by approximately 18 percent relative to the control mean in 2011 and 2014. Our estimates remain economically and statistically significant when (i) clustering errors at the subdistrict level (Column 2), (ii) excluding the sample of subdistricts with randomized treatment assignment employed in Section 2.3.2 (Column 3), including all pre-treatment periods (Column 4) and including subdistrict specific time trends (Column 5). Column 6 displays the treatment effects that give equal weight to the subdistricts regardless of their population size in 2005. This reveals an estimate of -0.591 suicides per 100,000 people, suggesting that the suicide reductions are larger in subdistricts with smaller population sizes.[19]

A key assumption underlying the difference-in-differences approach is that treatment and control subdistricts are on parallel trends. Figure 2.1 provides evidence supportive of the common trend assumption.[20] It displays pre-trends relative to the timing of the introduction of the conditional cash transfer program by treatment wave. Differences in the first period before the treatment are normalized to zero. The displayed coefficients are the difference-in-differences treatment effect estimates using subdistricts that had not received the PKH program until 2013 as counterfactuals. Moreover, as shown above, the effects remain both economically and statistically significant after controlling for subdistrict-specific trends. We find consistently

[18] We use population weights since the welfare relevant metric of interest are changes in the overall Indonesian suicide rate and not the average subdistrict suicide rate. Furthermore, the suicide rate is measured with less error in larger subdistricts which increases the precision of our estimates.

[19] Our results are also robust to using district fixed effects instead of subdistrict fixed effects, controlling for district trends, including controls in a specification without subdistrict fixed effects, and allowing for differential trends by baseline covariates (Table A2.9).

[20] We also provide further results consistent with parallel trends. Figure A2.1 displays the evolution of mean suicide rates over time. Table A2.3 shows that pre-trends are uncorrelated with timing of entry (Column 1) and that the level of the suicide rate in the period before a subdistrict received the program does not drive the treatment effects (Column 2).

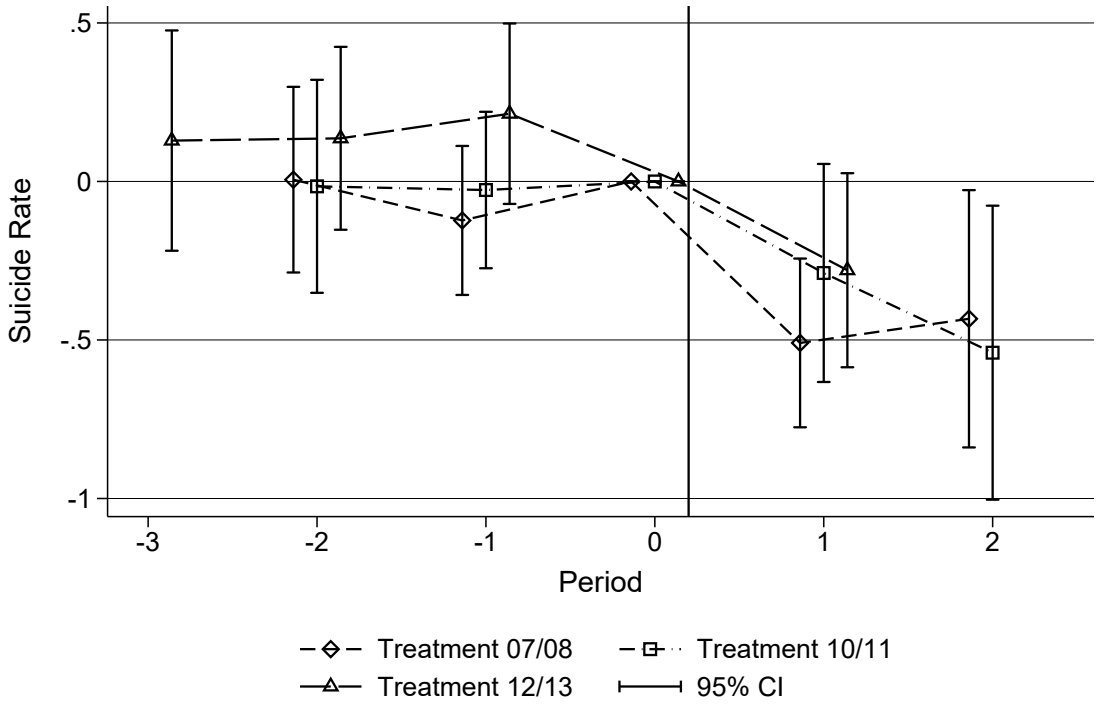
Table 2.2: Main results: roll-out

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable: Suicide rate						
Treatment	-0.358***	-0.358***	-0.381***	-0.366***	-0.267**	-0.591***
	(0.101)	(0.086)	(0.110)	(0.101)	(0.115)	(0.164)
Subdistrict FE	Y	Y	Y	Y	Y	Y
Time FE	Y	Y	Y	Y	Y	Y
Cluster subdistrict	N	Y	N	N	N	N
Exclude RCT data	N	N	Y	N	N	N
Pre-treat data	N	N	N	Y	Y	N
Subdistrict trends	N	N	N	N	Y	N
Population weights	Y	Y	Y	Y	Y	N
Control mean (11&14)	2.016	2.016	2.033	2.016	2.016	2.721
N	9414	9414	8484	15690	15690	9414
Census waves	05-14	05-14	05-14	00-14	00-14	05-14

Table 2.2 displays the difference-in-differences estimate of the effect of the conditional cash transfer program on the suicide rate. All specifications include year and subdistrict fixed effects. Estimates are weighted using the 2005 population size unless otherwise noted. Standard errors are clustered at district level unless otherwise noted. Column 2 reports standard errors clustered at subdistrict level. Column 3 excludes the RCT sample (which we employ in Table 2.3). Column 4 includes all pre-treatment periods. Column 5 further includes subdistrict time-trends. Column 6 does not use 2005 population weights to estimate treatment effects. The definition of the suicide rate per 100,000 people changes slightly over time. In 2014 it is defined as the number of suicides and suicide attempts per 100,000. In 2011 it is defined as the number of suicides per 100,000. From 2000 to 2005 the suicide rate is defined as the number of villages in a given subdistrict with at least one suicide per 100,000.* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

large negative and significant effects across a series of specifications. While we have provided evidence in support of the parallel-trend assumption, there is no formal test of the validity of this identification assumption.

Figure 2.1: Event-study: roll-out of conditional cash transfer program



Notes: Figure 2.1 displays treatment effects relative to he timing of the introduction of the conditional cash transfer program by treatment wave. The difference in the first period before the treatment are normalized to zero. The displayed coefficients are the difference-in-differences treatment effect estimates relative to subdistricts that had not received the treatment until 2013. All standard errors are clustered at the district level. The vertical red line indicates the program start. The definition of the suicide rate per 100,000 people changes slightly over time. In 2014 it is defined as the number of suicides and suicide attempts per 100,000. In 2011 it is defined as the number of suicides per 100,000. From 2000 to 2005 the suicide rate is defined as the number of villages in a given subdistrict with at least one suicide per 100,000. The vertical line indicates the program start.

2.3.2 Randomized Controlled Trial

In this subsection, we use those subdistricts in which the treatment was randomly assigned to test whether the experimental treatment effect estimates are in line with the non-experimental roll-out analysis from the previous section.

Balance As a first step, we test whether the treatment and control group are balanced in terms of observables. Let T_s denote the PKH program's original allocation, where $T_s = 1$

if the subdistrict was randomly assigned to receive the program, and $T_s = 0$ otherwise. We consider whether baseline balance holds for the original treatment assignment by comparing means and clustering standard errors at the subdistrict level. In Table A2.2, we provide evidence of baseline balance on a set of observables. We cannot reject the null hypothesis of global balance ($p = 0.57$).

Results We estimate treatment effects using the randomized assignment of the cash transfer program, clustering standard errors at the subdistrict level and weighting observations by population size in 2005.[21] Despite the randomization of cash transfers at the subdistrict level, we find some evidence that treated subdistricts were on an upward trend compared to control subdistricts (Figure A2.4) and some gap in the suicide rate before the program was launched (Table A2.2). As a result of the common trend violation and the slight baseline imbalance, difference-in-differences estimators might be upward biased and a more conservative way of evaluating treatment effects is to employ an ANCOVA and a post-estimator.

The ANCOVA specification in Column 1 of Table 2.3 shows that subdistricts randomly assigned to receive the same conditional cash transfer program as in Section 2.3.1 have, on average, a 0.337 lower suicide rate (about 19% of the control mean in 2011). Comparing mean suicide rates between treatment and control subdistricts in 2011 also yields an insignificant average decrease of 0.258 suicides per 100,000 (Column 2). While the magnitudes of the effects are economically meaningful and of very similar size to the roll-out estimates, they are statistically insignificant. We attribute this lack of significance to low statistical power.[22]

We also estimate the effect of the RCT using difference-in-differences specifications and find mostly significant negative treatment effects (Columns 3 to 6 of Table 2.3) confirming that the cash transfer program reduced suicides. However, given the pre-trend violation and slight baseline imbalance, the effect sizes of the ANCOVA and post-estimator are more credible. The results of the RCT are in line with the roll-out results. The key difference between the two pieces

[21] We do not use the 2014 census in this section. While contamination of the randomization was quite low in 2011 (with 10 percent of subdistricts having an actual treatment status differing from the randomly assigned one), the contamination of the program strongly increased over time. By 2014, 30 percent of control subdistricts received the program. Moreover, the cash transfer ended in 2012 and 2013 so that only a subset of subdistricts was still receiving the program during the relevant period from PODES 2014. We discuss the long-run effects of the PKH program in Section 2.3.3.

[22] Ex-post power calculations show that we had 80% power to detect a 0.770 effect size at the five percent level with the ANCOVA specification.

Table 2.3: Main results: randomized experiment

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent variable: Suicide rate							
Treatment	-0.337	-0.258	-0.665**	-0.665**	-0.466	-1.064*	-0.474
	(0.266)	(0.275)	(0.318)	(0.266)	(0.334)	(0.593)	(0.325)
Subdistrict FE	N	N	Y	Y	Y	Y	N
Time FE	N	N	Y	Y	Y	Y	N
Cluster district	N	N	N	Y	N	N	N
Pre-treat data	N	N	N	N	Y	Y	N
Subdistrict trends	N	N	N	N	N	Y	N
Baseline suicide	Y	N	N	N	N	N	Y
Population weights	Y	Y	Y	Y	Y	Y	N
Control mean (2011)	1.774	1.774	1.774	1.774	1.774	1.774	2.058
N	310	310	620	620	1240	1240	310
Census waves	11	11	05-11	05-11	00-11	00-11	11

Table 2.3 displays the results of the RCT experiment. The suicide rate is defined as explained in Table 2.2. Column 1 reports an ANCOVA specification using the 2011 suicide rates as the outcomes and controlling for the 2005 suicide rate. Column 2 displays the treatment effect estimate of a post comparison of treated and control subdistricts. Columns 3 to 6 report difference-in-differences estimates of the treatment effect and control for both subdistrict and time fixed effects. Standard errors are clustered at the subdistrict level unless otherwise noted. In Column 3 we report the baseline difference-in-differences specification using data from 2005 and 2011. Column 4 reports standard errors clustered at district level. Column 5 includes data from the 2003 and 2000 census waves. Column 6 further includes subdistrict specific time-trends on top of subdistrict and time fixed effects. Column 7 shows the ANCOVA specification from Column 1 without population weights.* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

of evidence is that the results from the roll-out are much more precisely estimated as they are based on a 10 times larger sample size than the estimates from the RCT.

2.3.3 Dynamics of Treatment Effects

In this subsection, we show that the cash transfer program had persistent treatment effects throughout the six-year duration of the program. To do so, we plot the evolution of the estimated treatment effect of the PKH program on suicide rates over time (Figure 2.2). To estimate the plotted coefficients we exploit the fact that the census data collection occurred at different points in time relative to the beginning of the treatment for different subdistricts. This means that each point estimate is obtained by comparing a different sample of treatment and control subdistricts.[23] Treated subdistricts in a given period t are defined as having received the cash transfer program exactly t before (or after) the census. For $t \geq 0$ control subdistricts are defined

[23]We do observe some treated subdistricts at two different points in time which allows us to directly compare treatment effects over time (Table A2.16). The treatment effects for the subsample of treated subdistricts we observe twice are in line with the results from the overall sample.

as subdistricts that had not received the treatment at the time of the census. For $t < 0$ control subdistricts are defined as not having received the treatment t years after the census. This also leads to differences in sample size and precision in the estimation of treatment effects, as apparent by variation in the width of confidence intervals.

A clear temporal pattern emerges from Figure 2.2. Similar to the roll-out analysis there are no detectable difference between treatment and control subdistricts in the years prior to receiving the cash transfer program. However, starting in the year subdistricts first receive the treatment cash transfer program the suicide rate declines by about -0.3 in line with the aggregate analysis. The difference in the suicide rate persists throughout the six-year duration of the cash transfer program without any obvious changes in the effect size.[24] The treatment effect also persists into year seven after the treatment, potentially suggesting a persistence beyond the receipt of the program. However, the periods are defined in calendar years so that we cannot rule out that persistence is driven by subdistricts who started receiving the program in late 2007 still receiving the program at the time of the census data collection in early 2014.

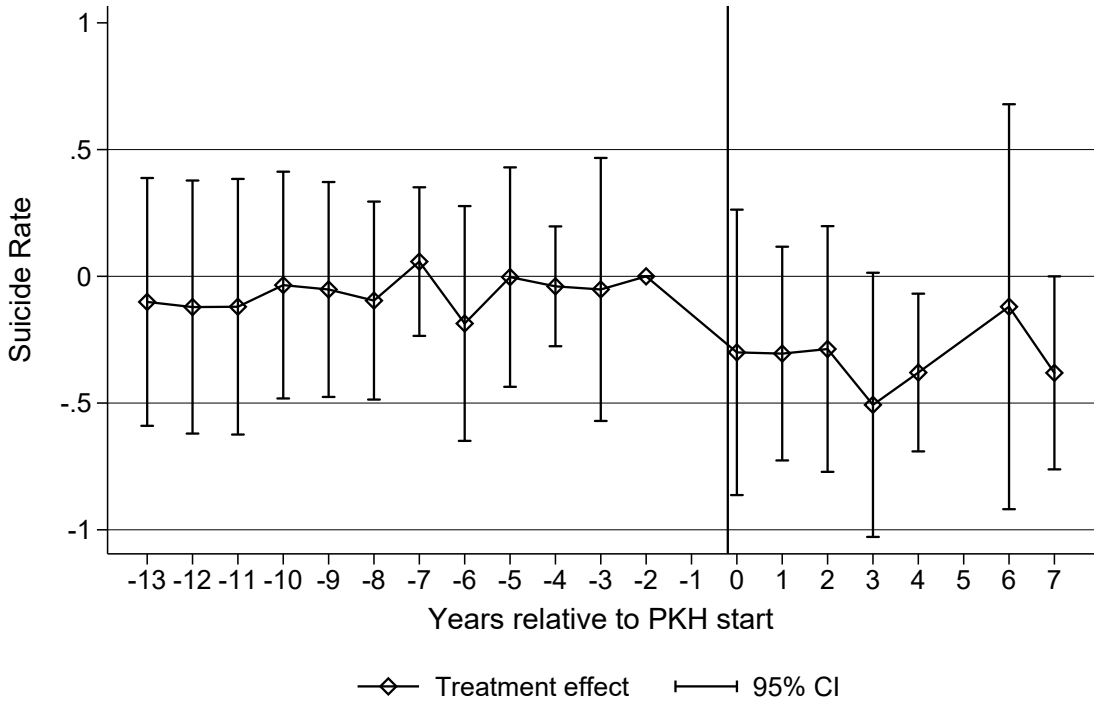
These findings also imply that the duration of exposure to the treatment does not affect the estimated effects of the program. To more formally test this we run a regression in which we include a treatment indicator as well as a variable on the number of years of treatment received by the subdistrict. Duration of receipt is not significantly related to suicide rates, and barely affects the coefficient on whether a given subdistrict received the program (Table A2.18). We find comparable effect sizes for short-run and medium-run effects of the program. This finding is inconsistent with a model of reference dependence over past income levels which predicts declining effects on suicides as soon as people's reference point adjusts to the new income level.

2.3.4 Robustness: Outcome Definitions

One concern with our analysis could be that the definitions of the suicide rate change over time. We provide three pieces of evidence that demonstrate the robustness of our results to using less rich but time-invariant measures of subdistrict suicides.

[24] This finding is in line with evidence of constant and persistent effects of the PKH program on child health, education, and child labor (Cahyadi et al., 2018).

Figure 2.2: Dynamics of treatment effects



Notes: Figure 2.2 displays the treatment effects on the suicide rate relative to the year of introduction of the conditional cash transfer program. The suicide rate is defined as in Figure 2.1. To estimate the coefficient for each year we use different sub-samples of subdistricts. Treated subdistricts in a given period t are defined as having received the cash transfer program exactly t before (or after) the census (e.g. to estimate the effect in year 1, we use observations from the 2011 census subdistricts that received PKH in 2010 and observations of the 2014 census of subdistricts first receiving the PKH in 2013). For $t \geq 0$ control subdistricts are defined as subdistricts that had not received the treatment at the time of census. For $t < 0$ control subdistricts are defined as not having received the treatment t years after the census. Coefficients are obtained conditional on year and subdistrict fixed effects. Standard errors are clustered at the district level. The vertical line indicates the program start. Differences two years before the treatment are normalized to zero.

First, we employ a measure of the suicide rate using the Poisson extrapolation described in Section 2.2.3. Using this extrapolated suicide rate as the outcome for all years, we see the same treatment effect patterns with similar effect sizes (Columns 1 to 5 of Table A2.4 and Figures A2.6 and A2.7). Second, we use a version of the suicide rate based on the number of villages in a given subdistrict that report at least one suicide. This is effectively a truncated version of our main outcome variable. Our treatment effect estimates with this outcome definition remain largely unchanged (Columns 6 to 10 of Table A2.4). Third, we employ a binary variable indicating whether any suicide occurred in a given subdistrict-year. The treatment effects are qualitatively similar to our main specification (Columns 11 to 15 of Table A2.4). Our preferred

specification implies that receiving the cash transfer reduces the likelihood of at least one suicide by 6 percentage points or 12 percent. Thus, it seems that changes in the survey structure over time do not affect our results.

Another concern could be that the cash transfer program affects migration patterns, and thereby shapes our treatment effect estimates. To test whether these potential changes in migration affect our results, we estimate treatment effects on two further outcomes not subject to this bias. The first one is the suicide rate per 100,000 inhabitants using the 2005 population for all years. This definition of the suicide rate is unaffected by changes in migration induced by the cash transfer. We find that with this measure treatment effects on the suicide rate are still highly significant (Columns 1 to 5 of Table A2.5). The second alternative outcome variables is the number of suicides.[25] Again, the treatment effect patterns remain largely unchanged. The point estimate of our preferred specification indicates that the cash transfer program decreased the number of suicides per subdistrict by 0.2 (Column 11 of Table A2.5). This set of results suggests that changes in migration do not drive our main results.[26]

2.3.5 Agricultural Productivity Shocks and Suicides

The evidence from previous sections shows that a positive economic shock, namely the receipt of a conditional cash transfer, can lower suicide rates. In this section, we examine whether agricultural productivity shocks, as measured by rainfall, also affect suicide rates.

Advantages of Rainfall Analysis The rainfall analysis has at least two advantages compared to the analysis of the cash transfer program: First, it enables us to examine whether positive and negative income shocks have symmetric effects on the incidence of suicide. Second, it allows us to retrieve estimates with no concerns regarding differential social desirability bias between the treatment and the control group.[27]

Data Our empirical strategy relies on the following two facts: First, the agricultural sector in Indonesia is to a large extent governed by seasonal monsoon rainfall. Second, Indonesian

[25] For this analysis we stick with the definition used for our main outcome variable and use the actual number of suicides when available. The results remain very similar when we use the number of villages with at least one suicide as the outcome (results available upon request).

[26] The results for the RCT are also robust to changing outcome definitions, though the results are more noisily measured (Tables A2.10 and A2.11).

[27] One may be concerned that village heads whose villages are in receipt of the PKH program report more favorable outcomes, but this critique does not apply to the rainfall analysis.

rainfall exhibits substantial variability within a given year across subdistricts, as well as within subdistricts over time. To examine the causal effect of agricultural productivity shocks, we leverage the ERA-Interim Reanalysis dataset which provides precipitation data from 1979 until 2016 on a 0.25×0.25 degree resolution (roughly a 27.5×27.5 kilometer grid at the equator). We define rainfall at the subdistrict level as the weighted average rainfall at the five grid points closest to the geometric center of the subdistrict.[28] Each grid point is weighted with the inverse of the squared distances to the subdistrict center. Reanalysis data is based on a mix of real weather observations (station and satellite data) and an atmospheric climate model. The main advantage of reanalysis data is the homogeneous data quality across time and space, which alleviates the concern of endogenous placement of weather stations. The rainfall data is matched to the 2000 subdistrict boundaries. We use suicide data from the 2000, 2003, 2005, 2011 and 2014 waves of the Indonesian village census. As before, our main outcome variable of interest, y_{st}, is the suicide rate in a given subdistrict, s at time t.

Specification and Results We follow Maccini and Yang (2009) and calculate rainfall in a particular year by focusing on rainfall in agricultural seasons (rather than in calendar years). As in Maccini and Yang (2009) we define rainfall, $zrain_{st}$, as the normalized deviation of rainfall from the long-term mean within a given subdistrict.[29] This measure of rainfall has been shown to significantly and strongly predict rice output (Levine and Yang, 2014). In all of our specifications we control for subdistrict level fixed effects, α_s, as well as time fixed effects, ϕ_t. As rainfall is heavily spatially correlated, we report Conley (1999) standard errors allowing for arbitrary spatial and temporal correlation of error terms in a 100km radius around the subdistrict center (we also report standard errors clustered at the district level). We estimate the following equation, using the 2005 population to weight the subdistrict observations to make the analysis comparable to the cash transfer estimates:

$$y_{st} = \gamma_1 \text{zrain}_{st} + \alpha_s + \phi_t + \varepsilon_{st} \tag{2.2}$$

Table 2.4 provides evidence that higher rainfall significantly reduces suicides. In Column 1 we show that increases in subdistrict-rainfall by one standard deviation from the long-run subdistrict

[28] We lack coordinates for 14 subdistricts. We use average rainfall and coordinates of other subdistricts in the same district for these subdistricts. For one district all subdistricts have missing coordinates. For those we use province level average rainfall and coordinates.

[29] We use rainfall data from 1979 to 2016 to construct the subdistrict specific leave-one-out long-run means and leave-one-out long-run standard deviations. As Indonesia is located around the equator, temperature is relatively constant and therefore does not have effects on agricultural yields (Kleemans and Magruder, 2018).

Table 2.4: Interactions between agricultural productivity shocks and the cash transfers

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable: Suicide rate						
Rain (z-scored)	-0.082*	-0.112**	-0.081*	-0.106**	-0.094*	-0.129**
	(0.045)	(0.050)	I (0.044)	(0.048)	(0.049)	(0.054)
	[0.039]	[0.043]	[0.038]	[0.042]	[0.039]	[0.046]
Treat		-0.216**		-0.255**		-0.193
		(0.109)		(0.101)		(0.117)
		[0.122]		[0.115]		[0.130]
Rain (z-scored)× Treat		0.118*		0.096		0.133*
		(0.072)		(0.068)		(0.079)
		[0.070]		[0.068]		[0.077]
Subdistrict FE	Y	Y	Y	Y	Y	Y
Time FE	Y	Y	Y	Y	Y	Y
Subdistrict trends	N	N	Y	Y	N	N
Include lagged rainfall	N	N	N	N	Y	Y
N	15690	15690	15690	15690	15690	15690
Census waves	00-14	00-14	00-14	00-14	00-14	00-14

Table 2.4 displays the impact of rainfall on the suicide rate. The suicide rate is defined as explained in Table 2.2. Odd columns report the impact of standardized rainfall on suicides. Even columns include the treatment variable from the conditional cash transfer roll-out and the interaction with rainfall. All specifications include subdistrict and year fixed effects. Columns 3 and 4 control for subdistrict time-trends. Columns 5 and 6 include the first, second and third lag of rainfall. Parentheses report Conley (1999) standard errors, allowing for arbitrary time correlation and two-dimensional spatial correlation within a 100 kilometre radius of the subdistrict centroid (used for stars). Standard errors clustered at the district level are in square brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

mean lowers the suicide rate by 0.08. In Column 3 we include subdistrict trends, to rule out the possibility that differential trends can explain our findings. Our results are virtually unchanged by the inclusion of trends and if anything become somewhat stronger.[30] Column 5 assesses the sensitivity of our estimates by also controlling for the first, second, and third lag of rainfall. This leaves our estimated coefficients largely unaffected. Our estimated coefficients increase further both in economic and statistical significance once we give equal weight to subdistricts, i.e. once we do not weight by population size (see Table A2.19).

Table A2.20 tests whether the relationship between positive and negative rainfall shocks and the suicide rate is approximately symmetric. To do so we augment the above equation by two dummy variables: posshock$_{st}$, taking value one for subdistricts experiencing a positive shock in

[30]We also estimate the impact of a linearly detrended (at the subdistrict level) measure of rainfall on the suicide rate (Tables A2.22 and A2.23). The results are slightly weaker, but qualitatively in line with the results from the main specification.

rainfall (top one third of the standardized rainfall distribution in our sample) and negshock$_{st}$, taking value one for districts experiencing a negative shock in rainfall (one third percent of the standardized rainfall distribution in our sample). Then we estimate the following equation:

$$y_{st} = \beta_1 \text{zrain}_{st} + \beta_2 \text{posshock}_{st} + \beta_3 \text{negshock}_{st} + \alpha_s + \phi_t + \varepsilon_{st} \tag{2.3}$$

We find little evidence of asymmetric responses to shocks. While the absolute value of the point estimates for β_2 are slightly larger than the estimates of β_3, we cannot reject that they are of equal size. In Figure 2.3 we non-parametrically assess the relationship between rainfall and suicide. To do so, we partial out time fixed effects, subdistrict fixed effects, and subdistrict-trends from both suicide rates and the rainfall measure. Then we use the predicted residuals from these regressions to run local polynomial regressions between these residuals. Figure 2.3 highlights a strong negative relationship between rainfall and the suicide rate, confirming our previous result that the responses to positive and negative rainfall shocks are fairly symmetric, and that the overall relationship is approximately linear. However, at the top end of the rainfall distribution we observe a slight flattening of the relationship consistent with a convex relationship between the suicide rate and income.

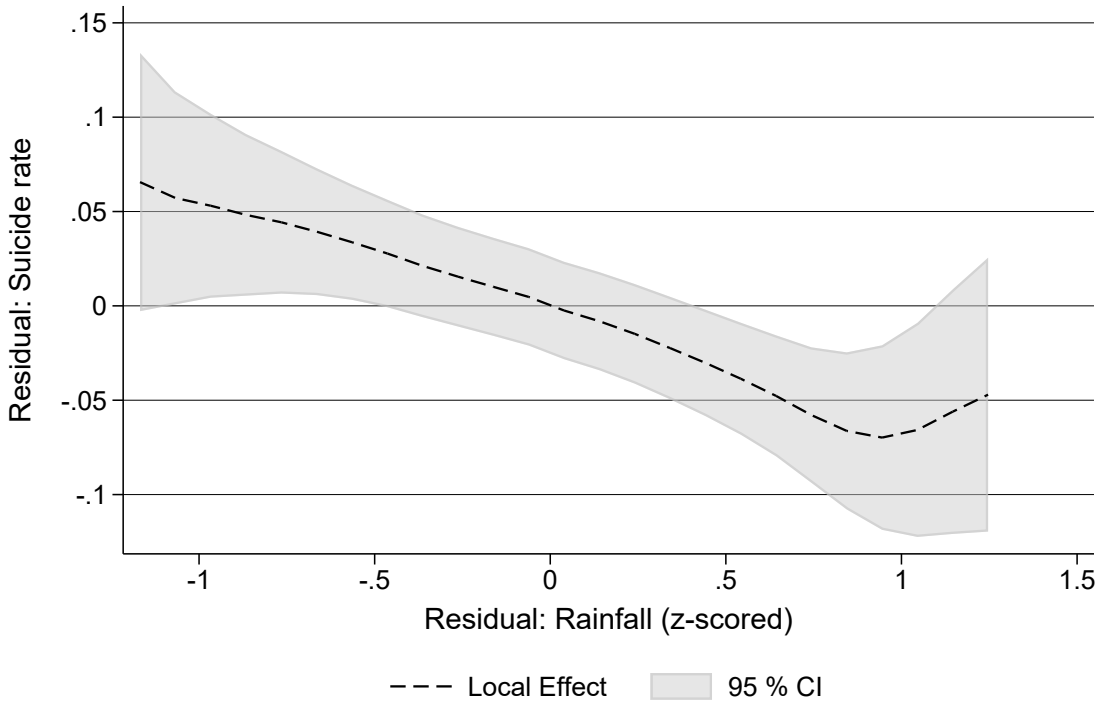
2.3.6 Cash Transfers and Agricultural Productivity Shocks

Do the cash transfers mitigate the adverse effects of agricultural productivity shocks on suicides? If economic hardship caused by negative economic shocks affects suicide rates, then we would expect cash transfers to more strongly reduce suicide rates in the face of negative economic shocks. We test whether there are significant interactions between receiving the cash transfer, Treat$_{st}$, and rainfall, zrain$_{st}$. We leverage the suicide data from 2000 to 2014, and estimate the following equation, reporting Conley standard errors and standard errors clustered at the district level (as in all other main specifications we use population size from 2005 to weight the observations):

$$y_{st} = \gamma_1 \text{zrain}_{st} + \gamma_2 \text{Treat}_{st} + \gamma_3 \text{zrain}_{st} \times \text{Treat}_{st} + \alpha_s + \phi_t + \varepsilon_{st} \tag{2.4}$$

Our key coefficient of interest is γ_3. Column 2 of Table 2.4 reveals that there is a significant interaction effect, consistent with the idea that cash transfers are more (less) effective at lowering suicides in years with lower (higher) agricultural productivity. Our estimates imply that cash

Figure 2.3: Agricultural productivity shocks and suicides: Local Polynomial Regression



Notes: Figure 2.3 displays a local polynomial regression of the residuals of the suicide rate and the residual of standardized rainfall using PODES 2000, 2003, 2005, 2011 and 2014. Residuals are obtained from a regression with time and subdistrict fixed effects and subdistrict time-trends. We employ population weights from 2005 in our estimations. We employ an Epanechnikov kernel with a bandwidth of 0.44. Top and bottom percentile of the residual rainfall distribution are omitted, because of concerns about biased estimates close to the limits (Li and Racine, 2006). The suicide rate is defined as described in Figure 2.1.

transfers lower suicides by 0.3 suicides per 100,000 inhabitants in a year with subdistrict rainfall one standard deviation below its long-run mean, but only lower suicides by 0.1 suicides per 100,000 inhabitants in a year with subdistrict rainfall one standard deviation above its long-run mean.

These effects become slightly smaller and statistically insignificant, but remain economically meaningful, once we control for subdistrict trends. This most likely reflects the limited power for studying interaction effects after controlling for subdistrict trends. The estimated effects of the interaction between the cash transfer and rainfall are much stronger in specifications that do not use population weighting (see Table A2.19). This stems from the fact that rainfall more strongly impacts incomes in more rural subdistricts with lower population size. We also ana-

lyze the effect of the cash transfer separately for positive and negative rainfall shocks in Table A2.21. While this analysis is limited by lower statistical power, the estimates suggest that the cash transfer program reduces the effects of both positive and negative agricultural productivity shocks symmetrically.

The observed heterogeneity in treatment effects suggests that social welfare programs can dampen the effects of negative and positive economic shocks. Put differently, the cash transfer program breaks the relationship between agricultural productivity shocks and suicides.[31] The finding of a positive interaction effect between rainfall and the receipt of the cash transfer suggests that the relationship between the suicide rate and income is convex and not linear or concave.

Further Heterogeneity Do other covariates predict heterogeneous responses to the cash transfer program? There is no statistically significant heterogeneity by any predetermined characteristics (fraction of farmers, fraction poor, per capita expenditure, per capita crimes, per capita social institutions, and per capita health institutions; see Table A2.27). However, our effective statistical power to detect statistical differences in treatment effects across groups is quite limited (see the minimum detectable effect sizes in Table A2.27).

2.4 Mechanism

In the next section we provide suggestive evidence in favor of depression as a channel through which economic circumstances could affect people's inclination to commit suicide. In particular, we show that economic shocks directly affect a measure of depression in line with the framework in Section 2.2.1. Finally, we examine the importance of several potential subdistrict-level mediators.

Economic Shocks and Depression: Micro-Evidence To provide evidence that economic circumstances may affect people's inclination to commit suicide through changes in mental health, we use unique individual-level data on depression from the Indonesian Family Life Survey (IFLS). The IFLS waves 4 and 5 (in 2007 and 2014) administer a ten question version of the CES-D depression scale (Radloff, 1997). Moreover, we use rich data on household expenditure available

[31] This interaction results also suggest that the effects of the cash transfer are not driven by differential social desirability bias between the treatment and the control group.

for all five waves of the IFLS (the sample is restricted to individuals observed in all 5 IFLS waves).

As in Section 2.3.5, we exploit agricultural productivity shocks to study the effects of economic circumstances. We assess the effects of rainfall, zrain$_{st}$, on depression, depression$_{ist}$, as measured by the CES-D score and monthly per capita household expenditure, exp$_{ist}$, in levels and logs.

Our object of interest are households with at least one agricultural worker, the group of households whose income is most strongly dependent on rainfall.[32] We employ the same subdistrict-level rainfall measure as in Section 2.3.5, and also report both Conley standard errors and standard errors clustered at the district level. We also include individual level fixed effects, α_i, in our regression which allows us to control for time-invariant individual-specific unobservables. Specifically, we estimate the following equation:

$$\mathsf{y}_{ist} = \delta_1 \mathsf{zrain}_{st} + \alpha_i + \phi_t + \varepsilon_{ist} \tag{2.5}$$

Table 2.5 shows that a one standard deviation increase in rainfall increases monthly per capita consumption of farmers by 40,000 Rupiah (Column 1), monthly log consumption by 6.9 percent (Column 2)[33], and decreases depression by 0.12 standard deviations (Column 3).[34] Columns 4 to 6 of Table 2.5 also provide evidence that the rainfall shocks operate through an economic mechanism by showing that both depression and consumption of non-farmers do not respond to rainfall shocks. This suggests that our estimated effects on depression do not operate through direct effects of weather on mental health. Indeed, the coefficients on the effects of rainfall on log consumption and mental health are statistically different between farmers and non-farmers ($p = 0.006$ and $p = 0.06$, respectively).[35] We also use the individual level data to study heterogeneous effects of agricultural productivity shocks on depression. We find that depression scores of individuals with higher baseline expenditure are more strongly affected by rainfall shocks (Table A2.24). This suggests a convex relationship between depression and income and is consistent with the finding that cash transfers more strongly reduce suicides in the presence of negative

[32] Households count as "working in agriculture" if any household member works in agriculture either in self-employment (without permanent employees), or as a casual or family worker in any of the five IFLS waves. There is no sector information for IFLS Wave 1. We therefore use working "in self-employment without permanent employees" or working as "temporary worker" as proxies.

[33] A 6.9 percent increase over the median per-capita consumption across all years corresponds to ca. 18,000 Rupiahs or roughly 1.8 USD at 2005 prices.

[34] We also find that individuals who received the PKH cash transfer by 2014 had higher depression scores in 2007 (0.14 standard deviation; see Table A2.25).

[35] The difference in effects on consumption levels are marginally insignificant ($p = 0.162$).

agricultural productivity shocks. The effects are also significantly larger for women and below median age individuals.

Table 2.5: Impact of rainfall on consumption and depression z-scores

	Working in agriculture			Not working in agriculture		
	(1) Cons.	(2) Log cons.	(3) Depression (z)	(4) Cons.	(5) Log cons.	(6) Depression (z)
Rain (z-scored)	39706.205***	0.069***	-0.124***	11687.256	0.008	-0.008
	(14587.567)	(0.022)	(0.042)	(15481.365)	(0.018)	(0.050)
	[9805.642]	[0.019]	[0.042]	[14631.741]	[0.017]	[0.053]
N	12790	12790	5192	8356	8356	3420
IFLS waves used	1-5	1-5	4-5	1-5	1-5	4-5

Table 2.5 displays the impact of rainfall on monthly per-capita household consumption and depression z-scores in the Indonesian Family Life Survey (IFLS). The sample is restricted to individuals tracked in all five waves of the IFLS. Columns 1 to 3 use a sample of individuals living in households with at least one member working in agriculture (self-employed without or with temporary workers, or as casual or family workers) in any of the five waves. Columns 4 to 6 use individuals from households without any agricultural worker. Consumption is measured in 2005 Indonesian Rupiah. Depression scores are measured using the 10 item CES-D scale (Radloff, 1997). All columns include individual and time fixed effects. Parentheses report Conley (1999) standard errors, allowing for arbitrary time correlation and spatial correlation within a 100 kilometre radius of the subdistrict centroid (used for stars). Standard errors clustered at the district level are in square brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Mediation Analysis What other factors could account for the effects of the cash transfer program on suicide rates? The cash transfer increased recipients' welfare by increasing their consumption and improving their health outcomes (The World Bank, 2011). Guided by this, we examine several potential subdistrict-level mediators including local crime rates, health, and education institutions, and social organizations through which the cash transfer program could lower the incidence of suicides. We therefore include time-varying endogenous controls at the subdistrict level in our main specification of interest. These controls could have been affected by the cash transfer in systematic ways and therefore act as a channel through which our treatment effects operate. However, we find little evidence that any subdistrict level institutions mediate our results. The treatment effect estimates hardly move when the potential mediators are included (Table A2.26). This mediation analysis is limited by the fact that we have to rely on subdistrict level mediators, and points to the importance of individual level mediators.

2.5 Interpreting Effect Sizes

The impact of cash transfers on suicides of cash transfer recipients is quantitatively very large if we assume no spillovers in suicides. Spillovers similar to those identified in previous work are consistent with moderately large direct effects. Implied direct per-dollar effects of income on suicides identified using rainfall are significantly smaller than those from cash transfers.

Cash Transfer Program There are two main factors to consider when calculating the size of the cash transfer program on recipients. First, there is a strong economic gradient in the suicide rate. The correlation between the fraction of individuals classified as poor and the suicide rate is large and positive (see Section 2.B.1). A 10% higher share of the poor population is, on average, associated with a 0.14 higher suicide rate in 2005 (Table A2.1). A back-of-the-envelope calculation suggests that poor individuals are 2.24 times more likely to commit suicide than non-poor households.[36]

Second, there are likely two types of spillover effects on non-treated individuals. The first is that suicides are highly contagious (Hedström et al., 2008). Moreover, the conditional cash transfer might have positive economic spill-overs to households not receiving the PKH program (Angelucci and De Giorgi, 2009). Thus, we would expect that reducing suicides among PKH recipients may also decrease suicides among non-recipients. Assuming no spillovers, our preferred treatment effect estimate (0.36 suicides per 100,000 people) implies that the suicide rate among

[36] For details for this and the following back of the envelope calculations see Section 2.A.

the poor decreased by 89% of the implied mean suicide rate or 3.6 suicides per 100,000. This effect size would imply that the suicide rate among PKH recipients who received the transfer is lower than the rate among non-recipients post-treatment. Therefore, it may be more reasonable to calculate the implied suicide rate reduction among PKH recipients assuming that the cash transfer equalized suicide rates between poor and non-poor households. This yields an estimate of a direct treatment effect of 58.7% of the implied control group mean for PKH recipients or 2.36 per 100,000 and an indirect effect of 0.12 for non-poor individuals (4.9% of the direct treatment effect). Our estimated effect size of spillovers is modest compared to estimates from the literature on suicide contagion (e.g. Hedström et al., 2008). We estimate that the program prevented about 1065 suicides, i.e. one suicide for every 672,000 USD spent.

Agricultural Productivity Shocks To estimate the per-dollar effect size of agricultural productivity shocks we use micro data from the Indonesian Family Life Survey. We show that a one standard deviation change in rainfall induces a 6.9% change in monthly per capita consumption for individuals in agricultural households (Table 2.5). This translates into a change of about 1.8 USD of monthly per capita consumption at the median consumption level. This is equivalent to a yearly change in per capita expenditure of 21.6 USD at 2005 prices per one standard deviation change, a similar magnitude as the annual cash transfer amount. We use this measure as our preferred dollar value to deal with outliers in the consumption data which might strongly influence the level estimates. Assuming no spillovers to non-farmers our estimates suggest that an increase in annual per capita expenditure of 21.6 USD decreases the suicide rate by 7.4% of the implied mean suicide rate for farmers across all periods or 0.14 per 100,000 farmers.

Effect Size Comparison Next, we compare the per-dollar impact of the cash transfer program and of the agricultural productivity shocks. This comparison requires strong assumptions, such as the baseline suicide rates of different groups of people. The implied effect sizes in terms of percentage changes in the suicide rate for a 10 USD change in income or consumption (assuming no spillovers) for the cash transfer program are roughly twelve times larger than of those of rainfall shocks (twenty-five times larger in absolute terms of the suicide rate). Extrapolating linearly, we find that an increase in the annual per capita income or consumption by 10 USD decreased the suicide rate of poor individuals and farmers by 39.64% (1.61 suicides) or 3.34% (0.06 suicides), respectively. We can reject the hypothesis of equality of the implied direct effect of the cash transfer and the rainfall shocks in dollar terms ($p<0.01$). This finding implies that

the income shock targeted at initially poorer individuals was more effective, and is therefore consistent with a convex relationship between the suicide rate and income.

The larger per-dollar impact of the cash transfer program compared to agricultural productivity shocks of the same value is in line with the predictions outlined in Section 2.2.1. There are several reasons why the cash transfer could have larger effects. First, the six year duration of the program substantially decreases uncertainty about future income streams. Second, it is possible that program recipients expected that the cash transfer may continue after the initial six year period, which would substantially increase the expected net present value of the program. Third, the conditionalities of the cash transfer may have an additional effect on mental health as they may induce more social interactions and may relieve stress related to the children. Fourth, individuals classified as poor have higher implied mean suicide rates than farmers. Thus, the cash transfer program targeted a high risk population and, therefore, likely had a larger impact. Fifth, we assume that all farmers are affected equally by rainfall, but there is a large degree of heterogeneity in how rainfall affects harvest depending, for example, on the type of crop. Sixth, it is possible that the cash transfers have a direct psychological effect. They may also act as a signal that the government may be willing to offer insurance from bad outcomes more generally, thereby shifting recipients' economic outlook.

2.6 Conclusion

We establish an important economic dimension in suicides. Using the nationwide roll-out of a conditional cash transfer program and a randomized experiment of the same program, we show that the program decreased suicides by approximately 0.36 per 100,000 inhabitants. We also show that agricultural productivity shocks, proxied by rainfall, significantly affect suicide rates. A one standard deviation increase in rainfall lowers the number of suicides per 100,000 inhabitants by approximately 0.08. Moreover, we establish that cash transfers lower suicides most effectively in subdistricts experiencing negative agricultural productivity shocks. This supports the idea that social welfare programs can mitigate the adverse effects of negative economic shocks on mental health. Our evidence points to an important role of government policies in alleviating the consequences of poverty on mental health.

Using micro-data from the Indonesian Family Life Panel we show that economic shocks may affect people's inclination to commit suicide through mental health. However, understanding the

exact mechanisms through which economic shocks ignite suicidal behavior leaves ample scope for future research. We believe that there are several fruitful avenues for future research. First, we need better micro-data on how economic circumstances affect mental health, for example through elevated stress levels or the formation of economic beliefs and preferences. Second, more research should be carried out to understand which combination of economic and psychological interventions are best-suited to increase mental health and prevent suicides. Third, we need a better understanding of which populations should be targeted to increase mental health and lower suicide rates most cost-effectively.

Acknowledgements I thank the Review of Economics and Statistics co-editor, Rohini Pande, and four anonymous referees for very useful suggestions. I would like to also thank Damian Clarke, Jon de Quidt, Esther Duflo, Jan Emanuelle De Neve, James Fenske, Thomas Graeber, Alexis Grigorieff, Ingar Haaland, Johannes Haushofer, Johannes Hermle, Matt Lowe, Climent Quintana-Domeque, Simon Quinn, Timo Reinelt, Marc Witte, and Johannes Wohlfart for comments and help. I thank Jessica Topping for excellent research assistance. We also thank various conference and seminar participants for comments on the paper. I am grateful to Sudarno Sumarto, Jan Priebe, Matthew Gudgeon, the National Team for the Acceleration of Poverty Reduction (TNP2K) and the World Bank for sharing several sources of data.

Appendices of Chapter 2

Summary of the Appendix

The Appendix starts with detailing the calculations used to obtain the direct effect sizes discussed in Section 2.5 of the paper.

In Section 2.B.1 we present descriptive statistics. Table A2.1 shows correlates of suicide rates. Table A2.2 shows evidence in favor of the integrity of the randomization. Section 2.B.2 examines the sensitivity of our results from the difference-in-differences strategy.

Table A2.3 provides further checks on the assumption that the timing of roll-out of the program was quasi-random. Tables A2.4 and A2.5 examine robustness of the roll-out results to using different definitions of the outcome variables. Table A2.6 examine sensitivity of our results to dropping partially treated subdistricts. Table A2.7 shows the main roll-out results without using population weights. Table A2.8 shows the results using 2006 subdistrict definitions. Table A2.9 shows further robustness checks.

Section 2.B.3 examines the sensitivity of the results from the randomized cash transfer program. Tables A2.10 and A2.11 examine robustness of the RCT results to using different definitions of the outcome variables. Table A2.12 examines sensitivity of our results to dropping partially treated subdistricts. Table A2.13 shows the RCT results without employing population weights. Table A2.14 shows the results using 2006 subdistrict definitions. Finally, Table A2.15 shows further robustness checks for the RCT results. Section 2.B.4 presents further analysis of the dynamics of treatment effects.

Table A2.16 present the results on "dynamic treatment effects" which using population weights. Table A2.17 present the results on "dynamic treatment effects" which does not apply population weights. Table A2.18 examines whether treatment effects vary with treatment intensity.

Section 2.B.5 provides further robustness checks for the rainfall analysis. Table A2.19 presents the main specifications without using population weights. Table A2.20 examines whether rainfall and suicide rates are symmetrically related. Table A2.21 examines whether the interaction of rainfall and conditional cash transfers are non-linearly related. Table A2.22 shows the main specification with rainfall detrended at the subdistrict level using population weights. Table A2.23 shows the main specification with rainfall detrended at the subdistrict level without using population weights.

Section 2.B.6 presents micro-evidence using the Indonesian Family Life Survey. Table A2.24 analyzes heterogeneous effect of rainfall shocks on depression symptoms by gender, age, and baseline expenditures. Table A2.25 examines whether PKH recipients have higher depression scores before the receipt of the cash transfer program.

Section 2.B.7 sheds further light on mechanisms. Table A2.26 examines whether our results are mediated by changes in crime rates, the quality of local institutions or social capital. Table A2.27 describes heterogeneous responses to the treatment.

Section 2.C presents additional figures. Figure A2.1 displays the evolution of mean suicide rates over time for different treatment groups. Figure A2.2 displays mean suicide rates over time by treatment wave without applying population weights. Figure A2.3 shows the event-study analysis without applying population weights. Figure A2.4 presents an event-study analysis for our RCT sample with effects normalized relative to 2005. Figure A2.4 presents an event-study analysis for our RCT sample with effects normalized relative to 2000. Figure A2.4 presents an event-study analysis for our RCT sample. Figure A2.6 displays mean suicide rates (based on a Poisson extrapolation) over time by treatment wave. Finally, Figure A2.7 showcases the event-study graph using the Poisson extrapolated suicide rates.

2.A Effect Size Calculations

This section explains how we calculate the implied direct effect on cash transfer recipients and farmers. This section is not meant to provide unambiguous estimates of direct effects, but rather to provide a back of the envelope calculation based on several simplifying assumptions. We conduct this analysis in the following steps for the cash transfer program and agricultural productivity shocks.

1. Calculate the economic gradient of suicide rates.
2. Calculate the implied direct treatment effect on treated individuals.
3. Account for potential spillover effects of changing suicide rates.

The Cash Transfer Program

First, we calculate the economic gradient of the suicide rate with respect to being classified as poor. We use this specification as a proxy for households eligible to receive PKH.[37] Column 1 of Table A2.1 suggests that a 10% higher share of the eligible household is, on average, associated with a 0.142 higher suicide rate 2005 before the program was rolled out. Assuming that this linear relationship between the fraction of poor individuals and the suicide rate holds we can use the constant as the suicide rate for non-eligible individuals ($\bar{s}_{\text{non-elig}} = 1.15$). The relationship between the share of eligible individuals ($x_{\text{elig}} = 0.1$) and the suicide rate is:

$$1.15 + 1.42 \cdot x_{\text{elig}} = (1 - x_{\text{elig}}) \cdot 1.15 + x_{\text{elig}} \cdot \bar{s}_{\text{elig}}$$

From this we obtain the implied suicide rate for poor individuals $\bar{s}_{\text{elig}} = 2.57$. Thus, we conclude that eligible individuals are 2.24 times more likely to commit suicide than non-eligible individuals.

Second, we calculate the implied direct effect on eligible individuals assume that this ratio of suicide rates for eligible and non-eligible individuals relationship remains the same for 2011 and 2014 for non-treated subdistricts. As a first step, we calculate the implied suicide rates for eligible and non-eligible individuals in 2011 and 2014. The control mean ($\bar{s}_{\text{control}} = 2.02$) is equal to a weighted average of eligible and non-eligible individuals:

$$\bar{s}_{\text{control}} = 2.02 = 0.9 \cdot s_{\text{non-elig}} + 0.1 \cdot \underbrace{2.24 \cdot s_{\text{non-elig}}}_{s_{\text{elig}}} \tag{2.6}$$

[37] We abstract from the fact that 15% of the population is classified as poor and assuming that 10% of the population is poor and that only these 10% are eligible for treatment. If anything this would make us underestimate the economic gradient between PKH recipient and non-recipients.

From this we obtain $s_{\text{non-elig}} = 1.80$ and $s_{\text{elig}} = 4.03$. Next, we use the calculated suicide rate in treated subdistricts $\bar{s}_{\text{treat}} = 2.02 - 0.36 = 1.66$ to calculate the implied direct treatment effect on eligible individuals. Assuming no spillovers on untreated individuals we obtain:

$$0.9 \cdot 1.8 + 0.1 \cdot (4.03 + \Delta s_{\text{elig}}) = 1.66 \tag{2.7}$$

.

This yields a direct treatment effect on PKH recipients as $\Delta s_{\text{elig}} = -3.6$ or a reduction of about 89% of the implied control suicide rate for eligible individuals.

These numbers imply that the suicide rate of eligible individuals receiving the cash transfer would be reduced to levels below the suicide rate of non-eligible households. Therefore, we conduct a further calculation allowing for the cash transfer program to have positive spillovers on non-eligible individuals, in line with evidence on cash transfers and suicides (Angelucci and De Giorgi, 2009; Hedström et al., 2008). Our benchmark specification calibrates the strength of within subdistrict spill-overs such that the ex-post suicide rates of eligible and non-eligible individuals in treated subdistricts is equalized.

We model the total treatment effect as consisting of a multiplier effect and a direct treatment effect. This requires an assumption about the functional form of the multiplier. We assume that the suicide rate of non-PKH recipients in treated subdistricts is a function of the direct treatment effect times a multiplier μ. The suicide rate in treated subdistricts can then be written as:

$$\bar{s}_{\text{treat}} = 0.9 \cdot (s_{\text{contr,non-elig}} + \underbrace{\mu(s_{\text{treat,elig}} - s_{\text{contr,elig}})}_{\text{Spill-overs}}) + 0.1 \cdot s_{\text{treat,elig}} \tag{2.8}$$

From this we can obtain the implied direct treatment effect on PKH recipients as:

$$\Delta s_{\text{elig}} = s_{\text{treat,elig}} - s_{\text{contr,elig}} = \frac{\bar{s}_{\text{treat}} - 0.9 \cdot s_{\text{contr, non-elig}} - 0.1 \cdot s_{\text{contr,elig}}}{0.9 \cdot \mu + 0.1} \tag{2.9}$$

Equation 2.9 confirms the intuition that the implied direct effect on poor individuals decreases with size of the multiplier μ. Setting ex-post suicide rates for poor and non-eligible individuals equal at the mean suicide rate in treated subdistricts of 1.66 suicides per 100,000, we obtain a multiplier of $\mu = 0.049$ and an implied direct effect on PKH recipients of 2.36 per 100,000 or 58.7% of the implied control group mean.

Effect on number of suicides Our preferred estimate for the number of suicides is Column 11 of Table A2.5 which shows an average treatment effect of 0.2 suicides. In total, there we observe 5324 subdistrict-years for treated subdistricts over time. Assuming constant treatment effects over time, we calculate that the PKH program prevented 1065 suicides.

Rainfall

To calculate the effect size of agricultural productivity shocks, we conduct a similar calculation.

First, we obtain the economic gradient of the suicide rate with the share of farmers using the results from Column 2 of Table A2.1. We observe that in 2005 an increase in the share of farmer by 10% is associated with an increase in the suicide rate of 1.19. The population weighted average share of farmers in 2005 is $x_{\text{farm}} = 0.56$.[38]

We again use the constant as the suicide rate for non-farmers ($\bar{s}_{\text{non-farm}} = 0.72$). The relationship between the share of farmers x_{farm} and the suicide rate is:

$$0.72 + 1.19 \cdot x_{\text{farm}} = (1 - x_{\text{farm}}) \cdot 0.71 + x_{\text{farm}} \cdot \bar{s}_{\text{farm}}$$

From this we obtain the implied suicide rate for farmers $\bar{s}_{\text{farm}} = 1.91$. Thus, farmers are 2.67 time more likely to commit suicide than non-farmers households. The calculations for the effect of rainfall are complicated by the fact that there is no clear control group. Therefore, we use the population weighted mean suicide rate across all years (1.39 suicides per 100,000). Assuming no spill-over effects, we then calculate the implied suicide rates for farmers and non-farmers using the following equation:

$$\bar{s} = 1.39 = 0.44 \cdot s_{\text{contr,non-farm}} + 0.56 \cdot \underbrace{2.67 \cdot s_{\text{contr, non-farm}}}_{= s_{\text{contr, farm}}}$$

We obtain mean suicide rates of $s_{\text{contr, farm}} = 1.91$ for farmers and $s_{\text{contr, non-farm}} = 0.72$ for non-farmers

Based on these implied means we can then calculate the implied direct effect size assuming no spill-overs. The average effect of a one-standard deviation in rainfall according to our preferred specification is -0.08. We obtain $\bar{s}_{\text{treat}} = 1.31$ by subtracting a one standard deviation treatment effect (0.08) from the population weighted mean suicide rate across all years (1.39). We then use an analogue equation to the cash transfer calculations to compute the implied direct effect:

[38] We use the 2005 share to make the calculation comparable to the cash transfer calculations.

$$\bar{s}_{\text{treat}} = 1.31 = 0.44 \cdot s_{\text{contr,non-farm}} + 0.56 \cdot (s_{\text{contr, farm}} + \Delta s_{\text{farm}})$$

With these assumptions, we obtain an implied direct effect size as $\Delta s_{\text{farm}} = -0.14$ which is equivalent to 7.4 percent of the implied long-term mean for farmers. Allowing for spillover with a multiplier in line with the cash transfer (0.049), the implied direct treatment effect decreases to -0.047 suicides per 100,000 or 2.5 percent of the implied long-term mean for farmers.

Calculating per-dollar impacts for implied direct effects

This section outlines how we calculate comparable monetary values for implied direct effect of the cash transfer and agricultural productivity shocks. For this exercise, we need to assume that the impact is linear in dollar amounts. For the cash transfer, we estimate that 22.45 USD lower the suicide rate by 3.6 suicides per 100,000 or 89 percent, so that 10 USD would lower the suicide rate by 1.61 suicides per 100,000 or 39.64 percent. For agricultural productivity shocks, we estimate that a yearly per-capita consumption increase by 21.6 USD reduces the suicide rate by 0.14 suicides per 100,000 or 7.4 percent. This implies that 10 USD would reduce the suicide rate by 0.06 suicides per 100,000 or 3.34 percent.

2.B Additional Tables

2.B.1 Descriptives

Table A2.1: Correlates of the suicide rate

	(1)	(2)	(3)	(4)	(5)
Dependent variable: Suicide rate					
Fraction poor	1.422**			-0.488	-0.155
	(0.591)			(0.721)	(0.684)
Fraction farmers		1.194***		1.109***	0.692***
		(0.129)		(0.168)	(0.204)
Average HH expenditure (per capita)			-0.273***	-0.075	-0.063
			(0.051)	(0.067)	(0.066)
Education (per capita)					-0.000
					(0.001)
Health facilities (per capita)					0.004***
					(0.001)
Social organisations (per capita)					0.000
					(0.000)
Crime (per capita)					0.013***
					(0.003)
Constant	1.147***	0.715***	2.035***	1.024***	0.120
	(0.114)	(0.070)	(0.144)	(0.294)	(0.348)
R2	0.006	0.030	0.017	0.030	0.066
N	3138	3138	3138	3138	3138
Census wave	05	05	05	05	05

Table A2.1 shows correlates of the suicide rate per 100,000 individuals in 2005 with district and subdistrict covariates. Our main outcome is the suicide rate per 100,000 individuals. Due to data constraints its definition varies slightly over time. In 2014 it is defined as the number of suicides and suicide attempts per 100,000. In 2011 it the number of suicides per 100,000. From 2000 to 2005 the suicide rate is defined as the number of villages in a given subdistrict with at least one suicide per 100,000 population. The fraction of individuals below the poverty line (poor) and the average per-capita household expenditure are measured at the district level. All other variables are measured at the village level and aggregated to the subdistrict level. Share of population classified as poor or farmers ranges from 0 to 1. Column 5 further includes per capita educational institutions, per capita health institutions, per capita social organizations, and crime rates. Regressions weighted by subdistrict population in 2005. Standard errors clustered at district level in parentheses. $^{*}\ p < 0.10$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$

Table A2.2: Baseline balance: randomized experiment

	Treatment	Control	Δ	se(Δ)	p(Δ=0)
Suicide rate	1.719	1.337	0.386	(0.282)	0.173
Any suicide	0.468	0.427	0.040	(0.059)	0.504
Education institutions per capita	91.763	91.713	-0.085	(5.338)	0.987
Health institutions per capita	100.784	100.573	0.401	(5.998)	0.947
% villages with asphalted road	0.651	0.662	-0.013	(0.037)	0.731
% villages with lighting	0.786	0.732	0.051	(0.040)	0.207
% rural villages	0.950	0.925	0.026	(0.026)	0.323
Number of villages	16.144	15.979	0.219	(0.783)	0.779
Population size	55811.295	54974.091	760.818	(3669.322)	0.836
Number of families	14758.525	14510.671	231.873	(927.610)	0.803
Percentage of farmers	68.063	68.620	-0.498	(2.524)	0.844
N	167	143	310	310	310

Table A2.2 compares population weighted baseline covariates of the RCT sample using the 2005 village census. The suicide rate is defined as described in Table A2.1. P-values are based on standard errors clustered at the subdistrict level. Test for joint significance of differences between the treatment and control group is not rejected (p=0.57). $^{*}\ p < 0.10$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$

2.B.2 Further sensitivity - Roll-Out of Cash Transfer Program

Table A2.3: Roll-out identification: regression tests

	(1) Δ Suicide rate	(2) Suicide rate
Treatment wave 1: 07-11	-0.002	
	(0.063)	
Treatment wave 2: 12-13	0.028	
	(0.064)	
Treatment		-0.399***
		(0.121)
Include pre treatment dummy	N	Y
N	6276	15690
Census waves	00-05	00-14

Table A2.3 displays two tests in support of the identification assumptions for the main roll-out specification. Column 1 presents a test for whether the timing of the entry into the program correlates with pre-trends. The dependent variable is changes in the suicide rate between rounds. The sample in Column 1 is restricted to pre-treatment periods. Column 2 displays a test for whether the treatment effect is driven by changes in the suicide rates right before the introduction of the treatment. It shows the treatment effect on the suicide rate controlling for an indicator variable for the period prior to treatment. The suicide rate is defined as described in Table A2.1. All specifications include time and subdistrict fixed effects and are weighted by 2005 population. Standard errors are clustered at the district level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A2.4: Roll-out: additional outcome definitions 1

	Suicide rate (extrapolated - Poisson)					Suicide rate (truncated)					Any Suicide				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Treatment	-0.298***	-0.474***	-0.313***	-0.284***	-0.240**	-0.268***	-0.418***	-0.282***	-0.265***	-0.206**	-0.061***	-0.061***	-0.062***	-0.046**	-0.068**
	(0.096)	(0.154)	(0.103)	(0.087)	(0.113)	(0.081)	(0.132)	(0.088)	(0.076)	(0.097)	(0.022)	(0.022)	(0.023)	(0.020)	(0.028)
Subdistrict FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Time FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Exclude RCT sample	N	N	Y	N	N	N	N	Y	N	N	N	N	Y	N	N
Include pre-treatment periods	N	N	N	Y	Y	N	N	N	Y	Y	N	N	N	Y	Y
Subdistrict trends	N	N	N	N	Y	N	N	N	N	Y	N	N	N	N	Y
Population weight	Y	N	Y	Y	Y	Y	N	Y	Y	Y	N	Y	N	N	N
Control Mean (2011 & 2014)	1.750	1.750	1.767	1.750	1.750	1.615	1.615	1.629	1.615	1.615	0.547	0.547	0.546	0.547	0.547
N	9414	9414	8484	15690	15690	9414	9414	8484	15690	15690	9414	9414	8484	15690	15690
Census waves	05 - 14	05 - 14	05 - 14	00 - 14	00 - 14	05 - 14	05 - 14	05 - 14	00 - 14	00 - 14	05 - 14	05 - 14	05 - 14	00 - 14	00 - 14

Table A2.4 displays the difference-in-differences estimate of the effect of the conditional cash transfer program on three alternative outcomes. Specifications use population weights unless otherwise noted. Columns 1 to 5 show treatment effects on the Poisson-extrapolated suicide rate (extrapolated from the subdistrict average of village-level incidences of at least one suicide in a given year). Columns 6 to 10 report the effects on a truncated version of the suicide rate based on the number of villages with at least one suicide in a subdistrict. Columns 11 to 15 report the treatment effect on a binary variable equal to one if there was at least one suicide in a subdistrict last year. All specifications include year and subdistrict fixed effects. Standard errors are clustered at district level. Column 2, 7, and 12 do not employ population weights. Columns 3, 8, and 13 exclude the RCT sample. Columns 4, 9, and 14 include all pre-treatment periods. Columns 5, 10, and 15 additionally include subdistrict time-trends. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A2.5: Roll-out: additional outcome definitions 2

	Suicide rate (2005 population)					Number of suicides (extrapolated - Poisson)					Number of suicides				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Treatment	-0.452***	-0.833***	-0.484***	-0.475***	-0.319***	-0.193***	-0.193**	-0.203***	-0.124***	-0.202***	-0.198***	-0.219***	-0.210***	-0.140***	-0.183**
	(0.109)	(0.163)	(0.119)	(0.113)	(0.117)	(0.053)	(0.086)	(0.057)	(0.044)	(0.067)	(0.055)	(0.084)	(0.060)	(0.050)	(0.091)
Subdistrict FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Time FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Exclude RCT sample	N	N	Y	N	N	N	N	Y	N	N	N	N	Y	N	N
Include pre-treatment periods	N	N	N	Y	Y	N	N	N	Y	Y	N	N	N	Y	Y
Subdistrict trends	N	N	N	N	Y	N	N	N	N	Y	N	N	N	N	Y
Population weight	Y	N	Y	Y	Y	N	Y	N	N	N	N	Y	N	N	N
Conrol mean (2011 & 2014)	2.226	2.226	2.041	2.226	2.226	0.920	1.127	0.922	0.920	0.920	1.061	1.318	1.326	1.061	1.061
N	9414	9414	8484	15690	15690	9414	9414	8484	15690	15690	9414	9414	8484	15690	15690
Census waves	05 - 14	05 - 14	05 - 14	00 - 14	00 - 14	05 - 14	05 - 14	05 - 14	00 - 14	00 - 14	05 - 14	05 - 14	05 - 14	00 - 14	00 - 14

Table A2.5 displays the difference-in-differences estimate of the effect of the conditional cash transfer program on three alternative outcomes. Specifications use population weights unless otherwise noted. Columns 1 to 5 show treatment effects on the suicide rate based on 2005 population numbers. Columns 6 to 10 report the effects on the Poisson extrapolated number of suicides (extrapolated from the subdistrict average of village-level incidences of at least one suicide in a given year). Columns 11 to 15 report the treatment effect on the number of suicides used to construct our main suicide rate. All specifications include year and subdistrict fixed effects. Standard errors are clustered at district level. Column 2, 7, and 12 do not employ population weights. Columns 3, 8, and 13 exclude the RCT sample. Columns 4, 9, and 14 include all pre-treatment periods. Columns 5, 10, and 15 additionally include subdistrict time-trends. $^{*}\ p < 0.10$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$

Table A2.6: Roll-out: drop partially treated

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent variable: Suicide rate							
Treatment	-0.336***	-0.336***	-0.358***	-0.339***	-0.442**	-0.252	-0.724*
	(0.103)	(0.087)	(0.112)	(0.103)	(0.180)	(0.283)	(0.403)
Subdistrict FE	Y	Y	Y	Y	Y		
Time FE	Y	Y	Y	Y	Y	N	Y
Cluster subdistrict	N	Y	N	Y	N	N	Y
Exclude RCT sample	N	N	Y	N	N	N	N
Include pre-treatment periods	N	N	N	Y	N	N	N
Subdistrict Trends	N	N	N	N	Y	N	N
Population weights	Y	Y	Y	Y	Y	N	N
						Y	N
Control mean (2011 & 2014)	2.016	2.016	2.033	2.016	2.016	1.680	1.835
N	8454	8454	7686	14090	14090	256	512
Census waves	05-14	05-14	05-14	00-14	00-14	11	05-11

Table A2.6 displays the difference-in-differences estimate of the effect of the conditional cash transfer on the suicide rate excluding partially treated subdistricts. The suicide rate is defined as described in Table A2.1. All specifications use population weights from 2005 unless otherwise noted. All specifications include year and subdistrict fixed effects. Standard errors are clustered at district level unless noted differently. Column 2 report standard errors clustered at subdistrict level. Column 3 excludes the RCT sample. Column 4 includes all pre-treatment periods. Column 5 further includes subdistrict time-trends. Column 6 does not use population weights. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A2.7: Main results: roll-out - not population-weighted

	(1)	(2)	(3)	(4)	(5)
Dependent variable: Suicide rate					
Treatment	-0.591***	-0.591***	-0.633***	-0.638***	-0.428**
	(0.164)	(0.143)	(0.179)	(0.158)	(0.174)
Subdistrict FE	Y	Y	Y	Y	Y
Time FE	Y	Y	Y	Y	Y
Cluster subdistrict	N	Y	N	Y	N
Exclude RCT sample	N	N	Y	N	N
Include pre-treatment periods	N	N	N	Y	Y
Subdistrict trends	N	N	N	N	Y
Control mean (2011 & 2014)	2.950	2.950	2.983	2.950	2.950
N	9414	9414	8484	15690	15690
Census waves	05-14	05-14	05-14	00-14	00-14

Table A2.7 displays the difference-in-differences estimates of the effect of the conditional cash transfer program on the suicide rate without using population weights. The suicide rate is defined as described in Table A2.1. All specifications include year and subdistrict fixed effects. Standard errors are clustered at the district level unless noted differently. Column 2 reports standard errors clustered at the subdistrict level. Column 3 excludes the RCT sample. Column 4 includes all pre-treatment periods. Column 5 additionally includes subdistrict time-trends. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A2.8: Main results: roll-out - 2006 subdistrict boundaries

	(1)	(2)	(3)	(4)
Dependent variable: Suicide rate				
Treatment	-0.264***	-0.264***	-0.255**	-0.347*
	(0.102)	(0.088)	(0.110)	(0.180)
Subdistrict FE	Y	Y	Y	Y
Time FE	Y	Y	Y	Y
Cluster subdistrict	N	Y	N	N
Exclude RCT sample	N	N	Y	N
Population weights	Y	Y	Y	N
Control mean (2011 & 2014)	2.170	2.170	2.183	3.046
N	14522	14522	13469	14522
Census waves	05-14	05-14	05-14	05-14

Table A2.8 displays the difference-in-differences estimates of the effect of the conditional cash transfer program on the suicide rate in subdistricts according to 2006 boundaries using population weights (unless otherwise noted). The suicide rate is defined as described in Table A2.1. All specifications include year and subdistrict fixed effects. Standard errors are clustered at the district level. Column 2 reports standard errors clustered at the subdistrict level. Column 3 excludes the RCT sample. Column 4 does not employ population weights. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A2.9: Main results: roll-out - additional robustness checks

	(1)	(2)	(3)	(4)
Dependent variable: Suicide rate				
Treatment	-0.359***	-0.255**	-0.393***	-0.284***
	(0.093)	(0.103)	(0.121)	(0.106)
Subdistrict FE	N	Y	N	N
Time FE	Y	Y	Y	Y
District FE	Y	N	N	N
District trends	N	Y	N	N
Baseline controls	N	N	Y	Y
Baseline controls × Post	N	N	N	Y
Population weights	Y	Y	Y	Y
Control mean (2011 & 2014)	2.016	2.016	2.033	2.016
N	9414	9414	8484	9414
Census waves	05-14	05-14	05-14	05-14

Table A2.9 displays estimates of the effect of the conditional cash transfer program on the suicide rate using the 2005 population weights. The suicide rate is defined as described in Table A2.1. Standard errors are clustered at the district level. Column 1 uses district instead of subdistrict fixed effects. Column 2 controls for district specific time trends. Column 3 includes baseline covariates instead of subdistrict fixed effects. Controls include per capita health, education, and social institutions as well as the fraction of farmers, fraction of individuals classified as poor, per-captita expenditure levels, and crime rates. Column 4 further includes interactions of these variable with time trends. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

2.B.3 Further sensitivity - Cash Transfer Program - RCT

Table A2.10: Randomized experiment - alternative outcomes 1

	Suicide rate (extrapolated - Poisson)					Suicide rate (truncated)					Any Suicide				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Treatment	-0.236	-0.160	-0.633**	-0.244	-0.318	-0.225	-0.192	-0.562**	-0.224	-0.306	-0.096	-0.084	-0.142	-0.078	-0.086
	(0.227)	(0.234)	(0.305)	(0.224)	(0.294)	(0.206)	(0.253)	(0.272)	(0.203)	(0.266)	(0.066)	(0.067)	(0.086)	(0.066)	(0.059)
Subdistrict FE	N	N	Y	Y	N	N	N	Y	Y	N	N	N	Y	Y	N
Time FE	N	N	Y	Y	N	N	N	Y	Y	N	N	N	Y	Y	N
Include pre-treatment periods	N	N	N	Y	N	N	N	N	Y	N	N	N	N	Y	N
Baseline suicide	Y	N	N	N	Y	Y	N	N	N	Y	Y	N	N	N	Y
Population weights	Y	Y	Y	Y	N	Y	Y	Y	Y	N	Y	Y	Y	Y	N
Control mean (2011)	1.514	1.514	1.514	1.514	1.775	1.420	1.420	1.420	1.420	1.666	0.578	0.578	0.578	0.578	0.545
N	310	310	620	1240	310	310	310	620	1240	310	310	310	620	1240	310
Census waves	11	11	05 - 11	00 - 11	11	11	11	05 - 11	00 - 11	11	11	11	05 - 11	00 - 11	11

Table A2.10 displays RCT treatment effects on three alternative outcomes. All specifications use population weights unless otherwise noted. Columns 1 to 5 show treatment effects on the Poisson-extrapolated suicide rate (extrapolated from the subdistrict average of village-level incidences of at least one suicide in a given year). Columns 6 to 10 report the effects on a truncated version of the suicide rate based on the number of villages with at least one suicide in a subdistrict. Columns 11 to 15 report the treatment effect on a binary variable equal to one if there was at least one suicide in a subdistrict last year. Columns 1, 6, and 11 report an ANCOVA specification. In Column 2, 7, and 12 we report treatment effect estimates of a post comparison of treated and control subdistricts. Columns 3, 8, and 13 report difference-in-differences estimates of the treatment effect. In Columns 4, 9, and 14 we report difference-in-differences results including all pre-treatment time periods. In Columns 5, 10, and 15 we report ANCOVA estimates without using baseline population weights. Standard errors are clustered at subdistrict level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A2.11: Randomized experiment - alternative outcomes 2

	Suicide rate (2005 population)					Number of suicides (extrapolated - Poisson)					Number of suicides				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Treatment	-0.344	-0.256	-0.663**	-0.313	-0.483	-0.099	-0.129	-0.345**	-0.118	-0.170	-0.149	-0.129	-0.365**	-0.167	-0.232
	(0.281)	(0.293)	(0.328)	(0.280)	(0.340)	(0.133)	(0.164)	(0.171)	(0.131)	(0.182)	(0.161)	(0.164)	(0.183)	(0.157)	(0.250)
Subdistrict FE	N	N	Y	Y	N	N	N	Y	Y	N	N	N	Y	Y	N
Time FE	N	N	Y	Y	N	N	N	Y	Y	N	N	N	Y	Y	N
Include pre-treatment periods	N	N	N	Y	N	N	N	N	Y	N	N	N	N	Y	N
Baseline suicide	Y	N	N	N	Y	Y	N	N	N	Y	Y	N	N	N	Y
Population weights	Y	Y	Y	Y	N	N	N	N	N	Y	N	N	N	N	Y
Control mean (2011)	1.883	1.883	1.883	1.883	2.167	0.878	0.878	0.878	0.878	1.014	1.035	1.035	1.035	1.035	1.228
N	310	310	620	1240	310	310	310	620	1240	310	310	310	620	1240	310
Census waves	11	11	05 - 11	00 - 11	11	11	11	05 - 11	00 - 11	11	11	11	05 - 11	00 - 11	11

Table A2.11 displays RCT treatment effects on three alternative outcomes. Specifications use population weights unless otherwise noted. Columns 1 to 5 show treatment effects on the suicide rate based on 2005 population numbers. Columns 6 to 10 report the effects the Poisson extrapolated number of suicides (extrapolated from the subdistrict average of village-level incidences of at least one suicide in a given year). Columns 11 to 15 report the treatment effect on the number of suicides used to construct our main suicide rate. Columns 1, 6, and 11 report an ANCOVA specification. In Column 2, 7, and 12 we report treatment effect estimates of a post comparison of treated and control subdistricts. Columns 3, 8, and 13 report difference-in-differences estimates of the treatment effect. In Columns 4, 9, and 14 we report difference-in-differences results including all pre-treatment time periods. In Columns 5, 10, and 15 we report ANCOVA estimates without using baseline population weights. Standard errors are clustered at subdistrict level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A2.12: Randomized experiment - drop partially treated

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent variable: Suicide rate							
Treatment	-0.326	-0.238	-0.640**	-0.640**	-0.453	-1.041*	-0.467
	(0.267)	(0.276)	(0.318)	(0.261)	(0.334)	(0.593)	(0.325)
Subdistrict FE	N	N	Y	Y	Y	Y	N
Time FE	N	N	Y	Y	Y	Y	N
Cluster district	N	N	N	Y	N	N	N
Include pre-treatment periods	N	N	N	N	Y	Y	N
Subdistrict trends	N	N	N	N	N	Y	N
Baseline suicide	Y	N	N	N	N	N	Y
Population weight	Y	Y	Y	Y	Y	Y	N
Control mean (2011)	1.774	1.774	1.774	1.774	1.774	1.774	2.058
N	282	282	564	564	1128	1128	282
Census waves	11	11	05-11	05-11	00-11	00-11	11

Table A2.12 reports the RCT effects on the suicide rate excluding partially treated subdistricts. All specifications use population weights from 2005 unless otherwise noted. The suicide rate is defined as described in Table A2.1. Column 1 reports an ANCOVA specification controlling for the 2005 suicide rate. Column 2 displays the treatment effect estimate of a post comparison of treated and control subdistricts. Columns 3 to 6 report difference-in-differences estimates of the treatment effect. Standard errors are clustered at the subdistrict level unless otherwise noted. In Column 3 we control for time fixed effects as well as subdistrict fixed effects. Column 4 reports standard errors clustered at district level. Column 5 includes data from the 2003 and 2000 census waves. Column 6 further includes subdistrict specific time-trends on top of subdistrict and time fixed effects. Column 7 shows the ANCOVA specification without population weights. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A2.13: Main results: randomized experiment - without population weights

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable: Suicide rate						
Treatment	-0.474	-0.258	-0.809**	-0.809**	-0.466	-1.302*
	(0.325)	(0.275)	(0.401)	(0.315)	(0.334)	(0.726)
Subdistrict FE	N	N	Y	N	Y	Y
Time FE	N	N	Y	Y	Y	Y
Cluster district	N	N	N	Y	N	N
Include pre-treatment periods	N	N	N	N	Y	Y
Subdistrict trends	N	N	N	N	N	Y
Baseline suicide	Y	N	N	N	N	N
Population weights	N	N	N	N	N	N
Control mean (2011)	2.058	2.058	2.058	2.058	2.058	2.058
N	310	310	620	620	1240	1240
Census waves	11	11	05-11	05-11	00-11	00-11

Table 2.3 displays the results of the RCT experiment without population weights. The suicide rate is defined as described in Table A2.1. Column 1 reports an ANCOVA specification controlling for the 2005 suicide rate. Column 2 displays the treatment effect estimate of a post comparison of treated and control subdistricts. Columns 3 to 6 report difference-in-differences estimates of the treatment effect. Standard errors are clustered at the subdistrict level unless otherwise noted. Estimates are weighted using population size from 2005 unless otherwise noted. In Column 3 we control for time fixed effects as well as subdistrict fixed effects. Column 4 reports standard errors clustered at district level. Column 5 includes data from the 2003 and 2000 census waves. Column 6 further includes subdistrict specific time trends on top of subdistrict and time fixed effects. Standard errors clustered at the subdistrict level in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A2.14: Main results: randomized experiment - 2006 boundaries

	(1)	(2)	(3)	(4)	(5)
Dependent variable: Suicide rate					
Treatment	-0.248	-0.192	-0.544*	-0.544*	-0.537
	(0.256)	(0.263)	(0.309)	(0.294)	(0.383)
Subdistrict FE	N	N	Y	Y	N
Time FE	N	N	Y	Y	N
Cluster district	N	N	N	Y	N
Baseline suicide	Y	N	N	N	Y
Population weights	Y	Y	Y	Y	N
Control mean (2011)	1.720	1.720	1.720	1.720	2.205
N	351	351	702	702	351
Census waves	11	11	05-11	05-11	11

Table A2.14 displays the results of the RCT experiment on the suicide rate according to 2006 subdistrict boundaries with 2005 population weights unless otherwise noted. The suicide rate is defined as described in Table A2.1. Standard errors are clustered at subdistrict level unless noted otherwise. Column 1 reports an ANCOVA specification controlling for the 2005 suicide rate. Column 2 displays the treatment effect estimate of a post comparison of treated and control subdistricts. Columns 3 and 4 report difference-in-differences estimates of the treatment effect. In Column 3 we control for time fixed effects as well as subdistrict fixed effects. Column 4 reports standard errors clustered at district level. Column 5 reports the ANCOVA specification without population weights. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A2.15: Main results: randomized experiment - additional robustness checks

	(1)	(2)	(3)	(4)
Dependent variable: Suicide rate				
Treatment	-0.665**	-0.649**	-0.809**	-0.665**
	(0.330)	(0.272)	(0.405)	(0.321)
Subdistrict FE	N	Y	N	N
Time FE	Y	Y	Y	Y
District FE	Y	N	N	N
District trends	N	Y	N	N
Baseline controls	N	N	Y	Y
Baseline controls × Post	N	N	N	Y
Population weights	Y	Y	Y	Y
Control mean (2011)	1.774	1.774	1.774	1.774
N	620	620	620	620
Census waves	05-11	05-11	05-11	05-11

Table A2.15 displays the results using the RCT experiment on the suicide rate with population weights. The suicide rate is defined as described in Table A2.1. Columns 1 to 4 report difference-in-differences estimates of the treatment effect. Column 1 uses district instead of subdistrict fixed effects, while also controlling for treatment status. Column 2 controls for district specific time trends. Column 3 includes baseline covariates instead of subdistrict fixed effects. Controls include per capita health, education, and social institutions as well as the fraction of farmers, fraction of individuals classified as poor, per-captita expenditure levels, and crime rates. Column 4 further includes interactions of these variable with time trends. Standard errors clustered in subdistrict level in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

2.B.4 Dynamics of treatment effects

Table A2.16: Dynamics of treatment effects

	(1) RCT	(2) Treat 07-08	(3) Treat 10-11	(4) Treat 12-13
Dependent variable: Suicide rate				
Panel A: First post treatment census wave				
Treatment	-0.665**	-0.509***	-0.289	-0.280*
	(0.318)	(0.136)	(0.175)	(0.156)
Years since launch of PKH	3-4	3-4	0-1	1-2
Census waves	05 & 11	05 & 11	05 & 11	11 & 14
Number of treated	167	621	233	1137
Number of counterfactuals	143	1186	1186	1186
Receiving Treatment	Y	Y	Partial	Y
Panel B: Second post treatment census wave				
Treatment	-0.273	-0.433**	-0.540**	
	(0.371)	(0.207)	(0.237)	
Years since launch of PKH	6-7	6-7	3-4	
Census waves	05 & 14	05 & 14	05 & 14	
Number of treated	167	621	233	
Number of counterfactuals	143	1186	1186	
Receiving treatment	Partial	Partial	Y	

Table A2.16 displays treatment effects on the suicide rate for each treatment wave relative to never treated subdistricts. The suicide rate is defined as explained in Table A2.1. Estimates are weighted using 2005 population size. Panel A shows treatment effects in the first PODES survey after subdistricts started receiving the conditional cash transfer program. Panel B shows treatment effects in the second PODES survey after subdistricts started receiving the program. Impacts in Columns 1 and 2 report treatment effect estimates past the six year duration of the program. Standard errors are clustered at the district level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A2.17: Dynamics of treatment effects: not population-weighted

	(1) RCT	(2) Treat 07-08	(3) Treat 10-11	(4) Treat 12-13
Dependent variable: Suicide rate				
Panel A: First post treatment census wave				
Treatment	-0.809**	-0.612***	-0.162	-0.628**
	(0.401)	(0.173)	(0.302)	(0.259)
Years since launch of PKH	3-4	3-4	0-1	1-2
Census waves	05 & 11	05 & 11	05 & 11	11 & 14
Number of treated	167	621	233	1137
Number of counterfactuals	143	1186	1186	1186
Receiving Treatment	Y	Y	Partial	Y
Panel B: Second post treatment census wave				
Treatment	-0.017	-0.629**	-0.733*	
	(0.486)	(0.305)	(0.440)	
Years since launch of PKH	6-7	6-7	3-4	
Census waves	05 & 14	05 & 14	05 & 14	
Number of treated	167	621	233	
Number of counterfactuals	143	1186	1186	
Receiving treatment	Partial	Partial	Y	

Table A2.17 displays treatment effects for each treatment wave relative to subdistricts who did not receive the treatment until 2013 without using population weights. The suicide rate is defined as described in Table A2.1. Panel A shows treatment effects in the first PODES survey after subdistricts started receiving the conditional cash transfer program. Panel B shows treatment effects in second PODES survey after subdistricts started receiving the program. Impacts in Columns 1 and 2 report treatment effect estimates past the six year duration of the program. Standard errors are clustered at the district level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A2.18: Roll-out effects: treatment intensity

	(1)	(2)
Dependent variable: Suicide rate		
Treatment	-0.358***	-0.336***
	(0.101)	(0.101)
Intensity		-0.012
		(0.024)
N	9414	9414
Census Waves	05-14	05-14

Table A2.18 displays treatment effects on the suicide rate controlling for treatment duration in years since start of the program. The suicide rate is defined as described in Table A2.1. All specifications include time and subdistrict fixed effects. Standard errors are clustered at the district level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

2.B.5 Rainfall

Table A2.19: Interactions between agricultural productivity shocks and the cash transfers - not population weighted

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable: Suicide rate						
Rain (z-scored)	-0.137**	-0.204***	-0.134***	-0.162***	-0.149***	-0.227***
	(0.054)	(0.061)	(0.051)	(0.053)	(0.055)	(0.064)
	[0.058]	[0.064]	[0.050]	[0.053]	[0.055]	[0.065]
Treat		-0.245		-0.233		-0.174
		(0.151)		(0.172)		(0.157)
		[0.161]		[0.184]		[0.170]
Rain (z-scored)× Treat		0.302***		0.190*		0.350***
		(0.106)		(0.108)		(0.116)
		[0.103]		[0.119]		[0.115]
Subdistrict FE	Y	Y	Y	Y	Y	Y
Time FE	Y	Y	Y	Y	Y	Y
Subdistrict trends	N	N	Y	Y	N	N
Include lagged rainfall	N	N	N	N	Y	Y
N	15690	15690	15690	15690	15690	15690
Census waves	00-14	00-14	00-14	00-14	00-14	00-14

Table A2.19 displays the impact of rainfall on suicides. The suicide rate is defined as described in Table A2.1. Odd columns report the impact of standardized rainfall on suicides. Even columns include the treatment variable from the conditional cash transfer roll-out and the interaction with rainfall. All specifications include subdistrict and year fixed effects. Columns 3 and 4 additionally control for subdistrict time trends. Columns 5 and 6 also include the first, second and third lag of rainfall. Parentheses report Conley (1999) standard errors, allowing for arbitrary time correlation and two-dimensional spatial correlation within a 100 kilometre radius of the subdistrict centroid (used for stars). Standard errors clustered at the district level are in square brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A2.20: Rainfall: symmetry of effects

	(1)	(2)	(3)
Dependent variable: Suicide rate			
Rain (z-scored)	0.021	0.001	0.020
	(0.063)	(0.060)	(0.061)
	[0.057]	[0.056]	[0.055]
Positive Shock	-0.186*	-0.160*	-0.210**
	(0.100)	(0.091)	(0.094)
	[0.064]	[0.060]	[0.065]
Negative Shock	0.118	0.076	0.136
	(0.107)	(0.102)	(0.101)
	[0.089]	[0.088]	[0.089]
p(|Positive Shock| = |Negative Shock|)	0.648	0.555	0.605
Subdistrict FE	Y	Y	Y
Time FE	Y	Y	Y
Subdistrict trends	N	Y	N
Include lagged rainfall	N	N	Y
N	15690	15690	15690
Census waves	00-14	00-14	00-14

Table A2.20 displays the impact of non-linear measures of rainfall on the suicide rate. The suicide rate is defined as described in Table A2.1. The Table reports the results of a regression of standardized rainfall and dummies indicating rainfall above the 66th and below the 33th percentile of z-scored rainfall in our sample. All specifications include subdistrict and year fixed effects. Column 2 also controls for subdistrict specific time trends. Column 3 includes the first, second and third lag of rainfall. Parentheses report Conley (1999) standard errors, allowing for arbitrary time correlation and two-dimensional spatial correlation within a 100 kilometre radius of the subdistrict centroid (used for stars). Standard errors clustered at the district level are displayed in square brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A2.21: Rainfall: non-linear interaction of rainfall and cash transfer

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable: Suicide rate						
Positive shock	-0.169*	-0.163*	-0.157**	-0.154*	-0.190**	-0.189**
	(0.086)	(0.087)	(0.080)	(0.081)	(0.083)	(0.086)
	[0.055]	[0.055]	[0.053]	[0.054]	[0.059]	[0.060]
Positive shock × Treat		0.275		0.113		0.370
		(0.396)		(0.296)		(0.398)
		[0.404]		[0.376]		[0.419]
Negative shock	0.098	0.161*	0.089	0.124	0.129	0.171*
	(0.096)	(0.098)	(0.088)	(0.092)	(0.092)	(0.097)
	[0.071]	[0.075]	[0.069]	[0.073]	[0.070]	[0.075]
Negative shock × Treat		-0.225		-0.158		-0.183
		(0.141)		(0.122)		(0.140)
		[0.075]		[0.073]		[0.075]
Treat		-0.198		-0.259***		-0.227*
		(0.120)		(0.091)		(0.118)
		[0.121]		[0.115]		[0.121]
Subdistrict FE	Y	Y	Y	Y	Y	Y
Time FE	Y	Y	Y	Y	Y	Y
Subdistrict trends	N	N	Y	Y	N	N
Include lagged rainfall	N	N	N	N	Y	Y
N	15690	15690	15690	15690	15690	15690
Census waves	00-14	00-14	00-14	00-14	00-14	00-14

Table A2.21 displays the impact of non-linear measures of rainfall on the suicide rate. The suicide rate is defined as described in Table A2.1. The Table reports the results of a regression of dummies indicating rainfall above the 66th and below the 33th percentile of z-scored rainfall in our sample. Even columns also display the interaction with receiving the conditional cash transfer. All specifications include subdistrict and year fixed effects. Columns 3 and 4 also control for subdistrict specific time trends. Columns 5 and 6 includes the first, second and third lag of z-scored rainfall. Parentheses report Conley (1999) standard errors, allowing for arbitrary time correlation and two-dimensional spatial correlation within a 100 kilometre radius of the subdistrict centroid (used for stars). Standard errors clustered at the district level are displayed in square brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A2.22: Interactions between agricultural productivity shocks and the cash transfers - detrended rainfall and population weighted

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable: Suicide rate						
Rain (z-scored)	-0.098**	-0.114**	-0.093**	-0.105**	-0.107**	-0.129**
	(0.043)	(0.046)	(0.041)	(0.044)	(0.045)	(0.050)
	[0.033]	[0.036]	[0.033]	[0.036]	[0.034]	[0.039]
Treat		-0.314***		-0.340***		-0.292***
		(0.093)		(0.088)		(0.096)
		[0.105]		[0.099]		[0.108]
Rain (z-scored)× Treat		0.080		0.058		0.106
		(0.079)		(0.076)		(0.085)
		[0.075]		[0.073]		[0.081]
Subdistrict FE	Y	Y	Y	Y	Y	Y
Time FE	Y	Y	Y	Y	Y	Y
Subdistrict trends	N	N	Y	Y	N	N
Include lagged rainfall	N	N	N	N	Y	Y
N	15690	15690	15690	15690	15690	15690
Census waves	00-14	00-14	00-14	00-14	00-14	00-14

Table A2.22 displays the impact of rainfall on suicides. The suicide rate is defined as described in Table A2.1. Rainfall is net of subdistrict level trends and standardized to have mean zero and standard deviation one. Odd columns report the impact of rainfall on suicides. Even columns include the treatment variable from the conditional cash transfer roll-out and the interaction with rainfall. All specifications include subdistrict and year fixed effects. Columns 3 and 4 additionally control for subdistrict time trends. Columns 5 and 6 also include the first, second and third lag of rainfall. Observations are weighted by 2005 subdistrict population. Parentheses report Conley (1999) standard errors, allowing for arbitrary time correlation and two-dimensional spatial correlation within a 100 kilometre radius of the subdistrict centroid (used for stars). Standard errors clustered at the district level are in square brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A2.23: Interactions between agricultural productivity shocks and the cash transfers - detrended and not population weighted

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable: Suicide rate						
Rain (z-scored)	-0.158***	-0.208***	-0.148**	-0.189***	-0.171***	-0.233***
	(0.059)	(0.063)	(0.058)	(0.061)	(0.060)	(0.067)
	[0.051]	[0.056]	[0.050]	[0.054]	[0.049]	[0.057]
Treat		-0.434***		-0.494***		-0.388***
		(0.131)		(0.129)		(0.132)
		[0.149]		[0.143]		[0.151]
Rain (z-scored)× Treat		0.289**		0.235**		0.341***
		(0.117)		(0.115)		(0.126)
		[0.110]		[0.108]		[0.121]
Subdistrict FE	Y	Y	Y	Y	Y	Y
Time FE	Y	Y	Y	Y	Y	Y
Subdistrict trends	N	N	Y	Y	N	N
Include lagged rainfall	N	N	N	N	Y	Y
N	15690	15690	15690	15690	15690	15690
Census waves	00-14	00-14	00-14	00-14	00-14	00-14

Table A2.23 displays the impact of rainfall on suicides. The suicide rate is defined as described in Table A2.1. Rainfall is net of subdistrict level trends and standardized to have mean zero and standard deviation one. Odd columns report the impact of rainfall on suicides. Even columns include the treatment variable from the conditional cash transfer roll-out and the interaction with rainfall. All specifications include subdistrict and year fixed effects. Columns 3 and 4 additionally control for subdistrict time trends. Columns 5 and 6 also include the first, second and third lag of rainfall. Parentheses report Conley (1999) standard errors, allowing for arbitrary time correlation and two-dimensional spatial correlation within a 100 kilometre radius of the subdistrict centroid (used for stars). Standard errors clustered at the district level are in square brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

2.B.6 Microdata on Depression: IFLS

Table A2.24: Microdata on depression: individual level heterogeneity

	Male	Aged ¡ 50 (median)	Baseline expenditure (z-scored)
	(1)	(2)	(3)
Dependent variable: depression z-score			
Rain (z-scored)	-0.175***	-0.067	-0.128***
	(0.052)	(0.047)	(0.042)
	[0.050]	[0.049]	[0.042]
Rain (z-scored) × Heterogeneity	0.107*	-0.152**	0.047**
	(0.059)	(0.066)	(0.022)
	[0.057]	[0.068]	[0.023]
Individual FE	Y	Y	Y
Time FE	Y	Y	Y
N	5192	5192	5188
IFLS Waves used	4-5	4-5	4-5

Table A2.24 displays heterogeneous effects of rainfall on depression z-scores by individual level characteristics using the Indonesian Family Life Survey (IFLS). Depression scores are measured in 2011 and 2014 using the 10 item CES-D scale (Radloff, 1997). Age is based on the respondent's age in IFLS wave 4. Baseline expenditure uses pre-determined per-capita expenditure from IFLS wave 3. All specifications control for differential trends. The sample is restricted to individuals tracked in all five IFLS waves. The sample is further restricted to individuals living in households with a member working in agriculture (self-employed without or with temporary workers, or as casual or family workers) in any of the five waves. IFLS wave 1 contains no coded sector information. Therefore, all individuals working as temporary workers or those who are self-employed without permanent workers are counted as working in agriculture. This makes up 60% of IFLS individuals. The coefficients of the heterogeneity variable are omitted to save space. All specifications are conditional on individual and time fixed effects and allowing for different trends by the heterogeneity variable. Parentheses report Conley (1999) standard errors, allowing for arbitrary time correlation and two-dimensional spatial correlation within a 100 kilometre radius of the subdistrict centroid (used for stars). Standard errors clustered at the district level are in square brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A2.25: Correlates of depression scores

	(1)	(2)	(3)
Dependent variable: Baseline depression z-score			
Bottom decile of expenditure	0.083*** (0.026)		0.056** (0.027)
CCT recipient in 2014		0.142*** (0.049)	0.131*** (0.049)
N	25532	19637	19637
IFLS wave used	4	4	4

Table A2.25 displays correlations with depression scores in the Indonesian Family Life Survey (IFLS) wave 4. IFLS 4 was collected before the roll-out of the conditional cash transfer program in 2007 and 2008. Depression scores are measured using the 10 item CES-D scale (Radloff, 1997). Column 1 shows the correlation between being in the bottom decile of the IFLS wave 4 per-capita household expenditure distribution and depression z-scores. Column 2 shows the correlation between having received the program at some point between IFLS wave 4 and wave 5 (collected in IFLS wave 5 (2014)). Column three includes both variables simultaneously. The sample is restricted to individuals tracked in IFLS wave 4 and wave 5. Standard errors are clustered at the subdistrict level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

2.B.7 Mechanisms and heterogeneity

Table A2.26: Including endogenous controls

	(1)	(2)	(3)	(4)	(5)
Dependent variable: Suicide rate					
Treatment	-0.358***	-0.327***	-0.366***	-0.325***	-0.271***
	(0.101)	(0.095)	(0.100)	(0.107)	(0.103)
Subdistrict FE	Y	Y	Y	Y	Y
Time FE	Y	Y	Y	Y	Y
Crime controls	N	Y	N	N	N
Institution controls	N	N	Y	N	N
Social capital controls	N	N	N	N	Y
Control mean (2011 & 2014)	1.797	1.797	1.797	1.797	1.797
N	9414	9414	9414	6276	6276
Census waves	05-14	05-14	05-14	05-11	05-11

Table A2.26 displays treatment effects on the suicide rate controlling for potential mediators. The suicide rate is defined as described in Table A2.1. All estimations are weighted using 2005 population size. All specifications include time and subdistrict fixed effects. In Columns (1) to (3) we use observations from the 2005, 2011 and 2014 census. In Columns (4) and (5) we use observations from the 2005 and 2011 village census, as data on social capital was not available in the 2014 census. Standard errors are clustered at the district level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

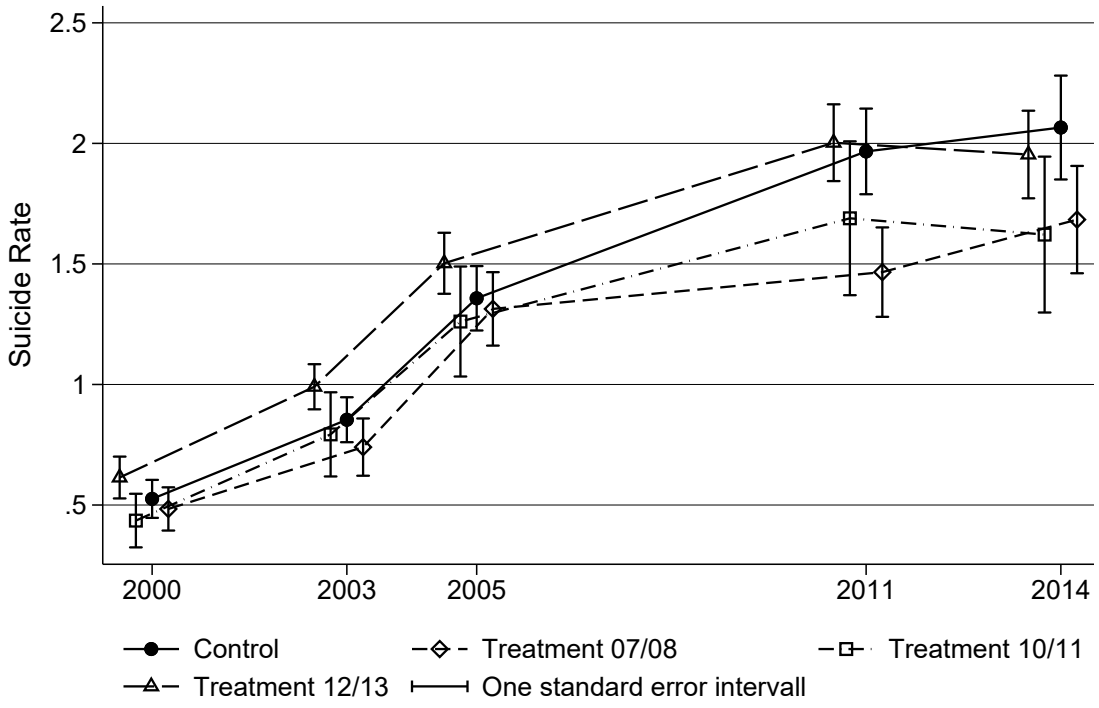
Table A2.27: Subdistrict level heterogeneous treatment effects

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent variable: Suicide rate							
Treatment	-0.234	-0.357***	-0.356***	-0.380***	-0.346***	-0.283**	-0.274*
	(0.147)	(0.101)	(0.113)	(0.103)	(0.108)	(0.139)	(0.145)
Treatment ×	0.218*	0.122					
Fract poor (z)	(0.122)	(0.095)					
Treatment ×	-0.105		-0.117				
Perc. Farmers (z)	(0.102)		(0.072)				
Treatment ×	0.120			0.026			
Per Capita Excp (z)	(0.086)			(0.059)			
Treatment ×	0.145				0.077		
Per Capita Crimes (z)	(0.218)				(0.183)		
Treatment ×	0.277					0.015	
Per Capita Social Institutions (z)	(0.198)					(0.191)	
Treatment ×	-0.244						-0.127
Per Capita Health Institutions (z)	(0.256)						(0.223)
Minimum detectable effect size		0.266	0.202	0.165	0.512	0.535	0.624
N	9414	9414	9414	9414	9414	9414	9414
Census waves	05-14	05-14	05-14	05-14	05-14	05-14	05-14

Table A2.27 displays heterogeneous treatment effects on the suicide rate for the cash transfer roll-out by baseline district level characteristics. The suicide rate is defined as described in Table A2.1. All specification include time and subdistrict fixed effects. All estimations are weighted using 2005 population size. All interaction term variables, interact_{sd}, are based on the 2005 census, and are z-scored using the mean and standard deviation of the sample. All specifications control for differential trends by interaction variables, $\text{interact}_{sd} \times \text{post}_t$, where post is defined as post-baseline years. "Minimum detectable effect size" provides us with the minimum detectable effect size at 0.8 power and significance levels of 0.05 for heterogeneous treatment effects for the different interactions variables. Estimates are based on years 2005 to 2014. Standard errors are clustered at the district level. $^{*}\ p < 0.10$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$

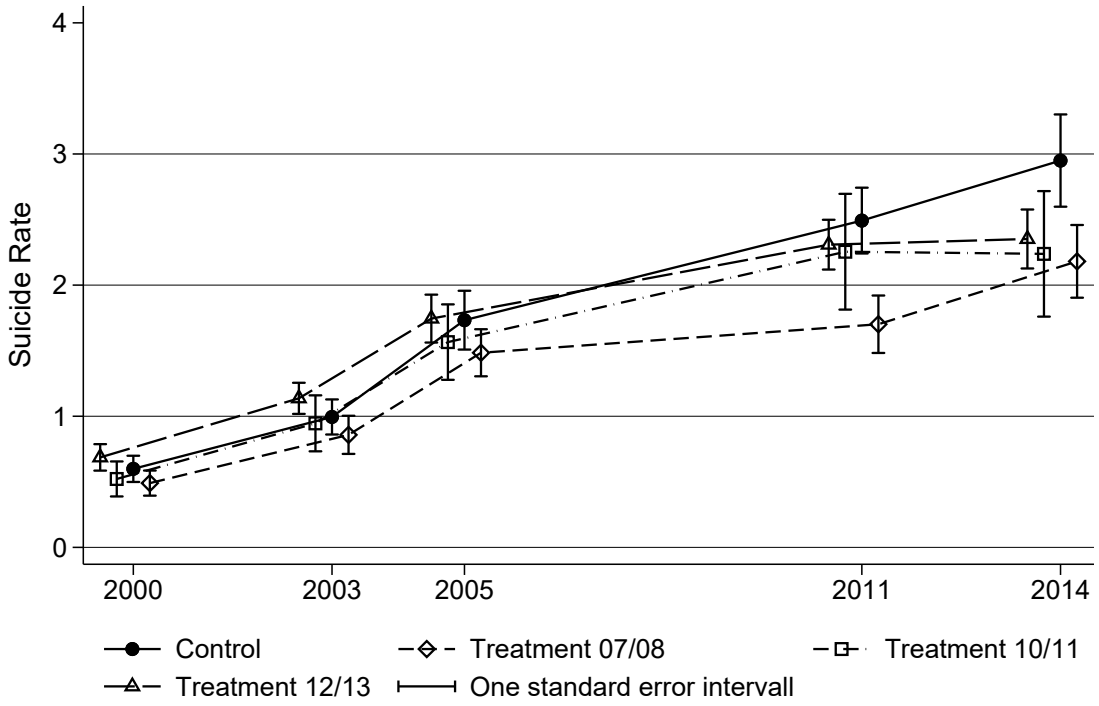
2.C Additional Figures

Figure A2.1: Mean suicide rates over time



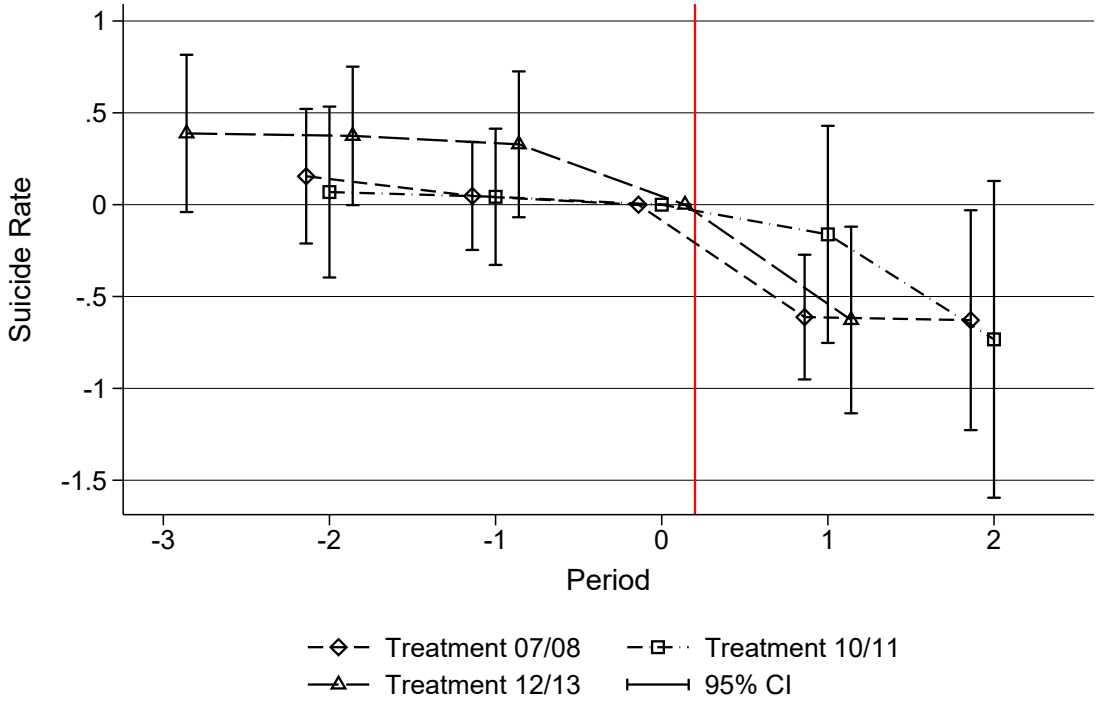
Notes: Figure A2.1 displays mean suicide rates over time by treatment wave. The definition of the suicide rate per 100,000 people changes slightly over time. In 2014 it is defined as the number of suicides and suicide attempts per 100,000. In 2011 it is defined as the number of suicides per 100,000. From 2000 to 2005 the suicide rate is defined as the number of villages in a given subdistrict with at least one suicide per 100,000.

Figure A2.2: Mean suicide rates by treatment group over time – not population-weighted



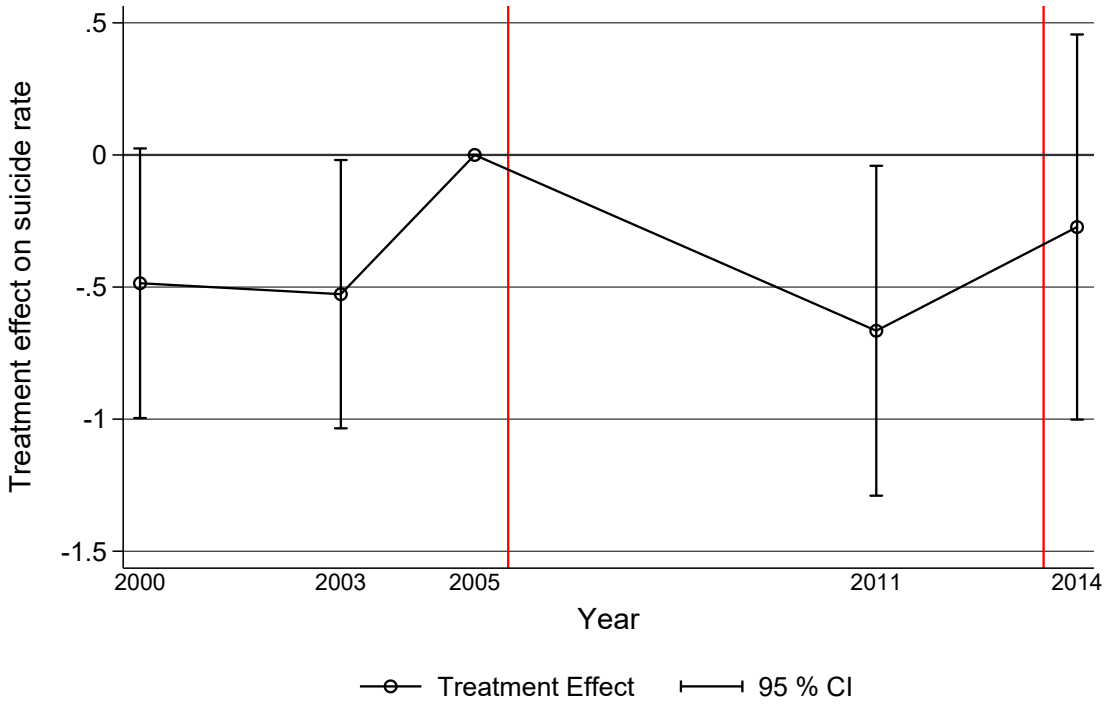
Notes: Figure A2.2 displays mean suicide rates over time by treatment wave without population weights. The suicide rate is defined as in Figure A2.1.

Figure A2.3: Event-study: roll-out of conditional cash transfer program – not population-weighted



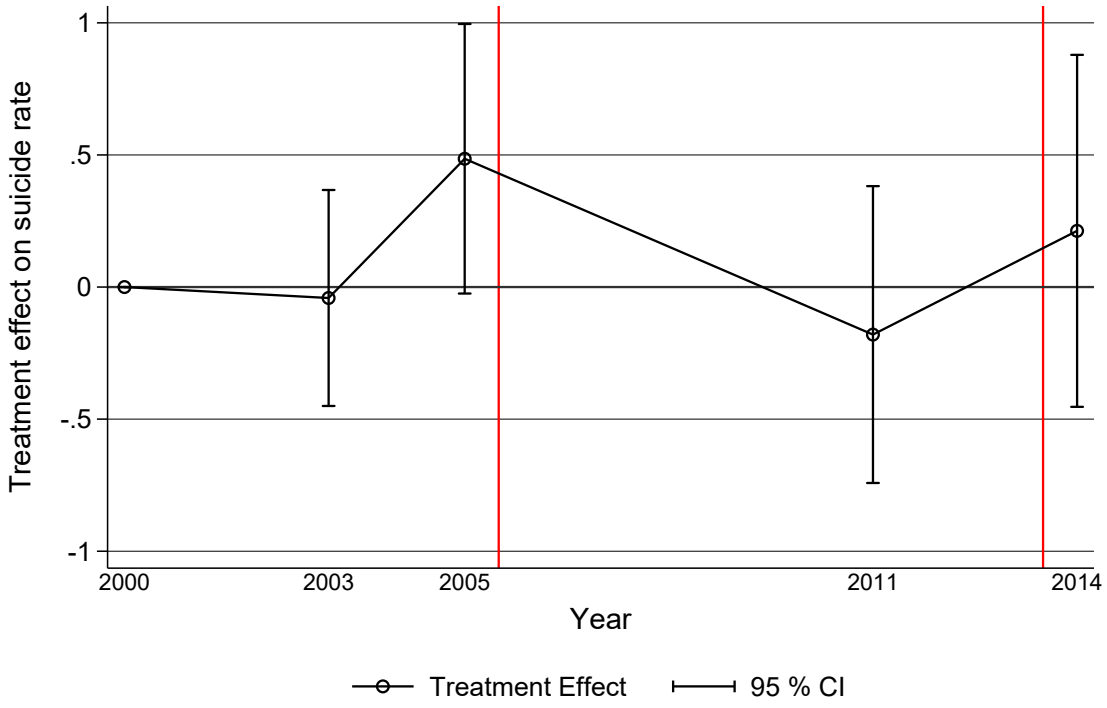
Notes: Figure 2.1 displays pre-trends relative to the timing of the introduction of the conditional cash transfer program by treatment wave without population weights. The suicide rate is defined as in Figure A2.1. Differences in the first period before the treatment are normalized to zero. The displayed coefficients are the difference-in-differences treatment effect estimates relative to subdistricts that had not received the treatment until 2014. All standard errors are clustered at the district level. The vertical red line indicates when the program was first introduced.

Figure A2.4: Event-study: randomized controlled trial



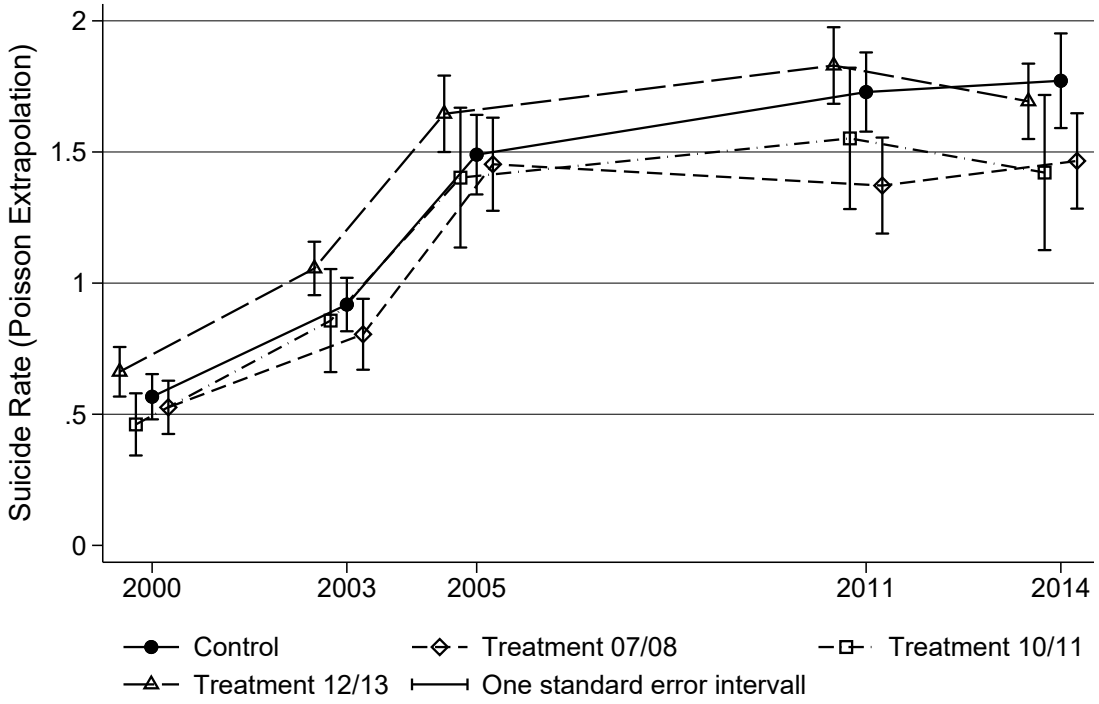
Notes: Figure A2.4 displays the coefficients of an interaction between an indicator for whether a given subdistrict was randomly chosen to receive the PKH cash transfer program and year dummies, conditional on year and subdistrict fixed effects. Effects in 2005 are normalized to zero. Standard errors are clustered at subdistrict level. The suicide rate is defined as in Figure A2.1. The vertical red lines indicate the treatment period.

Figure A2.5: Event-study relative to 2000: randomized controlled trial



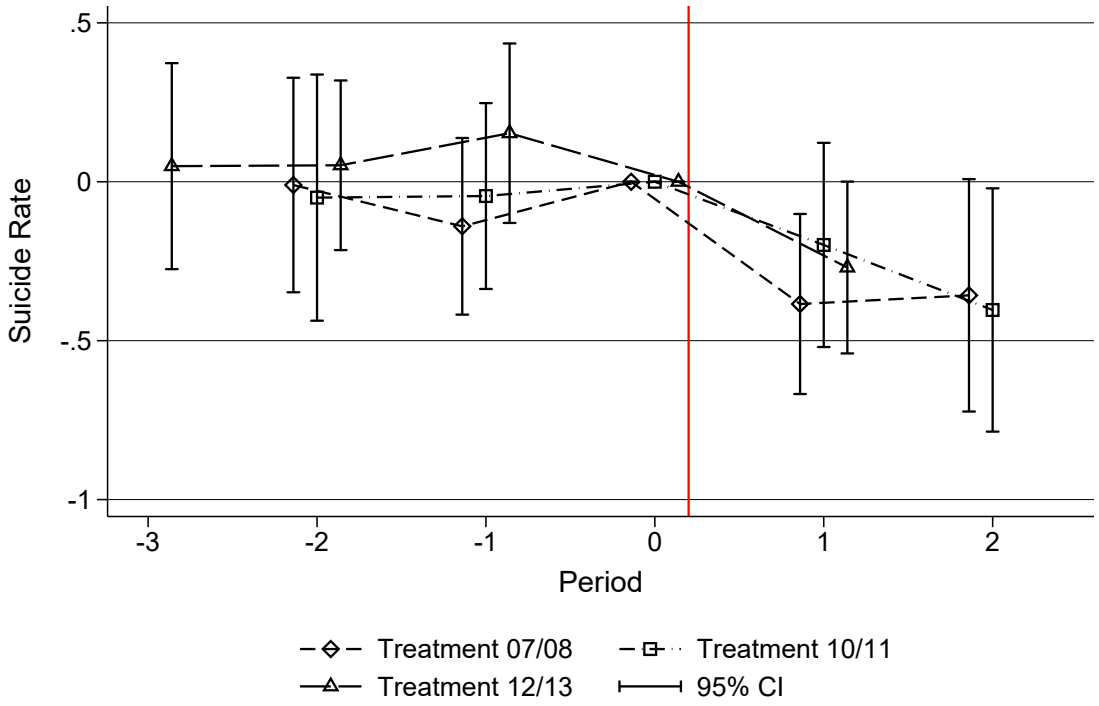
Notes: Figure A2.5 displays the coefficients of an interaction between an indicator for whether a given subdistrict was randomly chosen to receive the PKH cash transfer program and year dummies, conditional on year and subdistrict fixed effects. Effects in 2000 are normalized to zero. Standard errors are clustered at subdistrict level. The suicide rate is defined as in Figure A2.1. The vertical red lines indicate the treatment period.

Figure A2.6: Mean suicide rates over time (Poisson extrapolation)



Notes: Figure A2.6 displays mean Poisson-extrapolated suicide rates over time by treatment wave with 2005 population weights. Subdistrict number of suicides is extrapolated from mean incidence of at least one subdistrict assuming a homogeneous village level Poisson distribution within each subdistrict and year.

Figure A2.7: Event-study: roll-out of conditional cash transfer program (Poisson extrapolation)



Notes: Figure A2.7 displays pre-trends relative to the timing of the introduction of the conditional cash transfer program by treatment wave. The suicide rate is defined as in Figure A2.6. All specifications use 2005 population weights. Differences in the first period before the treatment are normalized to zero. The displayed coefficients are the difference-in-differences treatment effect estimates relative to never treated subdistricts. All standard errors are clustered at the district level. The vertical red line indicates between which periods the program was first introduced.

Chapter 3

Political Activists as Free-Riders: Evidence from a Natural Field Experiment

with **Anselm Hager**, University of Konstanz, **Johannes Hermle**, University of California - Berkeley, and **Christopher Roth**, Institute on Inequality and Beahvior

Abstract How does a citizen's decision to participate in political activism depend on the participation of others? We examine this core question of collective action in a natural field experiment in collaboration with a major European party during a recent national election. In a seemingly unrelated party survey, we randomly assign canvassers to true information about the canvassing intentions of their peers. Using survey evidence and behavioral data from the party's smartphone canvassing application, we find that treated canvassers significantly reduce both their canvassing intentions and behavior when learning that their peers participate more in canvassing than previously believed. These treatment effects are particularly large for supporters who have weaker social ties to the party, and for supporters with higher career concerns within the party. The evidence implies that effort choices of political activists are, on average, strategic substitutes. However, social ties to other activists can act as a force for strategic complementarity.

3.1 Introduction

Democracies rely on the efforts of political activists who participate in political campaigns. In the pursuit of improving collective outcomes, however, the individual action of an activist is unlikely to accomplish change. Instead, the impact of political activists hinges on the joint effort of the group, creating a situation of strategic interdependence: the effort of an activist might depend on the effort of her peers. Such strategic interdependence of individuals' actions is at the core of the collective action problem of political activism (Olson, 1965; Hardin, 2015). This paper presents a natural field experiment to investigate how the effort of a political activist depends on her belief about the participation of others.

In canonical models, political activism is viewed as a public goods game with incentives to free-ride (Olson, 1965): activists are motivated instrumentally, trading off private benefits against private costs of contributing. This class of models postulates that political activists reduce their effort when fellow activists contribute more to the public good. Or, put differently, the effort choices of activists are strategic substitutes.

In contrast, a large literature argues that participation in collective action is not mainly driven by instrumental concerns but instead by social motives leading to conditional cooperation (Ostrom, 2000; Falk and Fischbacher, 2006; Uhlaner, 1989). If such motives outweigh instrumental considerations, activists augment their effort in response to an increase in effort by their peers. Hence, activists' effort choices will exhibit strategic complementarity. Understanding the strategic nature of political activism is crucial to the refinement of the theoretical assumptions of models of collective action.

This paper presents the results of a natural field experiment to causally examine the strategic nature of political activism. In cooperation with a major political party in a Western European country, we implemented a pre-registered field experiment in the context of a large door-to-door canvassing campaign in the run-up to a nationwide general election. We examine whether and how party supporters' canvassing efforts depend on their beliefs about the canvassing efforts of their fellow party supporters. We focus on canvassing as it is an important form of political activism which directly influences political outcomes (Pons, 2014) and requires substantial time commitments from the activists.

Identifying the causal link between beliefs and behavior from correlational data faces the common challenges of causal inference. First, canvassers' effort choices might directly affect their beliefs, thereby giving rise to reverse causality. A concern in this context relates to motivated beliefs (Di Tella et al., 2015). For instance, canvassers might exaggerate the participation of others to exculpate their own lack of effort. Second, active canvassers might hold systematically different beliefs, potentially inducing omitted variable bias. For instance, party supporters who strongly believe in their party's goals might exert a high canvassing effort, but also overestimate the effort of their peers. In correlational data, these confounds could spuriously suggest either strategic substitutability or complementarity.

Our experimental strategy circumvents these confounds by exogenously manipulating beliefs in a natural field setting. Our design proceeds as follows. We use an unobtrusive survey distributed by the party via email eight weeks before the election with the stated purpose of gathering information to organize the campaign. In this survey, we first measure party supporters' ex-ante beliefs about the door-to-door canvassing intentions of their fellow party supporters. We then exogenously shift these beliefs in a treatment group by providing true information collected through a different survey conducted one month prior to the experiment. Supporters in a control group receive no such information. Subsequently, we elicit respondents' post-treatment beliefs about the actual canvassing turnout of their fellow party supporters. Finally, we elicit respondents' intention to go canvassing. After the survey, we collect unique, unobtrusive real-time data on canvassing behavior through a novel canvassing smartphone application in which door-to-door canvassers register the addresses they visit. To preserve the natural field setting we ensured that participants are at no point aware of their participation in an experiment.

We present four key results on the form, strength, and heterogeneity in the strategic interdependence of political activists' effort choices. First, on average, political activists' intentions follow the predictions of a public goods game with free-riding incentives: party supporters who learn that their peers plan to exert more effort than they previously expected significantly lower their intentions to participate in the party's campaign. The response is concentrated along the intensive margin. Supporters plan to canvass 1.10 days ($s.e. = 0.36$) less relative to a control mean of 4.03 days.

Second, we demonstrate that the reduction in canvassing intentions translates into a reduction in actual canvassing behavior. Using real-time canvassing data collected through the party's smartphone application, we estimate a reduction of 14.39 ($s.e. = 7.38$) canvassed doors, which is equivalent to a reduction of 38% relative to the control group mean of 38.35. Furthermore, we find a statistically significant reduction of 0.093 ($s.e. = 0.047$) standard deviations in a pre-specified index combining canvassing intentions and behavior. Our results thus imply that political activists' behavior exhibits strategic substitutability *on average*.

Who drives these treatment effects? The substantial heterogeneity in our sample allows us to investigate the underlying drivers of our treatment effects. Our third finding reveals that the effects are driven by party supporters with weaker social ties to the party (as proxied by prior canvassing experience, whether the respondent is a party member, and party membership duration). On the contrary, party supporters with strong ties do not exhibit a systematic pattern of strategic substitutability. Hence, these results highlight that social motives and connectedness can counterbalance free-rider incentives and act as a force for strategic complementarity.

Fourth, we demonstrate that the treatment effects are particularly large for party supporters with higher career concerns within the party and, thus, more concerns about signaling their commitment to the party. This finding is consistent with the predictions of models of signaling, where the marginal value of signaling decreases as others contribute more, as long as the fraction of individuals contributing is sufficiently low (Bénabou and Tirole, 2006).

Our study contributes to a growing body of literature investigating the motivation and behavior of political activists such as party supporters (Perez-Truglia and Cruces, 2017) and protesters (Cantoni et al., 2019; González, 2018; Enikolopov et al., 2019; Cantoni et al., 2016; Enikolopov et al., 2017; Passarelli and Tabellini, 2017; Manacorda and Tesei, 2018; Acemoglu et al., 2018; Madestam et al., 2013). Closest to our paper are Cantoni et al. (2019), who study protest participation in the context of the Hong Kong democracy movement. Their important study finds that beliefs about the turnout decision of protesters causally affect people's decision to participate in a protest. In line with our results, their findings provide evidence on strategic substitutability using self-reported protest participation. However, relative to their study and the existing literature, our setting and findings differ in several important ways.

First, our experimental data is distinct in two main respects: we draw on a unique combination of survey and behavioral outcome data collected through a smartphone application. This feature of the data allows us to study treatment effects on both self-stated intentions and actual behavior. Moreover, we provide evidence from a natural field experiment on behalf of the party, in which participants are not aware of their participation in an experiment. This feature reduces concerns about experimenter demand effects (de Quidt et al., 2018) and selection into the study, which commonly raises concerns about generalizability in non-natural field experiments (Harrison and List, 2004).

Second, in contrast to samples used in the previous literature, we leverage a heterogeneous sample of party supporters of all ages and with diverse backgrounds. The heterogeneity in our sample allows us to shed light on underlying drivers and mechanisms of our treatment effects. In particular, the heterogeneity in our data provides important evidence that strategic substitutability is most pronounced for party supporters with social ties to the party and those with higher career concerns.

Third, we study participation in a political campaign in a well-functioning democracy compared to the turnout to protests which previous literature has focused on. Compared to canvassing, protests may have different elements that will tend toward strategic substitutability: for example, the costs of turning out to a protest may increase in the perceived size of the protest due to an increased likelihood of a protest crackdown. This in turn makes the question of whether effort choices are strategic complements or substitutes conceptually distinct in the setting of canvassing. In addition, as we investigate the participation in a two month-long campaign compared to the turnout to a single protest, our setting also allows us to study responses along the intensive margin and dynamic effects over time. Importantly, in contrast to Cantoni et al. (2019), we do not find significant treatment effects along the extensive margin. Instead, supporters canvass on fewer days without completely abstaining from canvassing, thus reacting along the intensive margin.

We also contribute to the broader empirical literature examining whether and how beliefs affect political behavior (Gerber et al., 2019; Kendall et al., 2014; Cruz et al., 2019). While

previous studies have focused on voting (Bursztyn et al., 2018; Gerber et al., 2011; Green et al., 2013; DellaVigna et al., 2017), we examine canvassing which has previously been understudied.

Our findings also inform the theoretical literature investigating political behavior in democratic systems (Feddersen and Sandroni, 2006; Coate and Conlin, 2004; Downs, 1957; Palmfrey and Rosenthal, 1984). Our results provide evidence in line with the classical free-rider assumption in models of collective action (Olson, 1965). However, the systematic heterogeneity in the responsiveness to our treatment underscores that this assumption does not apply uniformly across the population. Instead, our findings highlight that theoretical models of political behavior have to account for heterogeneity in agents' motivation. This may be accomplished by introducing different types of agents, or by considering multiple motives simultaneously, both instrumental and social, as well as their interaction, as recently conceptualized by Jia and Persson (2017).

In this context, our results also highlight which assumptions on activists' motivation carry empirical relevance. In particular, our finding that stronger social ties within the party dampen strategic substitutability emphasizes the importance of social motives in counterbalancing free-rider incentives.[1] Furthermore, we also provide novel evidence on an additional form of party supporters' instrumental motivation: extrinsic career concerns within the party. Besides its relevance for the theoretical modeling of activists' motivation, this finding also holds practical relevance for party officials for the organization of political campaigns and the design of incentive schemes within the party.

Finally, we also contribute to the experimental literature on how strategic interaction affects public good provision in natural field settings (Fehr and Gächter, 2000a; Gallus, 2017). While our experiment concerns public goods at the nation-level, most field experiments on public good provision are in the domain of charitable giving (e.g. Frey and Meier, 2004; Shang and Croson, 2009) or the contribution to online communities (Chen et al., 2010). In contrast to our results, these field experiments tend to find patterns of strategic complementarity.

[1] This finding might also been seen as empirical support for the Uhlaner (1989) concept of relational goods, which are enjoyed only through the consumption of other group members by for instance enhancing a common identity.

This paper proceeds as follows: Section 3.2 provides a simple conceptual framework. Section 3.3 describes the setting and the experimental design. Section 3.4 characterizes pre-treatment beliefs about peer effort and belief updating in response to the treatment. Section 3.5 presents the main results on the impact of beliefs about peer effort on canvassing intentions and effort, followed by a conclusion in section 3.6.

3.2 Conceptual framework

Assumptions: To motivate our empirical design and guide our analysis, we present a simple partial-equilibrium model of an individual's decision to participate in canvassing. A canvasser's utility depends on her own canvassing effort, d_i, and her beliefs about fellow supporters' canvassing efforts, d_{-i}, according to

$$u_i(d_i) = (1-\alpha_i)g(d_i, d_{-i}) + \alpha_i h(d_i, d_{-i}) - c_i(d_i) - \tilde{c_i} \cdot \mathbb{1}(d_i > 0). \tag{3.1}$$

In this equation, $g(d_i, d_{-i})$ denotes an activist's instrumental utility from canvassing, while $h(d_i, d_{-i})$ represents the social utility gained from canvassing. Both $g(d_i, d_{-i})$ and $h(d_i, d_{-i})$ are assumed to be continuous and twice differentiable. Critically, both terms depend on own effort d_i but also on believed peer efforts d_{-i}. Instrumental and social utility have individual-specific relative weights $1-\alpha_i$ and α_i, respectively.

Costs of canvassing are captured by $c_i(d_i)$ with $c', c'' > 0$ and $c'(0) = 0$, and individual-specific fixed costs $\tilde{c_i} \sim F$. We assume that all idiosyncratic properties are summarized by type $i \in I$, distributed according to $\Phi \perp F$. Canvasser i chooses canvassing effort, d_i, to maximize her utility u_i given the reservation utility from not participating in canvassing $(1-\alpha_i)g(0, d_{-i})$.

We further clarify the properties of instrumental utility $g(d_i, d_{-i})$ and social utility $h(d_i, d_{-i})$. Instrumental utility $g(d_i, d_{-i})$ comprises two motives: first, the value of votes obtained through canvassed doors and, second, extrinsic career prospects within the party. In other words, activists may be motivated to participate in canvassing to signal their commitment to the party and thereby increase their chances of getting promoted. We thus assume that $\frac{\partial g(d_i, d_{-i})}{\partial d_i} > 0$. Critically, we also posit that the instrumental returns to canvassing are decreasing in peer effort (i.e. $\frac{\partial^2 g(d_i, d_{-i})}{\partial d_i \partial d_{-i}} < 0$). In the context of our experiment, this assumption is likely to hold for two reasons. First, instrumental returns to canvassing are decreasing in peer effort if marginal

benefits of additional votes or the returns to canvassing are decreasing, an assumption which is plausible in our setting. The party encouraged local canvassers to target the most promising areas first, which implies that the expected returns to additional canvassing are likely to decrease in the total canvassing activity.[2] Second, following the logic of models of signaling (Bénabou and Tirole, 2006), higher participation of others is likely to decrease the extrinsic signaling value of canvassing to indicate commitment to the party, which leads to an additional reduction in the instrumental return to d_i.[3]

The term $h(d_i, d_{-i})$ represents the social utility gained from canvassing for which we assume $\frac{\partial h(d_i,d_{-i})}{\partial d_i} \geq 0$. Most generally, this term captures the quality of canvassing as a relational good (Uhlaner, 1989) that is enjoyed only through the consumption of others. In line with a large literature on conditional cooperation in collective action (Ostrom, 2000), we hence assume that social returns to canvassing are increasing in peer effort, i.e. $\frac{\partial^2 h(d_i,d_{-i})}{\partial d_i \partial d_{-i}} > 0$. A specific psychological foundation for this assumption is provided by theories of reciprocity (Falk and Fischbacher, 2006; Fehr and Gächter, 2000b) in which individuals receive utility from rewarding kind actions of others.

Intensive margin responses: We first analyze intensive margin responses due to a change in beliefs about peer canvassing effort d_{-i}. The model yields that the strategic interaction between canvassers is determined by the relative importance of changes in the instrumental and social returns to canvassing.

- Effort choices will be **strategic complements ($\frac{\partial d_i^*}{\partial d_{-i}} > 0$)** iff $\alpha_i \frac{\partial^2 h(d_i,d_{-i})}{\partial d_i \partial d_{-i}} > -(1-\alpha_i)\frac{\partial^2 g(d_i,d_{-i})}{\partial d_i \partial d_{-i}}$, namely iff changes in social returns dominate changes in instrumental returns.

- Effort choices will be **strategic substitutes ($\frac{\partial d_i^*}{\partial d_{-i}} < 0$)** iff $-(1-\alpha_i)\frac{\partial^2 g(d_i,d_{-i})}{\partial d_i \partial d_{-i}} > \alpha_i \frac{\partial^2 h(d_i,d_{-i})}{\partial d_i \partial d_{-i}}$, namely iff changes in instrumental returns dominate changes in social returns.

Proofs of these results can be found in Appendix section 3.A.

[2] In addition, the party only gains little political power from winning votes beyond a threshold that ensures that it forms part of the government.

[3] In terms of the model developed by Bénabou and Tirole (2006), this holds for the lower tail of the distribution of prior beliefs about the participation of others where the 'virtue' channel dominates. This assumption seems suitable for this analysis as we mostly focus on individuals with relatively low beliefs about the participation of others.

Heterogeneity by strength of instrumental and social motives: How do the patterns of strategic interaction vary with the relative importance of social motives α_i? Intuitively, we show in Appendix section 3.A that individuals who put a sufficiently high weight on instrumental utility will exhibit strategic substitutability in their effort choices, whereas individuals who put a sufficiently high weight on social motives will exhibit strategic complementarity in their effort choices.

In the context of our experiment, we expect stronger social connectedness to the party to trigger a higher weight on social motives α_i, as stronger social connectedness plausibly increases the reciprocity to other party members and enhances identification with the party. We test this prediction empirically by estimating how the treatment effects vary with proxies such as having prior canvassing experience, being a member of the party, and the duration of membership. On the other hand, we expect individuals with stronger career concerns to give higher priority to extrinsic instrumental motives, and conversely to exhibit a lower relative weight α_i.

Extensive margin responses: Finally, we consider how extensive margin responses depend on individuals' beliefs about their peers canvassing effort. Note that an individual i will turn out canvassing if her cost of canvassing $\tilde{c}_i$ is lower than a cutoff value $\tilde{c}_i{}^*$. Hence, the fraction of individuals of type i deciding to participate in canvassing is equal to $F(\tilde{c}_i{}^*)$. We show in Appendix section 3.A that the marginal extensive margin response to an increase in d_{-i} is equal to

$$\Big((1-\alpha_i)\Big(\frac{\partial g(d_i^*, d_{-i})}{\partial d_{-i}} - \frac{\partial g(0, d_{-i})}{\partial d_{-i}}\Big) + \alpha_i \frac{\partial h(d_i^*, d_{-i})}{\partial d_{-i}}\Big) \cdot f(\tilde{c}_i{}^*),$$

where d_i^* is the optimal canvassing effort conditional on exerting positive canvassing effort. Note that the sign of the effect depends on the relative strength of changes in instrumental utility ($\frac{\partial g(d_i^*, d_{-i})}{\partial d_{-i}} - \frac{\partial g(d_{-i})}{\partial d_{-i}}$) and social utility ($\frac{\partial h(d_i^*, d_{-i})}{\partial d_{-i}}$). Most importantly, however, the effect size critically depends on the mass of marginal individuals $f(\tilde{c}_i{}^*)$. If $f(\tilde{c}_i{}^*)$ is low (high), the effect size on the extensive margin will be comparatively low (high) relative to the intensive margin. Differences in the effect sizes between the extensive and intensive margin may hence be interpreted as indicators of the mass of marginal individuals.

3.3 Experimental Design

3.3.1 Design and Sample

Setting: Our field experiment took place in the run-up to a recent general election in a Western European country. The experiment was implemented in collaboration with a major political party to study party supporters' motivation and actual participation in the party's door-to-door canvassing campaign. The analysis was pre-registered at the AEA RCT registry before the start of the data collection. The experimental manipulation was administered in an online pre-campaign survey sent out on behalf of the party roughly eight weeks before the election. After the intervention, we tracked party supporters' real canvassing efforts throughout the campaign until the election.

The party with which we cooperated strongly promoted canvassing as a campaigning tool through internal communication channels. All canvassing volunteers were instructed to record every canvassed door in a novel smartphone application as a way to help the party organize this and future campaigns. The data from the application provide unique behavioral outcomes on actual, real-time canvassing efforts.

While the overall level of canvassing activity was higher than in previous elections, there was still substantial potential to increase activity levels. Nationally, volunteers of the party reached out to 1.65% of all households. At the constituency level, the fraction of households canvassed ranged between 0 and 25% with a median of 0.5% and a 90^{th} percentile of 4.5%. These low absolute levels of canvassing in most places imply that there was scope for volunteers to increase their level of canvassing activity even though that could imply going to less promising areas and thus lower returns to canvassing activity.

Sampling and Procedures: Our original sample comprises all party supporters who had signed up to the party's campaign email list about eight weeks before the election. At the beginning of the electoral campaign, we contacted these supporters with an email invitation on behalf of the party. The email asked supporters to participate in the survey to help organize the campaign. The invitation email was designed by the party to preserve the natural environment and ensure that participants would not be aware of being part of an experiment. A reminder email was sent ten days later. In total 1,411 party supporters responded to the online survey

for this experiment.[4] Random assignment and experimental manipulation took place within the online survey.[5] The natural field setting mitigates concerns about selection into the study, which threatens the external validity in non-natural field experiments (Harrison and List, 2004).

Measuring and Manipulating Beliefs: We designed the experiment to provide causal evidence on how party supporters' motivation and actual canvassing effort depend on their beliefs about the efforts of their peers.

The experimental design is illustrated in Figure 3.1. In a first step, we elicit participants' pre-treatment beliefs about the share of party members who plan to go canvassing. Then, half of all respondents are randomly assigned to receive information about the canvassing plans of fellow party members (treatment group), whereas the remaining half receives no information (control group). More specifically, participants in the treatment group are truthfully informed that 37% of party members in a previous survey had stated an intention to go canvassing.[6] After the experimental manipulation, all respondents are asked to estimate the share of members who will actually go canvassing.[7] The elicitation of participants' post-treatment beliefs allows us to check whether the information provision successfully shifts beliefs.

Outcome Measures: We study the canvassing effort of party supporters in the campaign by combining both survey and behavioral outcome data. We use two pre-specified self-reported measures of canvassing intentions that are collected after the treatment administration: First, we measure whether a respondent intends to do any canvassing in the campaign. This allows us to shed light on movement along the extensive margin. Second, we elicit respondents' intended number of days of participation, enabling us to analyze responsiveness to the treatment along the intensive margin.[8]

We then assess whether changes in canvassing intentions translate into changes in canvassing behavior. We draw on unique behavioral outcome data from the smartphone application

[4] We simultaneously conducted a second experiment which is pre-registered in the same pre-analysis plan used for this paper. Supporters responding to the invite were randomly allocated between the experiment described in this paper and the experiment described in Hager et al. (2019a) which studies how beliefs about the effort choices of members of the main competing political party affect activists' effort choices.

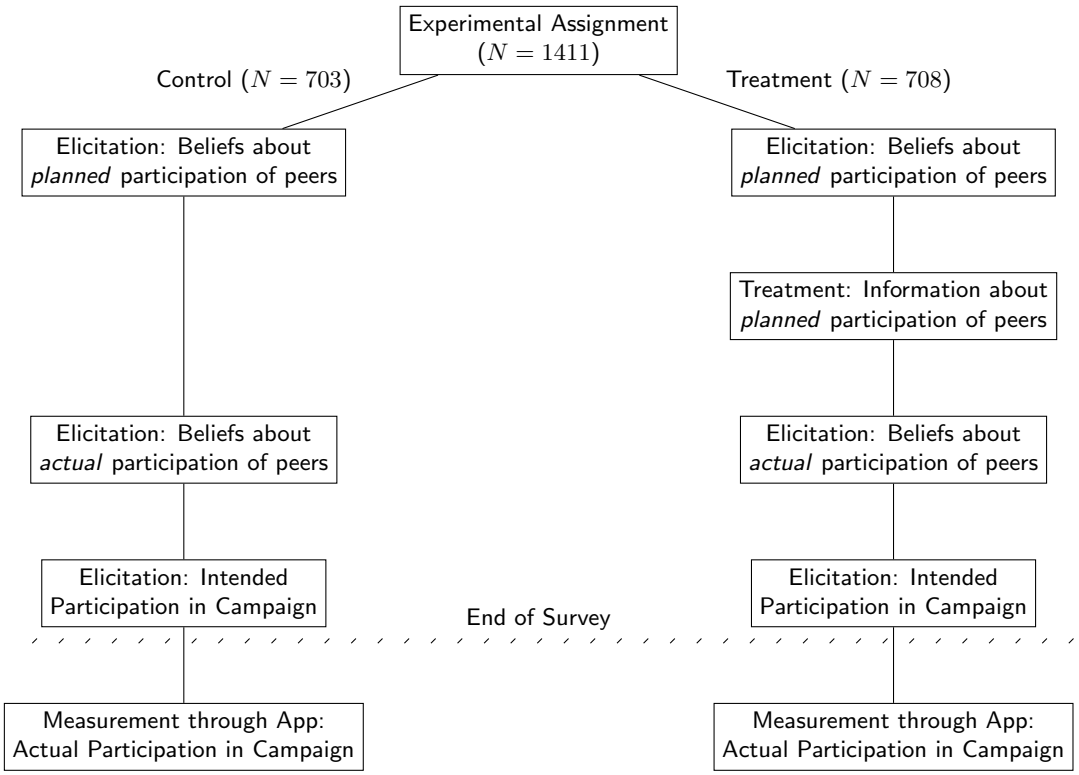
[5] We include all individuals who saw the treatment screen in our sample even when they did not complete the full survey as pre-specified in the pre-analysis plan. This leads to small variations in sample size across survey-based and behavioral outcomes.

[6] We collected this data in a separate survey amongst other party members three weeks before the experiment.

[7] We did not incentivize the belief elicitation to preserve the natural context of the survey. The organizers of the campaign were concerned that incentives would be perceived as very unusual by their supporters.

[8] The intended number of days for respondents who do not plan to canvass is coded as zero days.

Figure 3.1: Experimental design



Notes: Figure 3.1 illustrates the experimental design.

distributed by the party. The application allows us to assess three pre-specified behavioral outcomes: first, an indicator for whether a supporter knocks on any doors; second, the number of doors a supporter knocks on;[9] and third, the number of days a supporter goes canvassing.[10]

3.3.2 Descriptives and Balance

Panel A of Table 3.1 describes the sample characteristics elicited in the survey. 24% of supporters are women, and the average age is 41 years. The gender composition of our sample is comparable to the population of party members but supporters are significsntly younger than the average party member.[11] 83% of supporters are party members with an average membership duration of 12 years. Besides basic socio-demographic information, the survey also inquires about supporters' prior canvassing experience. 38% of participants had helped in a past campaign.

[9] We pre-specified to winsorize this variable at the 99 percentile to deal with outliers.

[10] Individuals who do not appear in the application data are coded as not having canvassed.

[11] This difference is possibly the result of the party's centralized mobilization effort Which relied on online technology geared towards younger supporters.

Table 3.1: Summary statistics: full sample

	Mean	SD	Median	Min.	Max.	Obs.
Panel A: pooled sample						
Predetermined variables						
Female	0.24	0.43	0.00	0.00	1.00	1411
Age	41.04	19.30	36.00	16.00	100.00	1411
Social ties to party (z-scored)	-0.00	1.00	-0.06	-1.76	2.42	1411
Is party member	0.83	0.38	1.00	0.00	1.00	1411
Years of party membership	12.19	14.22	6.00	0.00	60.00	1411
Has experience canvassing	0.38	0.48	0.00	0.00	1.00	1411
Participated in door-to-door workshop	0.21	0.41	0.00	0.00	1.00	1411
Downloaded app before survey	0.27	0.44	0.00	0.00	1.00	1411
Has canvassed before survey	0.08	0.27	0.00	0.00	1.00	1411
Days canvassed before survey	0.19	0.96	0.00	0.00	16.00	1411
Doors visited before survey	4.38	41.80	0.00	0.00	1071.00	1411
Intention outcomes						
Canvassing: yes	0.49	0.50	0.00	0.00	1.00	1395
Canvassing: days	3.85	7.33	0.00	0.00	60.00	1395
Behavioral outcomes						
Has canvassed after survey	0.12	0.33	0.00	0.00	1.00	1411
Days canvassed after survey	0.59	2.70	0.00	0.00	40.00	1411
Doors canvassed after survey	29.22	137.57	0.00	0.00	1045.00	1411
Panel B: underestimators						
Predetermined variables						
Female	0.21	0.41	0.00	0.00	1.00	1163
Age	41.31	19.04	36.00	16.00	100.00	1163
Social ties to party (z-scored)	0.07	0.98	0.01	-1.76	2.42	1163
Is party member	0.85	0.36	1.00	0.00	1.00	1163
Years of party membership	13.00	14.53	7.00	0.00	60.00	1163
Has experience canvassing	0.39	0.49	0.00	0.00	1.00	1163
Participated in door-to-door workshop	0.21	0.41	0.00	0.00	1.00	1163
Downloaded app before survey	0.26	0.44	0.00	0.00	1.00	1163
Has canvassed before survey	0.08	0.27	0.00	0.00	1.00	1163
Days canvassed before survey	0.20	1.02	0.00	0.00	16.00	1163
Doors visited before survey	5.06	45.87	0.00	0.00	1071.00	1163
Intention outcomes						
Canvassing: yes	0.48	0.50	0.00	0.00	1.00	1148
Canvassing: days	3.48	6.56	0.00	0.00	60.00	1148
Behavioral outcomes						
Has canvassed after survey	0.12	0.32	0.00	0.00	1.00	1163
Days canvassed after survey	0.62	2.86	0.00	0.00	40.00	1163
Doors canvassed after survey	30.80	143.39	0.00	0.00	1045.00	1163
Panel C: overestimators						
Predetermined variables						
Female	0.37	0.48	0.00	0.00	1.00	248
Age	39.74	20.48	33.50	16.00	92.00	248
Social ties to party (z-scored)	-0.33	1.01	-0.42	-1.76	2.24	248
Is party member	0.71	0.46	1.00	0.00	1.00	248
Years of party membership	8.41	12.01	2.00	0.00	55.00	248
Has experience canvassing	0.32	0.47	0.00	0.00	1.00	248
Participated in door-to-door workshop	0.21	0.41	0.00	0.00	1.00	248
Downloaded app before survey	0.30	0.46	0.00	0.00	1.00	248
Has canvassed before survey	0.09	0.28	0.00	0.00	1.00	248
Days canvassed before survey	0.17	0.63	0.00	0.00	5.00	248
Doors visited before survey	1.21	8.07	0.00	0.00	99.00	248
Intention outcomes						
Canvassing: yes	0.52	0.50	1.00	0.00	1.00	247
Canvassing: days	5.57	10.01	2.00	0.00	60.00	247
Behavioral outcomes						
Has canvassed after survey	0.15	0.36	0.00	0.00	1.00	248
Days canvassed after survey	0.46	1.79	0.00	0.00	21.00	248
Doors canvassed after survey	21.85	106.02	0.00	0.00	1045.00	248

Notes: Table 3.1 presents summary statistics. Panel A contains statistics for the full sample of our experiment. Panel B contains statistics for respondents who underestimate the share of fellow party members who plan to participate in the party's door-to-door canvassing campaign. Panel C contains statistics for respondents who overestimate the share of fellow party members who plan to participate in the party's door-to-door canvassing campaign.

Regarding our outcome variables, 49% of supporters intend to participate in door-to-door

canvassing with an average of 3.85 intended days. Turning to behavioral outcomes from the app, we observe much lower actual canvassing activity relative to stated canvassing intentions: 12% of party supporters in our sample actually participate in the campaign. The respondents canvass on average 0.59 days and knock on 29 doors. The unique link between the survey and the behavioral outcome data from the natural field setting also allows us to study how intentions and actual canvassing behavior are related. We find a sizable positive correlation between intended days and the actual number of days of canvassing (ρ =0.28 , visualized in Figure A3.1). Similarly, we find that people's intention to do any canvassing is significantly related to whether they actually canvass (ρ =0.33).

These sizable and statistically significant correlations indicate that intentions are predictive of subsequent actual behavior. Given that these correlations do not equal one, however, highlights that self-stated intentions and actual behavior of activists cannot be equated. This feature underscores the necessity of collecting behavioral outcomes in addition self-stated intentions in order to examine the strategic behavior of political activists.

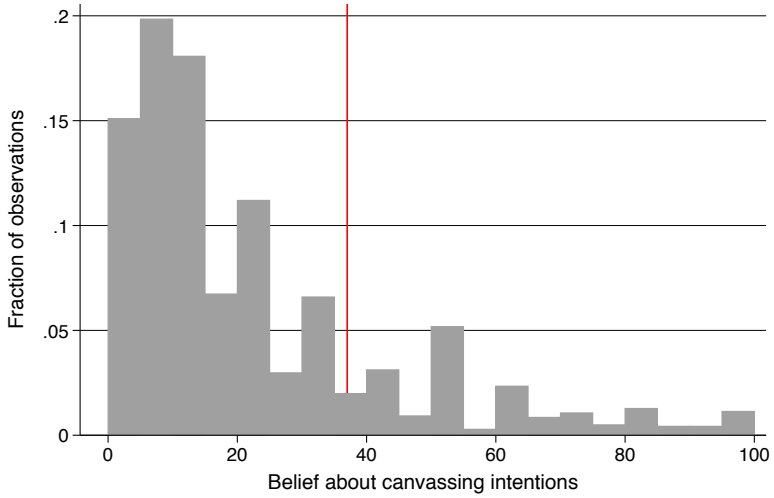
Finally, in terms of balance, we do not observe significant differences between the treatment and control group for any of the covariates (Table A3.1). We regress the treatment indicator on all covariates to test for joint significance. The p-value of this joint F-test is 0.59 indicating that the randomization produces two highly comparable groups.

3.4 Belief Updating

Before turning to the analysis of treatment effects on canvassing outcomes, we test for the successful manipulation of beliefs about peer canvassing effort.

Pre-treatment Beliefs: Figure 3.1 plots the distribution of pre-treatment beliefs about the percentage of party members who intend to go canvassing. The vertical red line corresponds to the treatment information which indicates that 37% of party members intend to go canvassing. We observe that the distribution is highly right skewed with a median belief of 10% of party members planning to go canvassing. Relative to the treatment information, 82% of participants underestimate their peers' canvassing intentions.

Figure 3.1: Distribution of pre-treatment beliefs



Notes: Figure 3.1 shows a histogram of pre-treatment beliefs about the fraction of party members who plan to participate in the party's door-to-door canvassing campaign. The vertical red line (37%) corresponds to the treatment information.

Belief Updating: The key qualification of our experimental design is that participants update their beliefs after receiving the treatment information. In particular, we expect underestimators to increase their post-treatment belief about the fraction of fellow party members who actually go canvassing. Vice versa, we expect overestimators to decrease their post-treatment beliefs.

Figure 3.2 shows the distribution of the difference between pre- and post-treatment beliefs over the the pre-treatment belief distribution. We see that underestimators in the treatment group update more positively than control group underestimators. The reverse is true for overestimators. Treated overestimators update more negatively compared to overestimators in the control group. We conclude that the information provision successfully shifted participants' beliefs about the canvassing effort of their peers.

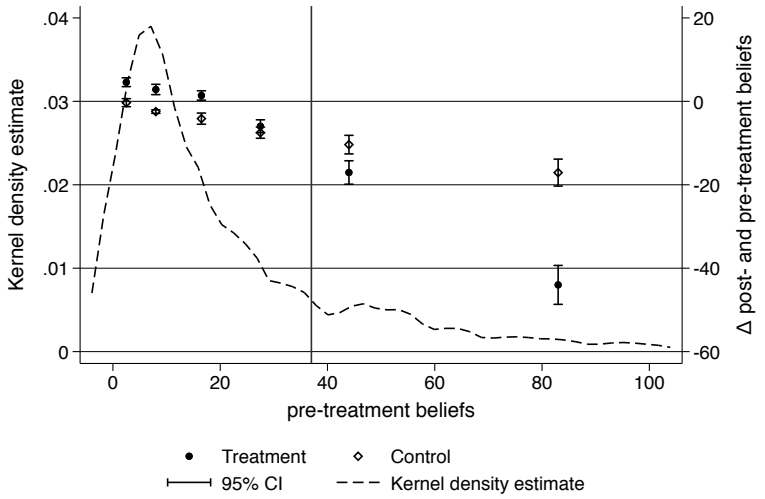
3.5 Results

3.5.1 Empirical Specification

Do these exogenous changes in beliefs affect party supporters' motivation and actual behavior in the campaign? We tackle this question in this section.

In our analysis we focus on participants who underestimate the share of fellow party members who plan to go canvassing. As only 18 percent of our respondents overestimate this statistic, we

Figure 3.2: Differences in beliefs between treatment and control group



Notes: Figure 3.2 shows the average difference between post- and pre-treatment beliefs about the fraction of party members who plan to go canvassing for the treatment and control group. Respondents are grouped according to treatment beliefs in six bins. Respondents who underestimate the treatment information are grouped in quartiles. Respondents who underestimate the treatment information are grouped in below and above median pre-treatment beliefs. The vertical line (37%) corresponds to the true treatment information, which is based on survey evidence.

lack the statistical power to estimate precisely measured treatment effects for overestimators.[12] We estimate the following specification using ordinary least squares:

$$Y_i = \beta_0 + \beta_1 \mathsf{T}_i + \zeta^T \mathbf{X_i} + \varepsilon_i \tag{3.2}$$

where Y_i is the outcome variable of interest. T_i is a dummy variable taking a value of one for people who receive the treatment information and zero otherwise. $\mathbf{X_i}$ is a set of pre-specified control variables: party membership, number of years of party membership, age, sex, whether a participant has already participated in a canvassing training, whether a participant has already downloaded the online application, whether a participant has participated in canvassing before this federal election, and whether a participant has participated in canvassing for this federal election.[13] To account for multiple comparisons, we also examine the effects on a pre-specified index as a joint measure of all five self-reported survey measures and behavioral outcomes jointly.[14] The key coefficient is β_1 which captures the strategic interaction between one's own

[12] Panel B and C of Table 3.1 provide descriptive statistics for under- and overestimators, respectively. We report the results for overestimators in Online Appendix Tables A3.12 to A3.20.

[13] Excluding control variables leads to similar results (Online Appendix Tables A3.21 to A3.29).

[14] The index takes into account (i) an indicator for whether a participant plans to go canvassing, (ii) the number of days that a participant plans to go canvassing, (iii) an indicator for whether a participant knocks on any door,

and peer effort. As we focus on underestimators, if $\beta_1 < 0$, own and peer effort exhibit strategic substitutability, if $\beta_1 > 0$ they exhibit strategic complementarity.

3.5.2 Main Results

Table 3.2 presents the main results. Participants who are informed that fellow party members devote more effort than previously thought, on average, reduce their willingness to participate in the party's campaign. This effect prevails on the intensive margin: participants intend to canvass 1.1 days ($s.e. = 0.36$) less after receiving the treatment information, which is equivalent to a reduction of 27% relative to the control group mean of 4 days. There is, however, no significant extensive margin effect on the dummy for the intention to engage in any canvassing, with a point estimate close to zero (0.002, $s.e. = 0.026$). Hence, party supporters intend to canvass less without abstaining from canvassing altogether.

Table 3.2: Main effects

	Posterior	Intentions		App Data			Index	App: Week 7/8	
	Belief	Any	Days	Any	Days	Door	Overall	Days	Doors
Treatment	5.027***	0.002	-1.098***	-0.013	-0.159	-14.388*	-0.093**	-0.145**	-6.657**
	(0.543)	(0.026)	(0.361)	(0.016)	(0.159)	(7.839)	(0.047)	(0.064)	(2.814)
Control mean	9.083	0.473	4.028	0.123	0.701	38.348	-0.029	0.297	13.318
Observations	1150	1148	1148	1163	1163	1163	1148	1163	1163

Notes: Table 3.2 presents treatment effects for respondents who underestimate the share of fellow party members who plan to participate in the party's door-to-door canvassing campaign. Outcome variables are as follows. *"Posterior"* captures the belief about the fraction of fellow party members who actually go canvassing. *"Intentions"* captures whether a participant intends to engage in any canvassing (*Any*) and the intended number of days (*Days*). *"App Data"* captures whether a participant actually engages in any canvassing (*Any*), as well as the number of days (*Days*) and number of doors a participant knocks on (*Door*). *"Index"* indicates a pre-specified index of all five outcome variables capturing canvassing intentions and behavior jointly. *"App: Week 7/8"* measures days and doors canvassed in weeks 7 and 8 after the treatment (one or two weeks before the election). All specifications but the last two columns are pre-specified in the pre-analysis plan. Treatment effects are obtained conditional on pre-specified control variables: party membership, number of years of party membership, age, sex, whether a participant has participated in a canvassing training, whether a participant has already downloaded the online application, whether a participant has participated in canvassing before this federal election and whether a participant has participated in canvassing for this federal election. $^{*}\ p < 0.10$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$

Does lower *intended* canvassing translate into lower *actual* canvassing? At the intensive margin we again find a significant reduction of 14.39 canvassed doors ($s.e. = 7.8$). This is

(iv) the number of doors that a participant knocks on, and (v) the number of days a participant goes canvassing. We construct the index by first standardizing each outcome using the control group mean and standard deviation, then calculating the total of the standardized variables, and finally re-standardizing the sum to have mean zero and standard deviation of one.

equivalent to a 38% reduction relative to the control group mean of 38.35 doors. Similarly, the point estimate on the impact on actual days canvassed indicates a reduction of 0.16 canvassed days ($s.e. = 0.16$), corresponding to a sizable, yet not statistically significant 23% reduction relative to the control group mean.[15] In line with the results on intentions, we again do not find an extensive margin effect.[16] The non-response on the extensive margin squares with the findings by Perez-Truglia and Cruces (2017) who examine interdependency in monetary campaign contributions. They find that party supporters contribute lower amounts at the intensive margin if they are informed that more of their peers contribute without changing behavior at the extensive margin.

Finally, we investigate the impact on the pre-specified index of all five outcomes capturing intentions and actual behavior jointly. We observe a decrease of 0.093 ($s.e. = 0.047$) standard deviations in this summary measure of canvassing intentions and behavior.[17] Taken together, these results indicate that, on average, increases in supporters' beliefs about their peers' efforts decrease both canvassing intentions and behavior. As such, our findings provide evidence on strategic substitutability in political activists' effort choices.[18]

Treatment Effects over Time: Did our treatment have a lasting impact or did we only temporarily affect behavior and intentions? One might suspect that the treatment effects are short-lived as treated individuals forget the provided information. In this case, the treatment effects should be driven by changes in behavior shortly after the treatment is administered. To investigate this issue systematically, panel (a) of Figure 3.1 shows the average number of doors knocked on separately for the treatment and control group, week-by-week for the 8-week period in the run-up to the election. Overall, the canvassing activity strongly increases as the election comes closer. However, a clear divergence between the control and treatment groups emerges,

[15] To show that our treatment effects are not driven by outliers, Figure A3.2 presents cumulative distribution functions of intended days of canvassing, actual days canvassing, and doors knocked on for treatment and control group.

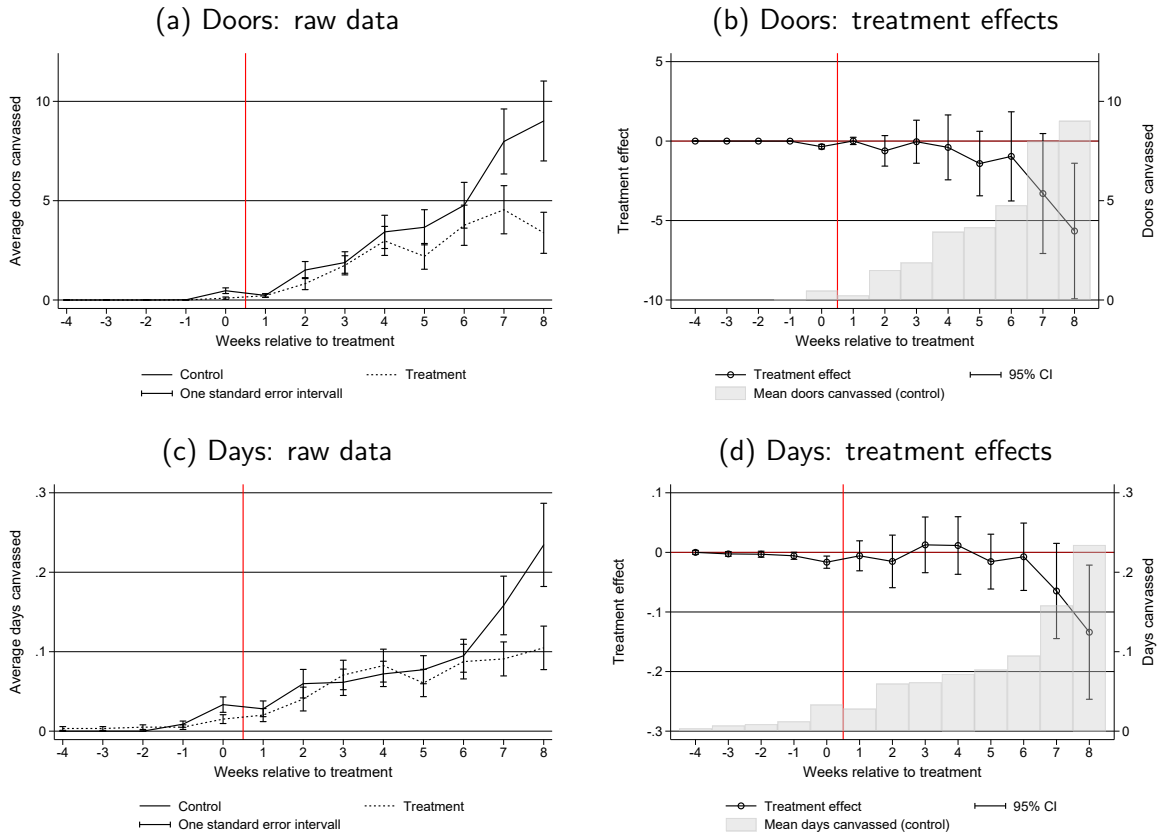
[16] Similarly, we do not find a significant effect on application download (Online Appendix Table A3.2), consistent with our findings on the extensive margin.

[17] All results are also significant when employing randomization inference (see Appendix Table A3.3).

[18] Among overestimators, we find imprecisely measured and small increases in canvassing effort in response to the information treatment. The effect is far from significant and we lack statistical power to draw firm conclusions for the subsample of overestimators. The positive effect, which is consistent with complementarity in effort choices, could arise from the fact that overestimators have in general larger social motives compared to underestimators and as they are less driven by instrumental concerns.

most noticeable in weeks seven and eight when the overall canvassing activity peaks.

Figure 3.1: Doors and days canvassed over time



Notes: Figure 3.1 (a) shows the average number of doors canvassed (winsorized at the 99th percentile) for each week after the treatment. Figure 3.1 (b) plots the estimates of treatment effects on doors canvassed (winsorized at the 99th percentile) for each week after the treatment. Estimates are obtained by estimating equation (3.2) separately for each week after the treatment. Pre-specified control variables include: party membership, number of years of party membership, age, sex, whether a participant has participated in a canvassing training, whether a participant has already downloaded the online application, whether a participant has participated in canvassing before this federal election and whether a participant has participated in canvassing for this federal election. Shaded bars indicate the average number of doors knocked on in the control group in a given week after the survey. Vertical red lines indicate the timing of treatment. The sample in both figures is restricted to respondents who underestimate the share of fellow party members who plan to participate in the party's door-to-door canvassing campaign. Figures 3.1 (c) and (d) present analogous evidence for days canvassed.

Panel (b) of Figure 3.1 shows treatment effects conditional on the pre-specified covariates obtained by estimating equation (3.2) separately for each week after the survey. As already suggested by the raw data, we observe a large and highly significant treatment effect in week eight after the survey at the time where overall canvassing activity peaks and we possess the statistical power to detect the treatment effects. The analogous pattern is also present when

we investigate the number of days canvassed in panels (c) and (d) of Figure 3.1. The last two columns of Table 3.2 show the results of our main regression analysis on days canvassed and doors knocked on in week seven and eight after the treatment. The estimates show statistically significant effect sizes relative to the control mean of roughly 50%.

The combination of an instantaneous impact on canvassing intentions with a long-term impact on behavioral outcomes indicates that our information provision persistently changed beliefs about fellow party members' participation in the party's campaign and produced long-lasting changes in canvassing behavior.

3.5.3 Heterogeneity in Treatment Effects

Heterogeneity by Social Ties to the Party: Who drives these treatment effects? In Section 3.2, we outlined a simple conceptual framework of canvassing effort. The model posits that individuals who put a relatively lower weight on social as compared to instrumental motives will exhibit a stronger pattern of strategic substitutability. Vice versa, for individuals with stronger social motives, the pattern of strategic substitutability should be weaker.

To test for heterogeneity by social motives, we use a variety of proxies for the strength of the social connection to the party. In the pre-analysis plan we had pre-specified analyzing heterogeneity by participants' (i) prior canvassing experience, (ii) whether they are a party member, and (iii) their years of party membership. To keep the analysis as parsimonious as possible, we first employ a simple principal component analysis (PCA) of these three variables to study heterogeneous treatment effects using the standardized first principle component of the three variables. Intuitively, we use the first principle component as an index capturing the strength of social ties to the party.

Before turning to the heterogeneity analysis, we validate this index as a well-suited measure capturing party supporters' social connectedness and social motives. We do so by investigating the predictive power of this index for three outcomes: (i) the number of party members a supporter knows personally, (ii) a supporter's identification with the party, and (iii) a supporter's identification with the party platform (Table A3.4). Using data from a post-election survey administered to a different sample of the same population of supporters, we find that supporters scoring above the median on the index measure know 23 ($p < 0.001$) more party members

than supporters below the median. Furthermore, they exhibit a 0.27 ($p < 0.001$) standard deviations higher identification with the party, and a 0.23 ($p < 0.001$) standard deviations higher identification with the party platform. Based on these results, we use the first principle component as a measure of party supporters' social connectedness and social motives in the following analysis.

In line with our hypothesis, supporters with weaker connections to the party exhibit a larger negative treatment effect relative to supporters with stronger connections (Table 3.3). The treatment effect heterogeneity is of statistical and economic significance for both intentions and behavior: treated supporters with one-standard deviation stronger connections to the party exhibit a reduction of the treatment effect on intended days of 0.76 $(s.e. = 0.37)$. The attenuation of the treatment effect on canvassing intentions translates into an attenuated treatment effect on canvassing behavior. Treated supporters with one standard deviation weaker connections to the party knock on 16.6 $(s.e. = 7.2)$ fewer doors. The complementing treatment effect heterogeneity for behavioral outcomes over time is illustrated by Figure 3.2. Finally, in terms of the summary index measuring canvassing intentions and behavior jointly, we find that treated supporters with one-standard deviation stronger connections to the party display a 0.084 $(s.e. = 0.045)$ standard deviation smaller decrease in canvassing. These results suggest that social connectedness and social motives can counterbalance strategic substitutability and act as a force for strategic complementarity in the effort choices of political activists.[19]

We also examine heterogeneous responses for each of the variables used in the PCA separately which further corroborates that supporters with weaker social ties to the party drive our treatment effects: treatment effects are driven by supporters without prior canvassing experience (Table A3.5), non-party members (Table A3.6) and for newer members (including non-members) (Table A3.7).

Heterogeneity by Career Concerns: Turning from social to instrumental motives, we next examine whether party supporters driven by instrumental motives are responsible for our treatment

[19] An alternative mechanism might be that supporters with weaker social connections to the party update their post-treatment beliefs more strongly. However, we do not find any differential treatment effects on the post-treatment belief (column 1, Table 3.3).

Table 3.3: Treatment effects and strength of connection to the party

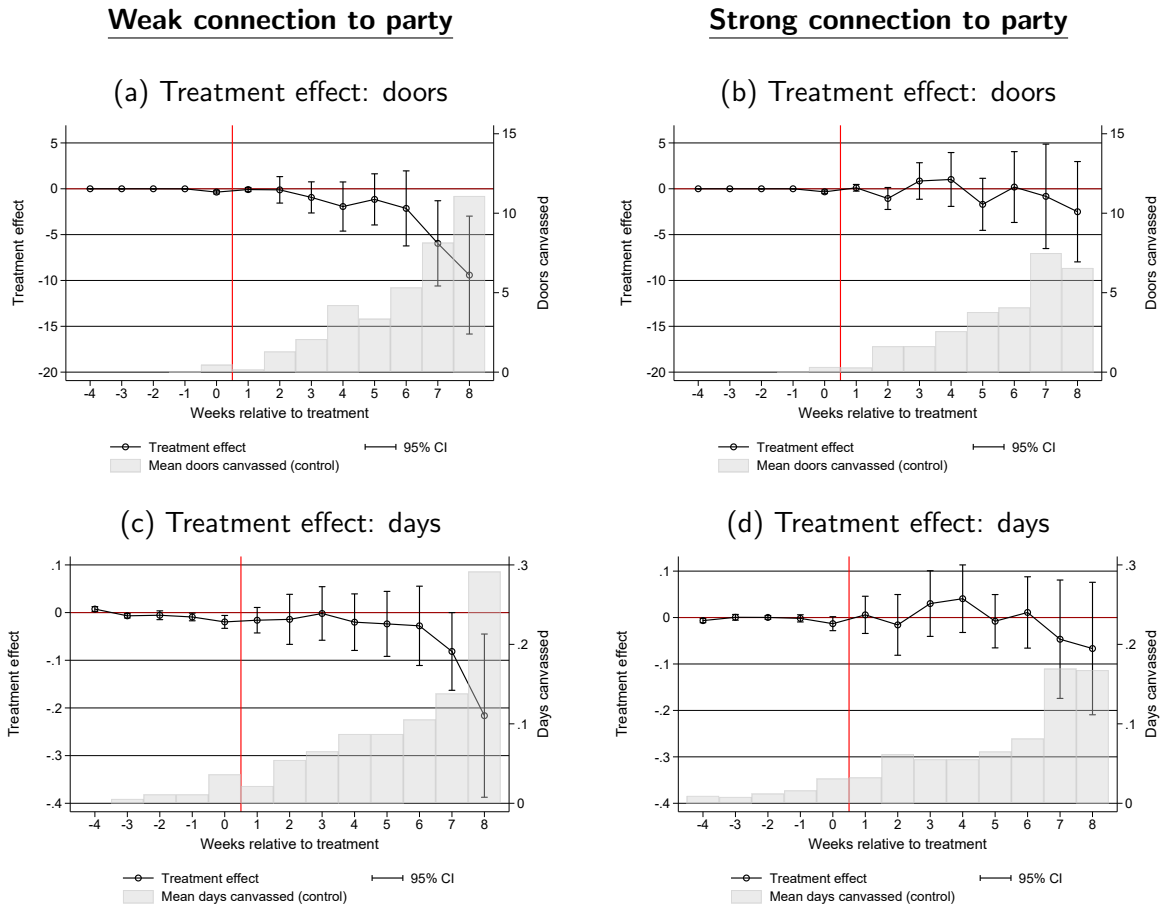
	Posterior	Intentions		App Data			Index	App: Week 7/8	
	Belief	Any	Days	Any	Days	Door	Overall	Days	Doors
Treatment	4.962***	0.001	-1.152***	-0.013	-0.168	-15.529*	-0.099**	-0.149**	-7.021**
	(0.556)	(0.026)	(0.364)	(0.016)	(0.157)	(7.920)	(0.047)	(0.064)	(2.828)
Treatment ×	0.909	0.024	0.756**	-0.000	0.141	16.661**	0.084*	0.060	5.310**
Strength of connection to party	(0.675)	(0.026)	(0.372)	(0.013)	(0.149)	(7.176)	(0.045)	(0.066)	(2.381)
Strength of connection to party	-2.151***	0.267***	1.199**	-0.012	-0.266	-19.693	0.118	-0.114	-6.506
	(0.754)	(0.039)	(0.579)	(0.023)	(0.251)	(12.318)	(0.074)	(0.109)	(4.394)
Control mean	9.083	0.473	4.028	0.123	0.701	38.348	-0.029	0.297	13.318
Observations	1150	1148	1148	1163	1163	1163	1148	1163	1163

Notes: Table 3.3 presents treatment effects interacted with a summary measure of connectedness to the party for respondents who underestimate the share of fellow party members who plan to participate in the party's door-to-door canvassing campaign. The measure for strength of connection to the party is the first principal component of three pre-specified dimensions of heterogeneity (previous canvassing experience, party membership and party membership duration). This summary measure is standardized to have mean zero and standard deviation one. Outcomes and pre-specified controls are as defined in Table 3.2. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

effects as hypothesized in our conceptual framework. In this regard, we analyze heterogeneous treatment effects by individuals' career concerns within the party as a specific form of extrinsic, instrumental motivation. As we could not ask respondents directly about their career concerns, we use a post-LASSO technique to predict career concerns based on data on career concerns from the post-election survey administered to the same population of supporters, but excluding individuals who also participated in the main experiment. We define individuals as having high career concerns if they answer "It is possible" to at least one of the following questions: "Can you imagine yourself taking on a political position within the party?" and "Can you imagine yourself running for office for this party?".[20] Based on the survey responses, we predict values of career concerns for participants in our experiment using a post-LASSO algorithm (Belloni and Chernozhukov, 2013). This procedure involves three steps. First, we use a LASSO algorithm to select the subset of best predictors among a large set of predetermined variables and all their pairwise interactions. Next, we perform an OLS regression of the dependent variable on the selected variables. We then use the obtained OLS coefficients to predict social connections for participants in our experiment based on their characteristics (see Table A3.8). We then standardize the variable to have mean zero and standard deviation one to facilitate interpretation. Using the standardized measure, we investigate the heterogeneity by predicted career concerns.

[20] The answer options were "No", "It is possible", and "I am already doing this".

Figure 3.2: Treatment effects over time by connection to the party



Notes: Figure 3.2 shows treatment effects over time separately for respondents with below median social connections to the party (panels (a) and (c)) and above median social connections to the party (panels (b) and (d)). The measure for strength of connection to the party is the first principal component of three pre-specified dimensions of heterogeneity (previous canvassing experience, party membership and party membership duration). Figures 3.2 (a) and (b) plot the estimates of treatment effects on doors canvassed (winsorized at the 99th percentile) for each week after the treatment. Figures 3.2 (c) and (d) present analogous evidence for days canvassed. The sample in all figures is restricted to respondents who underestimate the share of fellow party members who plan to participate in the party's door-to-door canvassing campaign. All estimates are obtained by estimating equation (3.2) separately for each week after the treatment. Pre-specified controls are as defined in Figure 3.1. Shaded bars indicate the average number of doors knocked on in the control group in a given week after the survey. Vertical red lines indicate the timing of treatment.

Table 3.4 provides evidence that strategic substitutability is stronger for respondents with a higher degree of predicted career concerns. Treated supporters with one-standard deviation higher career concerns display a 0.097 standard deviation ($s.e. = 0.045$) larger decrease in canvassing as measured by the summary index and knock on 18.2 fewer doors ($s.e. = 6.98$). The complementing treatment effect heterogeneity over time by career concerns is illustrated in Figure 3.3. The large negative treatment effects for supporters with higher career concerns is

consistent with dampened instrumental signaling returns: higher participation of others reduces the signaling value of exerting effort to indicate commitment to the party.[21]

Table 3.4: Treatment effects and career concerns

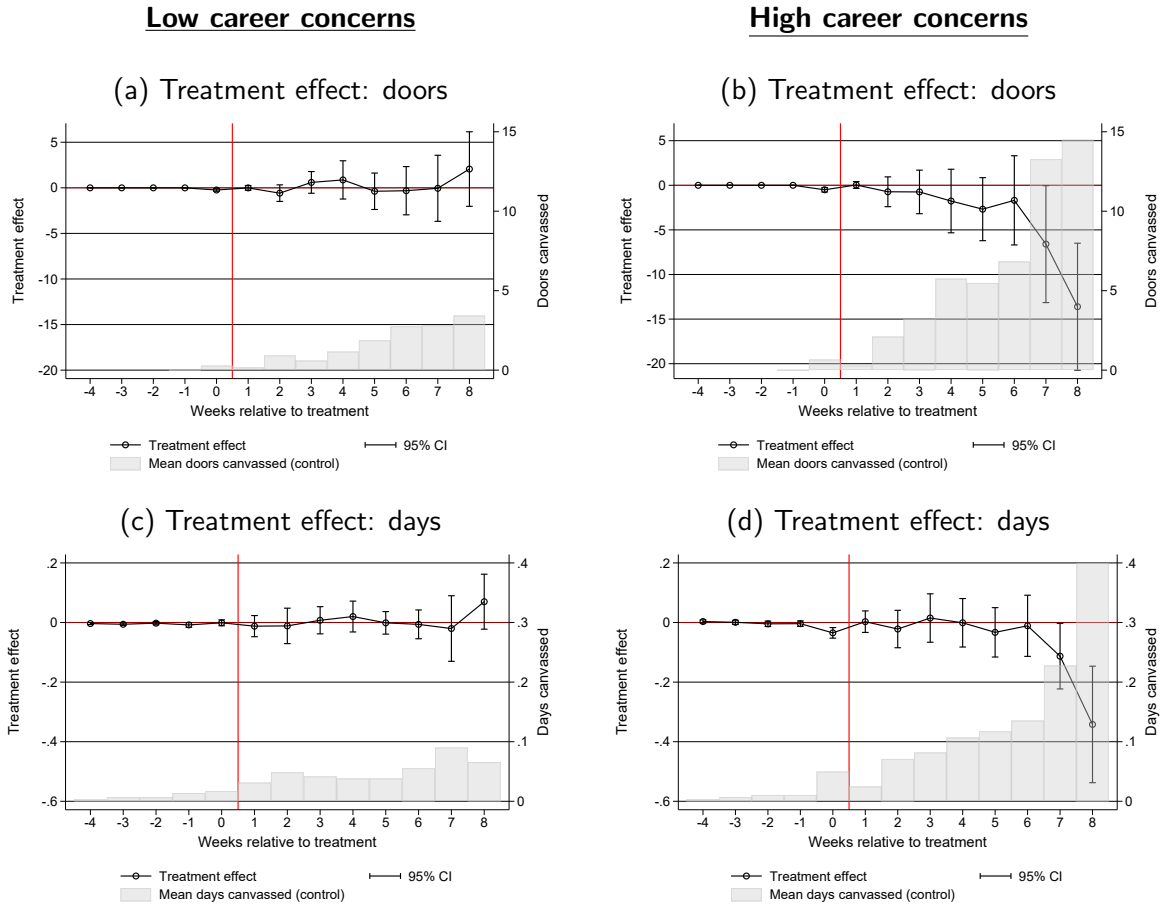
	Posterior	Intentions		App Data			Index	App: Week 7/8	
	Belief	Any	Days	Any	Days	Door	Overall	Days	Doors
Treatment	5.108***	0.001	-1.108***	-0.013	-0.165	-14.862*	-0.096**	-0.150**	-6.909**
	(0.541)	(0.026)	(0.361)	(0.016)	(0.161)	(7.915)	(0.048)	(0.064)	(2.847)
Treatment × Predicted career concerns	-0.644	-0.034	-0.726*	-0.002	-0.167	-18.207***	-0.097**	-0.124**	-6.598***
	(0.541)	(0.026)	(0.370)	(0.015)	(0.135)	(6.975)	(0.045)	(0.051)	(2.244)
Predicted career concerns	2.789***	-0.015	0.351	-0.005	-0.007	4.358	0.010	-0.018	-0.454
	(0.707)	(0.035)	(0.628)	(0.025)	(0.269)	(11.414)	(0.073)	(0.082)	(3.684)
Control mean	9.083	0.473	4.028	0.123	0.701	38.348	-0.029	0.297	13.318
Observations	1150	1148	1148	1163	1163	1163	1148	1163	1163

Notes: Table 3.4 presents treatment effects interacted with a supporter's career concerns for respondents who underestimate the share of fellow party members who plan to participate in the party's door-to-door canvassing campaign. We do not directly observe career concerns for participants in our experiment. We therefore predict career concerns using a post-election survey administered two months after the election to the same population of party supporters (excluding individuals who participated in the main experiment). Career concerns are measured as choosing "It is possible" as an answer to at least one the following questions "Can you imagine yourself taking on a political position within the party?" and "Can you imagine yourself running for office for this party?" (the answer options were "No", "It is possible", and "I am already doing this"). We use a post-LASSO algorithm to predict career concerns for our experimental sample based on pre-determined covariates (Belloni and Chernozhukov, 2013). The predicted variable is standardized to have mean zero and standard deviation one. The results of the post-LASSO regression are displayed in Table A3.8. Outcomes and pre-specified controls are as defined in Table 3.2. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

In sum, our results are consistent with the prediction that the overall treatment effects are driven by individuals with higher instrumental motivation and weaker social ties to the party. There is substantial heterogeneity in the extent of strategic substitutability even among supporters of the same party. Depending on the relative strength of instrumental and social motives, the degree of strategic substitutability and potentially even the qualitative nature of the strategic interaction vary between individuals.

[21] An alternative mechanism might be that supporters with higher career concerns update their post-treatment beliefs more strongly. However, we do not find any differential treatment effects on the post-treatment belief (column 1, Table 3.4).

Figure 3.3: Treatment effects over time by career concerns



Notes: Figure 3.3 shows treatment effects over time separately for respondents with below median career concerns (panels (a) and (c)) and above median career concerns (panels (b) and (d)). We do not directly observe career concerns for participants in our experiment. We therefore predict career concerns using a post-election survey administered two months after the election to the same population of party supporters (excluding individuals who participated in the main experiment).
Career concerns are measured as choosing "It is possible" as an answer to at least one the following questions "Can you imagine yourself taking on a political position within the party?" and "Can you imagine yourself running for office for this party?" (the answer options were "No", "It is possible", and "I am already doing this"). We use a post-LASSO algorithm to predict career concerns for our experimental sample based on pre-determined covariates (Belloni and Chernozhukov, 2013). Figures 3.3 (a) and (b) plot the estimates of treatment effects on doors canvassed (winsorized at the 99th percentile) for each week after the treatment. Figures 3.3 (c) and (d) present analogous evidence for days canvassed. The sample in all figures is restricted to respondents who underestimate the share of fellow party members who plan to participate in the party's door-to-door canvassing campaign. All estimates are obtained by estimating equation (3.2) separately for each week after the treatment. Pre-specified controls are as defined in Figure 3.1. Shaded bars indicate the average number of doors knocked on in the control group in a given week after the survey. Vertical red lines indicate the timing of treatment.

3.5.4 Spillovers into Alternative Effort Domains

Do party supporters substitute their reduced canvassing effort with increased effort in alternative contribution domains? This question pertains to the cross-substitutability between different forms of political activism. Activists might shift their effort between different contribution domains and activities to maximize their impact. In this case, different forms of political activism exhibit cross-substitutability. Alternatively, activists might be constrained or unwilling to respond with increased effort in alternative activities, indicating that different forms of political activism exhibit no cross-substitutability. To investigate this question, we leverage unique data from the party's application regarding activity on social media. Specifically, we investigate whether respondents who learn that fellow supporters engage in more canvassing become more likely to share news stories pertaining to their party on Facebook through the application.

We find that respondents who learn that fellow supporters engage in more canvassing than previously believed do not respond with an increase in sharing of party news stories (Table A3.9). The data indicate that treated supporters show a close-to-zero decrease of 1.1%-points ($s.e. = 1.4\%$) in their likelihood to share any news story as well as a decrease of 0.148 ($s.e. = 0.22$) in the average number of news stories shared. This evidence suggests that there is no cross-substitutability between supporters' canvassing behavior and their activity on social media pertaining to the party.

3.5.5 Interpreting Effect Sizes

IV Estimates: Finally, we assess the quantitative impact of beliefs about the effort of fellow party members on canvassing effort. To answer this question we employ the following instrumental variable specification:

$$\begin{aligned} y_i &= \pi_0 + \pi_1 \widehat{\mathsf{PB}}_i + \zeta^T \mathbf{X_i} + \varepsilon_i \\ \mathsf{PB}_i &= \kappa_0 + \kappa_1 \mathsf{T}_i + \xi^T \mathbf{X_i} + \vartheta_i \end{aligned}$$

where PB_i is a respondent's post-treatment belief about the fraction of party members who actually go canvassing and all other variables are defined as before. We instrument the post-treatment belief with the treatment indicator, ensuring that our estimates are solely identified from variation induced by the experimental treatment. Taking this estimate at face value, we

obtain the causal quantitative impact of beliefs on respondents' effort choices.

Table 3.5 presents the results. For the full sample (Panel A) we find that a one percentage point increase in the belief about the fraction of fellow party members who participate in canvassing leads to a decrease in planned days of 0.22 ($s.e. = 0.077$). In addition, we find a decrease of 3 ($s.e. = 1.6$) doors and a 0.02 ($s.e. = 0.01$) standard deviation decrease in the index. Again, the effect sizes are much stronger for party supporters with a weak connection to the party (Panel B) with a decrease of 0.35 ($s.e. = 0.13$) intended days and 6.03 ($s.e. = 2.74$) doors per one percentage point belief change.[22] Similarly, supporters with high career concerns (Panel E) display a high quantitative magnitude of the treatment effect with a decrease of 0.39 ($s.e. = 0.13$) intended days and 7.13 ($s.e. = 3.21$) doors per one percentage point belief change. In contrast, supporters exhibiting a strong connectedness to the party exhibit a much weaker effect size (Panel C) and those with low career concerns (Panel D) show an effect size close to zero.

Comparing OLS with the IV Estimates: Do the IV estimates differ from correlational evidence based on OLS estimation? In Table A3.10, we present for the control group OLS regressions of the outcome data on the belief about the fraction of fellow party members who participate in canvassing. We find that the IV results stand in contrast to the OLS estimates which suggest moderate complementarity in effort choices. A one percentage point increase in the belief about peer effort is associated with 1.2 ($s.e. = 0.69$) more canvassed doors and a 0.007 ($s.e. = 0.004$) standard-deviation increase in the index.

Using a Hausman-style test for the exogeneity of beliefs about peer effort, we reject exogeneity at the 5% level for intended days, doors, and the index.[23] The differences in results between the OLS estimates and the IV estimates could be explained by omitted variable bias. For example, individuals who believe that canvassing is particularly effective might be more likely to both canvass and believe that others do the same. This divergent evidence highlights the methodological necessity of exogenous belief manipulation for the isolation of causal effects.

[22] We define supporters with a weak or strong connection to the party as those scoring below or above the median of the summary measure of connectedness to the party.

[23] The tests we use are defined as the difference of two Sargan-Hansen statistics: one for the equation without instruments, where the belief about peer effort is not instrumented, and one for the equation where it is instrumented by the treatment. Under the null hypothesis that the belief about peer effort is exogenous, the test statistic is distributed as chi-squared with one degree of freedom.

Table 3.5: IV estimates: post-treatment beliefs and effort

	Intentions		App Data			Index	App: Week 7/8	
	Any	Days	Any	Days	Door	Overall	Days	Doors
Panel A: Main Effect (IV)								
Posterior	0.000	-0.220***	-0.003	-0.033	-2.990*	-0.019*	-0.030**	-1.368**
	(0.005)	(0.077)	(0.003)	(0.032)	(1.623)	(0.010)	(0.013)	(0.590)
Observations	1148	1148	1163	1163	1163	1148	1163	1163
Panel B: Weak connection to party (IV)								
Posterior	-0.003	-0.346***	-0.002	-0.068	-6.034**	-0.033**	-0.046**	-2.495***
	(0.008)	(0.132)	(0.005)	(0.046)	(2.739)	(0.016)	(0.021)	(0.966)
Observations	547	547	554	554	554	547	554	554
Panel C: Strong connection to party (IV)								
Posterior	0.003	-0.116	-0.003	-0.002	-0.587	-0.006	-0.016	-0.465
	(0.007)	(0.093)	(0.004)	(0.045)	(2.093)	(0.013)	(0.018)	(0.784)
Observations	601	601	609	609	609	601	609	609
Panel D: Low career concerns (IV)								
Posterior	0.007	-0.059	-0.002	0.005	0.657	0.002	0.004	0.189
	(0.007)	(0.094)	(0.004)	(0.034)	(1.270)	(0.010)	(0.014)	(0.541)
Observations	547	547	554	554	554	547	554	554
Panel E: High career concerns (IV)								
Posterior	-0.005	-0.394***	-0.004	-0.079	-7.131**	-0.041**	-0.067***	-3.101***
	(0.007)	(0.127)	(0.005)	(0.055)	(3.210)	(0.017)	(0.024)	(1.142)
Observations	601	601	609	609	609	601	609	609

Notes: Table 3.5 presents IV estimates for the impact of post-treatment beliefs for respondents who underestimate the share of fellow party members who plan to participate in the party's door-to-door canvassing campaign (Panel A). Post-treatment beliefs about actual participation are instrumented with the treatment indicator. Outcomes and pre-specified controls are as defined in Table 3.2. Panels B and C show effects for supporters with above and below median strength of connection to the party, respectively. The measure for strength of connection to the party is the first principal component of three pre-specified dimensions of heterogeneity (previous canvassing experience, party membership and party membership duration). Panels D and E show effects for supporters with low and high career concerns, respectively. We do not directly observe career concerns for participants in our experiment. We therefore predict career concerns using a post-election survey administered two months after the election to the same population of party supporters (excluding individuals who participated in the main experiment).

Career concerns are measured as choosing "It is possible" as an answer to at least one the following questions "Can you imagine yourself taking on a political position within the party?" and "Can you imagine yourself running for office for this party?" (the answer options were "No", "It is possible", and "I am already doing this"). We use a post-LASSO algorithm to predict career concerns for our experimental sample based on pre-determined covariates (Belloni and Chernozhukov, 2013). The predicted variable is standardized to have mean zero and standard deviation one. The results of the post-LASSO regression are displayed in Table A3.8. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

3.6 Conclusion

How do political activists' effort choices depend on their beliefs about the effort of fellow activists? This paper presents a natural field experiment to provide evidence on the strategic interdependence of political activists' intentions and behavior, a core question of the collective action problem of political activism. In collaboration with a major political party in a Western European country, we exogenously manipulate party supporters' beliefs about their peers' canvassing effort in a large door-to-door canvassing campaign in the run-up to a nationwide general election. We study how the belief manipulation affects the effort provision in the campaign.

Our findings demonstrate, on average, a pattern of strategic substitutability: political activists lower their effort when learning that fellow party members are more likely to canvass. This result holds for the self-stated willingness to canvass as well as actual canvassing effort measured through a smartphone application. Treatment effects are driven by behavior several weeks after the treatment, underscoring that our intervention had a long-lasting impact beyond the intentions reported right after the treatment. The estimated effect sizes are large (up to 38% of the mean in the control group), suggesting that strategic considerations are quantitatively important in shaping political activists' motivations and behavior.

These findings underscore the empirical relevance of the classical free-rider assumption in models of collective action (Olson, 1965). However, the substantial heterogeneity in responsiveness to our treatment also highlights that this assumption is unlikely to apply uniformly across the population. In particular, while respondents motivated by career concerns are more likely to exhibit strategic substitutability, respondents displaying stronger ties to the party are less likely to exhibit strategic substitutability. These findings highlight that contributions to collective action can be driven by multiple motives, in particular social considerations which can be a force for strategic complementarity (Ostrom, 2000; Falk and Fischbacher, 2006; Uhlaner, 1989). Models of collective action may incorporate such heterogeneity in their assumptions for instance by incorporating different types of actors, or considering multiple motives and their interaction simultaneously as in a recently proposed model by Jia and Persson (2017).

Besides the theoretical implications of our results, the evidence also provides practical implications for policy makers and campaign organizers. Anecdotal evidence suggests that party

officials and campaign organizers rely on the assumption of strategic complementarity. For instance, our cooperating party tried to motivate its supporters by highlighting high levels of canvassing participation through e-mail notifications like "*Dear supporter, our diligent canvassers have already knocked on over [number anonymized] doors. And every day the number increases! [...] Everybody goes from door to door! Participate as well!*" Our results suggest that such promotion strategies might actually be counterproductive on average. Instead, our findings indicate that instrumental motives, such as career concerns within the party, can be an important driver of party supporters' participation decisions. Hence, our results might also be informative for party officials in designing effective incentive schemes for the party's supporters.

Acknowledgements For helpful comments, I thank Johannes Abeler, Ned Augenblick, Eric Avis, Lukas Balafoutas, Ernesto Dal Bo, Davide Cantoni, Stefano Caria, Stefano DellaVigna, Dominik Duell, Christine Exley, Claudio Ferraz, Fred Finan, Simon Gächter, Nicola Gennaioli, Thomas Graeber, Don Green, Michael Kosfeld, Alexander Haas, Macartan Humphreys, Ulrike Malmendier, Suresh Naidu, Muriel Niederle, Ricardo Perez-Truglia, Thorsten Persson, Gautam Rao, Leah Shiferaw, David Yang, Noam Yuchtman, and Florian Zimmermann and seminar audiences at the University of California Berkeley, WZB Berlin, the Institute on Behavior and Inequality in Bonn, University of Cologne, EUI Florence, University of Konstanz, King's College London, University of Innsbruck, University of Oxford, University of Tilburg, University of Zurich, the Demand for Democracy conference in Munich, and the 2018 ECBE conference in Bergen. The pre-analysis plan for this study is available at the AEA RCT Registry following this link: `https://www.socialscienceregistry.org/trials/2358`. The research described in this article was approved by the University of Oxford Department of Economics Research Ethics Committee.

Appendices of Chapter 3

Appendix Section 3.A provides supplementary proofs and derivations. Appendix Section 3.B provides the invitation email, followed by Appendix Section 3.C providing the survey instrument. Appendix Section 3.D describes deviations from the pre-analysis plan.

Appendix Section 3.E contains additional figures. Figure A3.1 displays the relationship between intentions and actual canvassing behavior. Figure A3.2 show the cumulative distribution of intended days as well as actual days and doors canvassed. Appendix Section 3.F contains additional tables. Table A3.1 shows balance tests. Table A3.2 displays treatment effects on application download. Table A3.3 presents the main results using randomization inference. Table A3.4 compares high and low social connection respondents in the post-election survey. Table A3.5 shows heterogeneity by previous canvassing experience. Table A3.6 shows heterogeneity by party membership. Table A3.7 displays heterogeneity by party membership duration. Table A3.8 displays the results of the post-LASSO regressions to predict career concerns. Table A3.9 shows treatment effects on social media campaigning activity. Table A3.10 shows the relationship between beliefs about the canvassing effort of fellow party membes and canvassing outcomes in the control group. Table A3.11 shows the effects by application download prior to the experiment. Tables A3.12 to A3.20 provide results for overestimators. Tables A3.21 to A3.29 provide results for underestimators without pre-specified controls.

3.A Mathematical Appendix

We model agent i's utility of choosing her own canvassing effort d_i given fellow supporters' canvassing effort (d_{-i}) according to:

$$u_i(d_i) = (1-\alpha_i)g(d_i, d_{-i}) + \alpha_i h(d_i, d_{-i}) - c_i(d_i) - \tilde{c}_i \cdot \mathbb{1}(d_i > 0) \tag{3.3}$$

where $g(d_i, d_{-i})$ represents the instrumental utility gained from either the overall level of canvassing activity (i.e. additional votes) or the utility gained from signaling one's commitment. We assume $\frac{\partial g(d_i,d_{-i})}{\partial d_i} > 0$ and $\frac{\partial^2 g(d_i,d_{-i})}{\partial d_i \partial d_{-i}} < 0$. The term $h(d_i, d_{-i})$ represents social motives. We assume $\frac{\partial h(d_i,d_{-i})}{\partial d_i} \geq 0$ and $\frac{\partial^2 h(d_i,d_{-i})}{\partial d_i \partial d_{-i}} > 0$. The term α_i denotes an individual-specific relative weight on social motives. $c_i(d_i)$ represents individual-specific costs of canvassing for which we assume $c'(0) = 0$, and $c', c'' > 0$. For $g(\cdot)$, $h(\cdot)$, and $c(\cdot)$ we assume continuity and that they are twice differentiable. $\tilde{c}_i$ is an individual-specific fixed costs of doing any canvassing distributed according to F. Agent i chooses d_i to maximize her utility u_i, where the reservation utility from not engaging in canvassing is equal to $(1-\alpha_i)g(0, d_{-i})$. We assume that all idiosyncratic properties are summarized by type $i \in I$, distributed according to $\Phi \perp F$.

Formally, agents choose $d_i \geq 0$ such that

$$d_i^* = \text{argmax}_{d_i}(1-\alpha_i)g(d_i, d_{-i}) + \alpha_i h(d_i, d_{-i}) - c_i(d_i) - \tilde{c}_i \cdot \mathbb{1}(d_i > 0) \tag{3.4}$$

if $u_i(d_i^*) \geq (1-\alpha_i)g(0, d_{-i})$ and $d_i^* = 0$ otherwise.

First, we investigate intensive margin responses and focus on interior solutions of this optimization problem. In this case, agent i sets her canvassing effort d_i according to the following first order condition:

$$\frac{\partial u(d_i)}{\partial d_i} = (1-\alpha_i)\frac{\partial g(d_i, d_{-i})}{\partial d_i} + \alpha_i \frac{\partial h(d_i, d_{-i})}{\partial d_i} - \frac{\partial c_i(d_i)}{\partial d_i} = 0 \tag{3.5}$$

Without further functional form assumptions there is no closed form solution for the optimal effort choice d_i^*. However, it is possible to analyze i's optimal response to changes in d_{-i} using implicit differentiation:

$$\frac{\partial d_i^*}{\partial d_{-i}} = -\frac{(1-\alpha_i)\frac{\partial^2 g(d_i,d_{-i})}{\partial d_i \partial d_{-i}} + \alpha_i \frac{\partial^2 h(d_i,d_{-i})}{\partial d_i \partial d_{-i}}}{(1-\alpha_i)\frac{\partial^2 g(d_i,d_{-i})}{\partial d_i \partial d_i} + \alpha_i \frac{\partial^2 h(d_i,d_{-i})}{\partial d_i \partial d_i} - \frac{\partial^2 c(d_i)}{\partial d_i \partial d_i}} \tag{3.6}$$

Note that the denominator is negative, as we require d^* to be at a local maximum. This is guaranteed, for instance, if we assume $\frac{\partial^2 g(d_i,d_{-i})}{\partial d_i \partial d_i} < 0$ and $\frac{\partial^2 h(d_i,d_{-i})}{\partial d_i \partial d_i} \leq 0$. This immediately

yields the intensive margin result presented in the main text.

Next, we investigate how $\frac{\partial d_i^*}{\partial d_{-i}}$ varies with the strength of social motives. Consider the case where $\alpha_i = 0$, that is the weight on social motives is zero. Then,

$$\frac{\partial d_i^*}{\partial d_{-i}}\Big|_{\alpha_i=0} = -\frac{\frac{\partial^2 g(d_i,d_{-i})}{\partial d_i \partial d_{-i}}}{\frac{\partial^2 g(d_i,d_{-i})}{\partial d_i \partial d_i} - \frac{\partial^2 c(d_i)}{\partial d_i \partial d_i}} < 0. \tag{3.7}$$

In the case where $\alpha_i = 1$,

$$\frac{\partial d_i^*}{\partial d_{-i}}\Big|_{\alpha_i=1} = -\frac{\frac{\partial^2 h(d_i,d_{-i})}{\partial d_i \partial d_{-i}}}{\frac{\partial^2 h(d_i,d_{-i})}{\partial d_i \partial d_i} - \frac{\partial^2 c(d_i)}{\partial d_i \partial d_i}} > 0. \tag{3.8}$$

As $\frac{\partial d_i^*}{\partial d_{-i}}$ is continuous in α_i, we can draw two conclusions. There exists a cutoff-level $\underline{\alpha_i}$ such that individuals with $\alpha_i < \underline{\alpha_i}$ exhibit strategic substitutability. Vice versa, there exists a cutoff-level $\bar{\alpha_i}$ such that individuals with $\alpha_i > \bar{\alpha_i}$ exhibit strategic complementarity.

Next, we consider extensive margin responses. Note that for each individual of type i there exists a cutoff value $\tilde{c_i}^* = (1-\alpha_i)g(d_i^*, d_{-i}) + \alpha_i h(d_i^*, d_{-i}) - c_i(d_i^*) - (1-\alpha_i)g(0, d_{-i})$ such that the individual chooses to engage in canvassing iff $\tilde{c_i} \leq \tilde{c_i}^*$. Hence, the fraction of individuals of type i deciding to engage in canvassing is equal to

$$F\Big((1-\alpha_i)g(d_i^*, d_{-i}) + \alpha_i h(d_i^*, d_{-i}) - c_i(d_i^*) - (1-\alpha_i)g(0, d_{-i})\Big). \tag{3.9}$$

Taking the derivative with respect to d_{-i} yields

$$\Big((1-\alpha_i)\Big(\frac{\partial g(d_i^*, d_{-i})}{\partial d_{-i}} - \frac{\partial g(0, d_{-i})}{\partial d_{-i}}\Big) + \alpha_i \frac{\partial h(d_i^*, d_{-i})}{\partial d_{-i}}\Big) \cdot f(\tilde{c_i}^*). \tag{3.10}$$

Given g is concave in d_i, $\frac{\partial g(d_i^*, d_{-i})}{\partial d_{-i}} - \frac{\partial g(0, d_{-i})}{\partial d_{-i}}$ will be unambiguously negative. Hence, for low α_i the above term will be negative. If we assume $\frac{h(d_i,d_{-i})}{\partial d_{-i}} \geq 0$, i.e. social utility increases in the participation of others, reflecting the quality of canvassing as a relational good, the sign of the effect is ambiguous and depends on the relative strength of social motives α_i. Most importantly, however, extensive margin responses depend critically on the mass of individuals $f(\tilde{c_i}^*)$ who are marginal. Hence, we expect the relative strength of intensive and extensive margin responses to depend on the degree of how many individuals are marginal in their decision to do any canvassing. Aggregating over all individuals yields

$$\int_{i\in I}\Big((1-\alpha_i)\Big(\frac{\partial g(d_i^*,d_{-i})}{\partial d_{-i}}-\frac{\partial g(0,d_{-i})}{\partial d_{-i}}\Big)+\alpha_i\frac{\partial h(d_i^*,d_{-i})}{\partial d_{-i}}\Big)\cdot f(\tilde{c_i}^*)d\Phi. \tag{3.11}$$

3.B Invitation email

Dear 'name',

the critical stage of the election campaign is imminent. We have conducted workshops in almost all constituencies and the feedback was resoundingly positive. Now it's up to us. We are all out to canvass to help [party name] win the election.

To help our campaign succeed, we ask you to respond to a short survey. We would like to know if you have any suggestions and to what extent you plan to participate in the campaign. Your answers will of course be treated confidentially.

Here is the link to the survey:

Survey

It would be great if you could support us with this survey. Just click on the link right now. The survey only takes five minutes.

3.C Survey instrument

- **Introduction**

  Dear 'name',

  we are conducting a short survey among our supporters to plan our election campaign. Your participation helps us to use our campaign resources optimally. We will treat your answers confidentially. The survey only takes 5 minutes (10 questions).

  Thank you very much for your help!

- **Sex**

  What is your sex?

- **Age**

  How old are you?

- **Party member**

  Are you a member of [party name] party?

- **Years of party membership** (asked if respondent is party member)

  For how many years have you been a member of [party name] party?

- **Canvassing workshop**

  Have you ever participated in a canvassing training workshop?

- **Canvassing experience**

  Do you have experiences from canvassing in previous election campaigns?

- **Pre-treatment belief**

  Think of 100 typical [party name] party members.

  What do you think: How many of these 100 [party name] party members plan to engage in canvassing during this election campaign?

- **Treatment text**

  You said X of 100 [party name] party members.

  According to a survey of [party name] party members, 37 of 100 [party name] party members plan to engage in canvassing during this election campaign.

- **Post-treatment belief**

  What do you think: How many of these 100 [party name] party members will actually engage in canvassing during this election campaign?

- **Extensive margin**

  Do you plan to canvass during this election campaign?

- **Intensive margin** (asked if extensive margin is yes)

  On how many days do you plan to canvass during this election campaign?

- **Debrief**

  Now let's go! And don't forget to download the [party name]-application. Here for iOS and Android.

  With the [party name]-application you can actively participate in our election campaign and keep up to date with the campaign progress. Also, the application is fun!

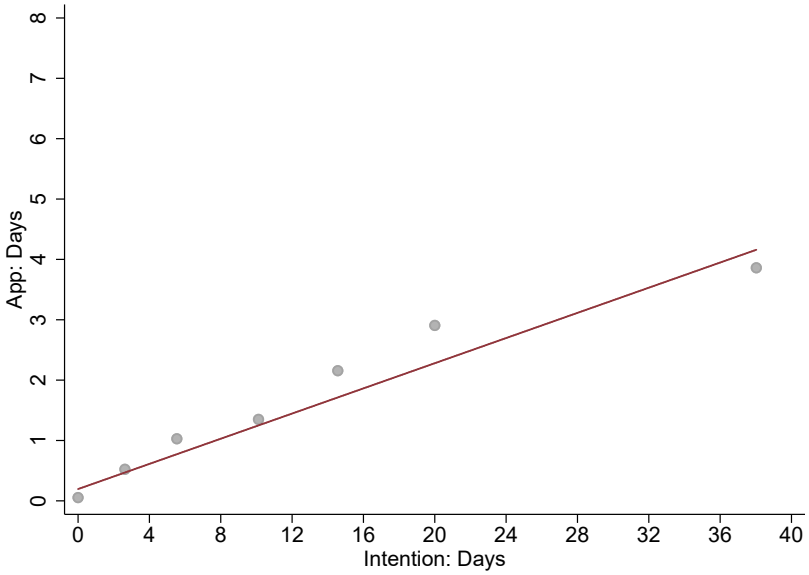
3.D Deviations from pre-analysis plan

This section describes where the analysis presented in the paper deviates from the pre-analysis plan uploaded to the AEA RCT registry (AEARCTR-0002358). The pre-analysis plan contains pre-registered analyses for two separate experiments ("own party" and "other party" group). We present the results of these experiments in two different papers. This paper describes the results of the "own party" experiment. The results of the "other party" group are presented in Hager et al. (2019a) which is available upon request from the authors.

For the analysis of the experiment presented in this paper, we deviated in the following points from the pre-analysis plan:

- We conduct additional analysis not covered in the pre-analysis plan. This includes the following analysis:
  - Any analysis of temporal patterns of treatment effects, including the impacts on doors and days in week 7/8.
  - Heterogeneity by strength of connection to the party (PCA of three pre-specified control variables) and predicted career concerns.
  - The impact of the treatment on social media campaigning.
  - An IV-analysis instrumenting post-treatment beliefs with treatment assignment.
- We also include randomization inference as an additional robustness check.

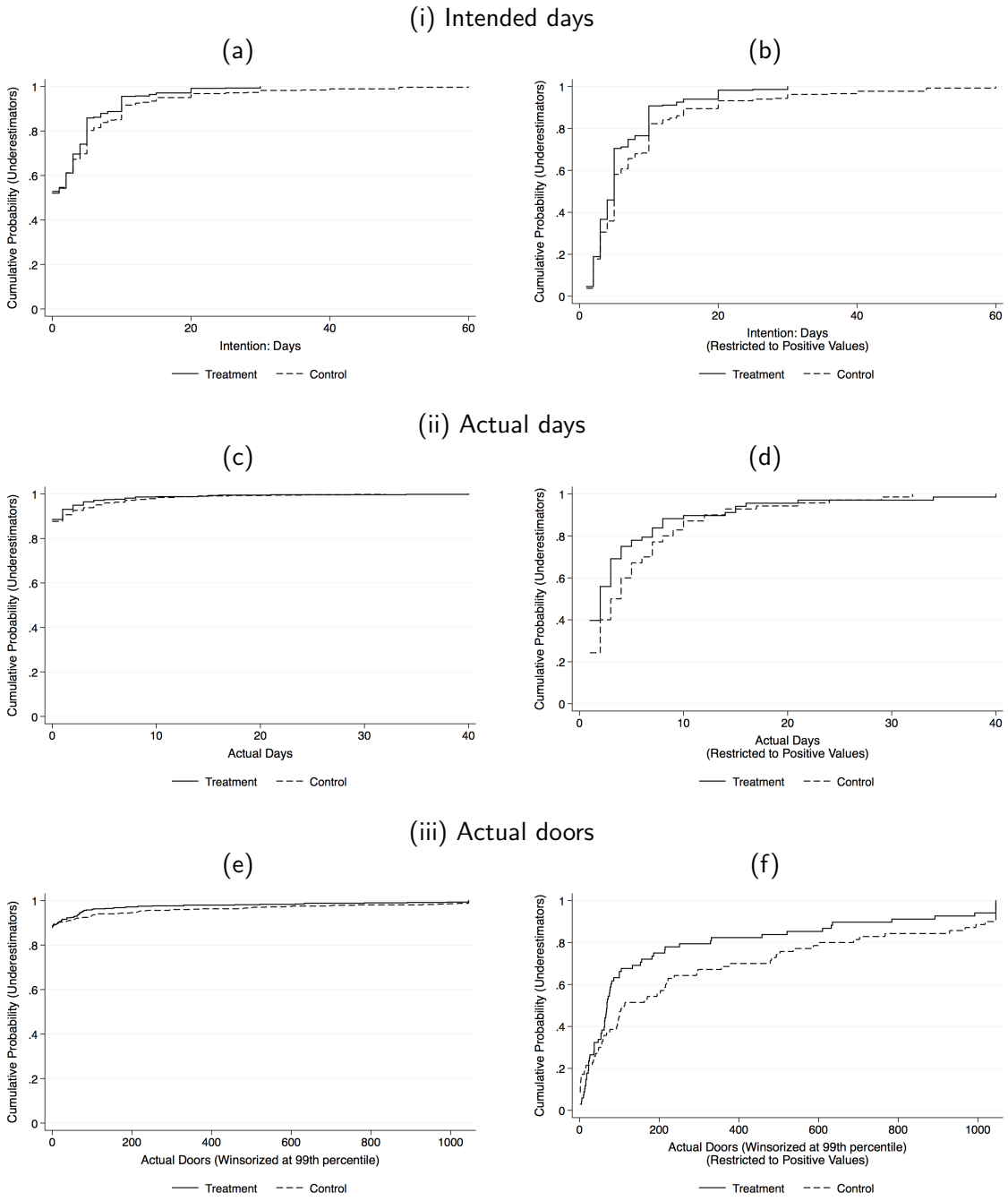
Figure A3.1: Intentions vs. behavior (data from survey and online application)



Notes: Figure A3.1 shows a binscatter plot between intended number of days and number of days canvassing as recorded through the online application. The regression line indicates the best linear fit.

3.E Additional Figures

Figure A3.2: Cumulative distribution function of key outcomes



Notes: Figures A3.2 a-f show cumulative distribution functions (cdfs) of three key outcomes: (i) intended days canvassing, (ii) actual days canvassing, (iii) actual doors knocked on (winsorized at 99th percentile). The cdfs are plotted for respondents who underestimate the share of fellow party members who plan to participate in the party's door-to-door canvassing campaign. Post-treatment beliefs about actual participation are instrumented with the treatment indicator. The panels (a), (c), and (e) show the distribution for all observations, panels (b), (d), and (f) show the distribution for positive observations only.

3.F Additional Tables

Table A3.1: Balance tests

	Treatment	Control	Δ	se(Δ)	p(Δ=0)
Female	0.229	0.245	-0.016	(0.023)	0.484
Age	40.274	41.805	-1.532	(1.027)	0.136
Is party member	0.821	0.832	-0.012	(0.020)	0.568
Years of party membership	11.731	12.654	-0.922	(0.757)	0.223
Has experience canvassing	0.384	0.371	0.013	(0.026)	0.617
Participated in door-to-door workshop	0.201	0.220	-0.020	(0.022)	0.359
Downloaded app before survey	0.287	0.256	0.031	(0.024)	0.195
Has canvassed before survey	0.081	0.078	0.002	(0.014)	0.875
Days canvassed before survey	0.185	0.203	-0.018	(0.051)	0.719
Doors visited before survey	3.192	5.576	-2.384	(2.229)	0.285
Prior Belief: % of party members who canvass	19.860	20.727	-0.867	(1.123)	0.440
Number of observations	708	703			

Notes: Table A3.1 presents balance tests for the treatment and control group in our experiment. Columns 1 and 2 report variable means. Column 3 reports difference in means. Column 4 reports the associated heteroskedasticity robust standard error. Column 5 the p-value of a test of equality of means. We regress the treatment indicator on all covariates to test for joint significance. The p-value of this joint F-test is 0.59.

Table A3.2: Treatment effects on application download

	Survey	App Data		
	click on app-link	within 24h	within one week	any time after survey
Treatment	-0.006	0.005	-0.005	0.010
	(0.012)	(0.008)	(0.009)	(0.014)
Observations	1163	1163	1163	1163
Control group mean	0.04	0.02	0.03	0.05

Notes: Table A3.2 presents treatment effects on application download for respondents who underestimate the share of fellow party members who plan to participate in the party's door-to-door canvassing campaign. Pre-specified controls are as defined in Table 3.2.

Table A3.3: Main effects - randomization inference

	Posterior	Intentions		App Data			Index	App: Week 7/8	
	Belief	Any	Days	Any	Days	Door	Overall	Days	Doors
Treatment	5.027	0.002	-1.098	-0.013	-0.159	-14.388	-0.093	-0.145	-6.657
	[0.0000]	[0.9299]	[0.0025]	[0.4168]	[0.3145]	[0.0648]	[0.0477]	[0.0219]	[0.0128]
Control mean	9.083	0.473	4.028	0.123	0.701	38.348	-0.029	0.297	13.318
Observations	1150	1148	1148	1163	1163	1163	1148	1163	1163

Notes: Table A3.3 presents treatment effects for respondents who underestimate the share of fellow party members who plan to participate in the party's door-to-door canvassing campaign. p-values based on randomization inference with 10,000 draws are displayed in brackets. Outcomes and pre-specified controls are as defined in Table 3.2.

Table A3.4: Comparison of supporters with weak and strong connection to the party

	Weak connections to party	Strong connections to party	P-value(Weak con. = Strong con.)	N
Number of known party member	26.99	49.50	0.000	902
Identification with party (z-scored)	-0.13	0.14	0.000	884
Identification with party platform (z-scored)	-0.11	0.12	0.000	877

Notes: Table A3.4 presents summary statistics for supporters with below and above median strength of connections to the party. Connection to the party is measured as first principal component of the three pre-specified dimensions of heterogeneity (canvassing experience prior to the federal election, party membership dummy, and party membership duration). Columns 1 and 2 report variable means. Column 3 reports the p-value of a test of equality of means. Column 4 reports the number of observations. The data was collected about two months after the election through a sample of the same population of party supporters but excluding participants in the main experiment. Row 1 reports the self-reported number of party members individuals know personally. Row 2 reports the means of z-scored responses to the question "How close are you to the party?" on a seven point Likert-scale. Row 3 reports the means of z-scored responses to the question "How much do you support the party platform?" on a seven point Likert-scale.

Table A3.5: Treatment effects by canvassing experience

	Posterior	Intentions		App Data			Index	App: Week 7/8	
	Belief	Any	Days	Any	Days	Door	Overall	Days	Doors
Panel A: Inexperienced supporters									
Treatment	4.566***	-0.014	-1.356***	-0.010	-0.295*	-25.112***	-0.142**	-0.205***	-11.075***
	(0.738)	(0.032)	(0.447)	(0.018)	(0.171)	(9.638)	(0.056)	(0.077)	(3.474)
Control mean	9.872	0.376	3.382	0.101	0.657	39.110	-0.126	0.300	14.187
Observations	701	700	700	710	710	710	700	710	710
Panel B: Experienced supporters									
Treatment	5.675***	0.025	-0.722	-0.015	0.088	3.527	-0.012	-0.045	0.780
	(0.785)	(0.042)	(0.608)	(0.028)	(0.323)	(13.850)	(0.086)	(0.114)	(4.884)
Control mean	7.850	0.626	5.041	0.158	0.770	37.158	0.122	0.293	11.959
Observations	449	448	448	453	453	453	448	453	453

Notes: Table A3.5 presents treatment effects for respondents who underestimate the share of fellow party members who plan to participate in the party's door-to-door canvassing campaign. Panel A contains the sample of supporters without previous canvassing experience, Panel B contains the sample of supporters with previous canvassing experience. Outcomes and pre-specified controls are as defined in Table 3.2.

Table A3.6: Treatment effects by party membership

	Posterior	Intentions		App Data			Index	App: Week 7/8	
	Belief	Any	Days	Any	Days	Door	Overall	Days	Doors
Panel A: No party member									
Treatment	2.590	0.002	-2.420**	-0.029	-0.506	-36.275*	-0.227*	-0.196	-10.681
	(2.096)	(0.064)	(1.037)	(0.026)	(0.343)	(20.542)	(0.116)	(0.156)	(7.353)
Control mean	13.012	0.367	4.266	0.061	0.573	41.854	-0.129	0.183	9.976
Observations	168	167	167	172	172	172	167	172	172
Panel B: Party member									
Treatment	5.461***	0.004	-0.857**	-0.009	-0.104	-10.520	-0.069	-0.142*	-6.193**
	(0.535)	(0.028)	(0.386)	(0.018)	(0.180)	(8.608)	(0.053)	(0.073)	(3.135)
Control mean	10.623	0.475	4.311	0.113	0.621	47.782	-0.016	0.258	12.597
Observations	982	981	981	991	991	991	981	991	991

Notes: Table A3.6 presents treatment effects for respondents who underestimate the share of fellow party members who plan to participate in the party's door-to-door canvassing campaign. Panel A contains the sample of non-party members, Panel B contains the sample of party members. Outcomes and pre-specified controls are as defined in Table 3.2.

Table A3.7: Treatment effects by party membership duration

	Posterior	Intentions		App Data			Index	App: Week 7/8	
	Belief	Any	Days	Any	Days	Door	Overall	Days	Doors
Panel A: Below med. membership dur.									
Treatment	4.612***	-0.033	-1.751***	-0.007	-0.215	-25.117*	-0.156**	-0.218**	-10.721**
	(0.797)	(0.036)	(0.542)	(0.023)	(0.229)	(12.932)	(0.071)	(0.089)	(4.364)
Control mean	9.900	0.523	4.674	0.141	0.827	52.676	0.076	0.363	17.359
Observations	575	574	574	582	582	582	574	582	582
Panel B: Above med. membership dur.									
Treatment	5.465***	0.044	-0.419	-0.020	-0.135	-5.952	-0.034	-0.087	-3.219
	(0.723)	(0.036)	(0.475)	(0.021)	(0.223)	(9.087)	(0.063)	(0.096)	(3.739)
Control mean	8.278	0.424	3.392	0.105	0.575	24.070	-0.133	0.232	9.291
Observations	575	574	574	581	581	581	574	581	581

Notes: Table A3.7 presents treatment effects for respondents who underestimate the share of fellow party members who plan to participate in the party's door-to-door canvassing campaign. Panel A contains the sample of supporters who have a below median party membership duration (including non-members), Panel B contains the sample of supporters who have above median membership duration. Outcomes and pre-specified controls are as defined in Table 3.2.

Table A3.8: Predicting career concerns

	(1) Career concern
Age	-0.0116***
	(0.00200)
Experienced supporter	-0.276***
	(0.0647)
Canvassed before experiment	0.0699
	(0.0750)
Canvassing workshop	-0.0910
	(0.0979)
Years of party membership	-0.0192***
	(0.00489)
Years party membership squared	0.000233***
	(0.0000870)
Age × party member	0.00504**
	(0.00207)
Age × canvassing workshop	-0.00109
	(0.00273)
Female × experienced supporter	0.104
	(0.0829)
Female × canvassed before experiment	-0.171
	(0.106)
Female × party member	-0.0858
	(0.146)
Female × canvassing workshop	0.160**
	(0.0810)
Female × years party membership	-0.00177
	(0.00348)
Experienced supporter × canvassing workshop	0.186**
	(0.0774)
Experienced supporter × years party membership	0.00622**
	(0.00245)
Experienced supporter × canvassed before experiment	0.186**
	(0.0946)
Canvassed before experiment × years party membership	-0.0166***
	(0.00321)
Canvassing workshop × years party membership	-0.00220
	(0.00406)
Female	-0.0630
	(0.133)
Constant	1.102***
	(0.0511)
R-squared	0.270
Number of observations	744

Notes: Table A3.8 presents the results of a OLS regression of variables selection by a LASSO algorithm on career concerns. Career concerns are defined as choosing "It is possible" as an answer to at least one the following questions "Can you imagine yourself taking on a political position within the party?" and "Can you imagine yourself running for office for this party?" (the answer options were "No", "It is possible", and "I am already doing this"). The sample is obtained using a post-election survey administered two months after the election to the same population of party supporters (excluding individuals who participated in the main experiment). The post-LASSO algorithm has two steps. First, we selected variables using a LASSO algorithm with the penalty parameter λ chosen optimally through 10-fold cross-validation. The set of predetermined variables includes the following variables and all pairwise interactions: age, age squared, party membership, years of party membership, years of party membership squared, sex, a dummy for whether a participant has participated in a canvassing workshop, a dummy indicating any prior canvassing experience before this federal election, and dummies for whether the supporter has already canvassed in this election before the survey. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A3.9: Effects on effort in alternative domain: social media activity

	Shared social media message		
	Any	Days	Total
Treatment	-0.011	-0.052	-0.148
	(0.014)	(0.120)	(0.220)
Control mean	0.074	0.388	0.782
Observations	1163	1163	1163

Notes: Table A3.9 presents treatment effects on social media activity for respondents who underestimate the share of fellow party members who plan to participate in the party's door-to-door canvassing campaign. *"Any"* takes value one if the respondent shares any party news story on Facebook through the application. *"Days"* denotes the total number of days a respondent shares a party news story on Facebook through the application. *"Total"* is the total number of party news stories shared by the respondent on Facebook through the application. Pre-specified controls are as defined in Table 3.2. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A3.10: OLS estimates: beliefs about peers and effort (control group)

	Intentions		App Data			Index	App: Week 7/8	
	Any	Days	Any	Days	Door	Overall	Days	Doors
Posterior	0.002	0.044	0.001	0.016	1.209*	0.007*	0.007	0.464*
	(0.002)	(0.031)	(0.001)	(0.011)	(0.695)	(0.004)	(0.005)	(0.247)
Observations	562	562	569	569	569	562	569	569

Notes: Table A3.10 presents OLS estimates for the impact of beliefs about fellow supporters' canvassing effort for control group respondents who underestimate the share of fellow party members who plan to participate in the party's door-to-door canvassing campaign. Outcomes and pre-specified controls are as defined in Table 3.2. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A3.11: Treatment effects by application download

	Posterior	Intentions		App Data			Index	App: Week 7/8	
	Belief	Any	Days	Any	Days	Door	Overall	Days	Doors
Panel A: No app download									
Treatment	4.341***	-0.001	-0.813**	-0.003	-0.144	-8.556*	-0.062*	-0.091*	-3.218*
	(0.664)	(0.031)	(0.404)	(0.011)	(0.093)	(4.536)	(0.038)	(0.049)	(1.696)
Control mean	9.438	0.366	2.908	0.026	0.206	10.848	-0.304	0.114	4.241
Observations	844	842	842	855	855	855	842	855	855
Panel B: App download									
Treatment	6.816***	0.000	-1.985**	-0.037	-0.119	-26.455	-0.168	-0.259	-13.605
	(0.839)	(0.045)	(0.786)	(0.051)	(0.531)	(26.353)	(0.142)	(0.193)	(9.169)
Control mean	8.000	0.799	7.439	0.418	2.206	121.823	0.808	0.851	40.872
Observations	306	306	306	308	308	308	306	308	308

Notes: Table A3.11 presents treatment effects for respondents who underestimate the share of fellow party members who plan to participate in the party's door-to-door canvassing campaign. Panel A contains the sample of supporters who had not downloaded the application before the treatment, Panel B contains the sample of supporters who had downloaded the application before the treatment. Outcomes and pre-specified controls are as defined in Table 3.2.

Results for overestimators

Table A3.12: Main effects (overestimators)

	Posterior	Intentions		App Data			Index	App: Week 7/8	
	Belief	Any	Days	Any	Days	Door	Overall	Days	Doors
Treatment	-16.686***	-0.081	-0.926	-0.029	-0.097	-9.009	-0.125	0.003	-2.561
	(1.973)	(0.053)	(1.227)	(0.038)	(0.186)	(12.265)	(0.086)	(0.055)	(4.030)
Control mean	9.083	0.473	4.028	0.123	0.701	38.348	-0.029	0.297	13.318
Observations	247	247	247	248	248	248	247	248	248

Notes: Table A3.12 presents treatment effects for respondents who overestimate the share of fellow party members who plan to participate in the party's door-to-door canvassing campaign. Outcomes and pre-specified controls are as defined in Table 3.2.

Table A3.13: Treatment effects and strength of connection to the party (overestimators)

	Posterior	Intentions		App Data			Index	App: Week 7/8	
	Belief	Any	Days	Any	Days	Door	Overall	Days	Doors
Treatment	-14.664***	-0.082	-0.197	-0.019	-0.005	-3.978	-0.073	0.018	-1.254
	(1.868)	(0.056)	(1.511)	(0.039)	(0.242)	(13.765)	(0.099)	(0.069)	(4.431)
Treatment × Strength of connection to party (PCA)	6.173***	-0.002	2.226	0.029	0.282	15.503	0.158	0.046	4.029
	(2.099)	(0.054)	(1.519)	(0.035)	(0.248)	(13.686)	(0.099)	(0.069)	(3.704)
Strength of connection to party (PCA)	-3.895	0.326***	1.460	0.047	0.486*	22.224	0.356***	0.166*	4.632
	(2.816)	(0.076)	(1.552)	(0.050)	(0.258)	(17.751)	(0.112)	(0.086)	(5.305)
Control mean	45.313	0.545	5.754	0.149	0.440	22.515	0.034	0.119	7.157
Observations	247	247	247	248	248	248	247	248	248

Notes: Table A3.13 presents treatment effects interacted with a summary measure of connectedness to the party for respondents who overestimate the share of fellow party members who plan to participate in the party's door-to-door canvassing campaign. The measure for strength of connection to the party is the first principal component of three pre-specified dimensions of heterogeneity (previous canvassing experience, party membership and party membership duration) obtained from a principal component analysis. This summary measure is standardized to have mean zero and standard deviation one. Outcomes and pre-specified controls are as defined in Table 3.2. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A3.14: Treatment effects and career concerns (overestimators)

	Posterior	Intentions		App Data			Index	App: Week 7/8	
	Belief	Any	Days	Any	Days	Door	Overall	Days	Doors
Treatment	-16.324***	-0.082	-0.973	-0.030	-0.112	-9.033	-0.130	-0.006	-2.707
	(1.998)	(0.054)	(1.237)	(0.037)	(0.174)	(11.829)	(0.086)	(0.052)	(3.854)
Treatment × Predicted career concerns	-4.774**	0.021	-1.095	0.007	-0.077	-6.900	-0.043	0.020	1.171
	(1.959)	(0.053)	(1.652)	(0.039)	(0.190)	(9.404)	(0.101)	(0.057)	(2.440)
Predicted career concerns	3.481	-0.046	3.829**	0.015	0.590	19.131	0.217	0.158	0.486
	(3.124)	(0.074)	(1.831)	(0.062)	(0.459)	(25.621)	(0.151)	(0.122)	(8.348)
Control mean	45.313	0.545	5.754	0.149	0.440	22.515	0.034	0.119	7.157
Observations	247	247	247	248	248	248	247	248	248

Notes: Table A3.14 presents treatment effects interacted with a supporter's career concerns for respondents who overestimate the share of fellow party members who plan to participate in the party's door-to-door canvassing campaign. We do not directly observe career concerns for participants in our experiment. We therefore predict career concerns using a post-election survey administered two months after the election to the same population of party supporters (excluding individuals who participated in the main experiment). Career concerns are measured as choosing "It is possible" as an answer to at least one the following questions "Can you imagine yourself taking on a political position within the party?" and "Can you imagine yourself running for office for this party?" (the answer options were "No", "It is possible", and "I am already doing this"). We use a post-LASSO algorithm to predict career concerns for our experimental sample based on pre-determined covariates (Belloni and Chernozhukov, 2013). The predicted variable is standardized to have mean zero and standard deviation one. The results of the post-LASSO regression are displayed in Table A3.8. Outcomes and pre-specified controls are as defined in Table 3.2. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A3.15: IV estimates: post-treatment beliefs and effort choices (overestimators)

	Intentions		App Data			Index	App: Week 7/8	
	Any	Days	Any	Days	Door	Overall	Days	Doors
Panel A: Main Effect (IV)								
Posterior	0.005	0.055	0.002	0.006	0.541	0.007	-0.000	0.155
	(0.003)	(0.072)	(0.002)	(0.011)	(0.730)	(0.005)	(0.003)	(0.241)
Observations	247	247	248	248	248	247	248	248
Panel B: Weak connection to party (IV)								
Posterior	0.005	0.089	0.002	0.011	0.767	0.010**	0.002	0.192
	(0.003)	(0.062)	(0.002)	(0.008)	(0.698)	(0.005)	(0.002)	(0.180)
Observations	151	151	151	151	151	151	151	151
Panel C: Strong connection to party (IV)								
Posterior	0.005	-0.126	-0.005	-0.086	-4.228	-0.023	-0.020	-1.066
	(0.009)	(0.261)	(0.006)	(0.071)	(3.624)	(0.020)	(0.020)	(1.116)
Observations	96	96	97	97	97	96	97	97
Panel D: Low career concerns (IV)								
Posterior	0.005	0.002	0.001	-0.006	-0.751	0.002	0.001	0.019
	(0.006)	(0.153)	(0.003)	(0.015)	(0.816)	(0.009)	(0.005)	(0.277)
Observations	151	151	151	151	151	151	151	151
Panel E: High career concerns (IV)								
Posterior	0.003	0.061	0.001	0.007	0.838	0.007	-0.002	0.027
	(0.003)	(0.069)	(0.003)	(0.015)	(0.916)	(0.006)	(0.004)	(0.227)
Observations	96	96	97	97	97	96	97	97

Notes: Table A3.15 presents IV estimates for the impact of post-treatment beliefs for all overestimators (Panel A). Post-treatment beliefs are instrumented with the treatment indicator. Parallel to Table 3.5, Panels (B) and (C) distinguish between supporters with weak and strong connections to the party, whereas panels (D) and (E) distinguish between supporters with low and high career concerns. Outcomes and pre-specified controls are as defined in Table 3.2.

Table A3.16: Treatment effects on application download (overestimators)

	Survey	App Data		
	click on app-link	within 24h	within one week	any time after survey
Treatment	-0.003	0.014	0.009	-0.021
	(0.029)	(0.025)	(0.026)	(0.033)
Observations	248	248	248	248
Control group mean	0.06	0.03	0.04	0.08

Notes: Table A3.16 presents treatment effects on application download for respondents who overestimate the share of fellow party members who plan to participate in the party's door-to-door canvassing campaign. Pre-specified controls are as defined in Table 3.2.

Table A3.17: Treatment effects by canvassing experience (overestimators)

	Posterior	Intentions		App Data			Index	App: Week 7/8	
	Belief	Any	Days	Any	Days	Door	Overall	Days	Doors
Panel A: Inexperienced supporters									
Treatment	-18.640***	-0.059	-1.435	-0.054	-0.182	-8.344	-0.159*	-0.064	-5.094
	(2.439)	(0.064)	(1.170)	(0.045)	(0.129)	(9.971)	(0.091)	(0.048)	(4.500)
Control mean	46.211	0.442	5.084	0.137	0.326	15.895	-0.079	0.105	6.737
Observations	168	168	168	168	168	168	168	168	168
Panel B: Experienced supporters									
Treatment	-12.179***	-0.101	-0.384	0.034	0.310	-1.601	-0.010	0.168	3.191
	(3.111)	(0.099)	(2.696)	(0.062)	(0.584)	(38.897)	(0.196)	(0.158)	(10.280)
Control mean	43.128	0.795	7.385	0.179	0.718	38.641	0.309	0.154	8.179
Observations	79	79	79	80	80	80	79	80	80

Notes: Table A3.17 presents treatment effects for respondents who overestimate the share of fellow party members who plan to participate in the party's door-to-door canvassing campaign. Panel A contains the sample of respondents without prior canvassing experience, Panel B contains the sample of respondents with prior canvassing experience. Outcomes and pre-specified controls are as defined in Table 3.2.

Table A3.18: Treatment effects by party membership (overestimators)

	Posterior	Intentions		App Data			Index	App: Week 7/8	
	Belief	Any	Days	Any	Days	Door	Overall	Days	Doors
Panel A: No party member									
Treatment	-24.088***	-0.123	-4.222**	-0.022	-0.101	-11.433	-0.266*	0.031	-1.727
	(4.891)	(0.121)	(1.998)	(0.076)	(0.211)	(21.474)	(0.148)	(0.084)	(5.324)
Control mean	49.556	0.528	7.722	0.194	0.500	45.111	0.177	0.111	11.250
Observations	73	73	73	73	73	73	73	73	73
Panel B: Party member									
Treatment	-13.650***	-0.075	0.880	-0.000	0.101	6.505	0.014	0.002	0.706
	(2.034)	(0.065)	(1.872)	(0.048)	(0.313)	(15.453)	(0.122)	(0.088)	(5.200)
Control mean	43.755	0.551	5.031	0.133	0.418	14.214	-0.019	0.122	5.653
Observations	174	174	174	175	175	175	174	175	175

Notes: Table A3.18 presents treatment effects for respondents who overestimate the share of fellow party members who plan to participate in the party's door-to-door canvassing campaign. Panel A contains the sample of non-party members, Panel B contains the sample of party members. Outcomes and pre-specified controls are as defined in Table 3.2.

Table A3.19: Treatment effects by party membership duration (overestimators)

	Posterior	Intentions		App Data			Index	App: Week 7/8	
	Belief	Any	Days	Any	Days	Door	Overall	Days	Doors
Panel A: Below med. membership dur.									
Treatment	-20.156***	-0.112*	-1.287	-0.017	-0.189	-16.278	-0.167	0.013	-2.027
	(2.538)	(0.067)	(1.332)	(0.051)	(0.228)	(15.984)	(0.103)	(0.057)	(3.915)
Control mean	46.313	0.590	6.108	0.169	0.518	27.675	0.104	0.108	5.711
Observations	160	160	160	161	161	161	160	161	161
Panel B: Above med. membership dur.									
Treatment	-10.188***	-0.027	0.934	-0.034	-0.001	2.230	-0.003	-0.023	-1.467
	(3.044)	(0.099)	(2.787)	(0.055)	(0.240)	(9.935)	(0.160)	(0.091)	(5.715)
Control mean	43.686	0.471	5.176	0.118	0.314	14.118	-0.079	0.137	9.510
Observations	87	87	87	87	87	87	87	87	87

Notes: Table A3.19 presents treatment effects for respondents who overestimate the share of fellow party members who plan to participate in the party's door-to-door canvassing campaign. Panel A contains the sample of supporters that have a below median party membership duration (including non-members), Panel B contains the sample of supporters who have above median membership duration. Outcomes and pre-specified controls are as defined in Table 3.2.

Table A3.20: Treatment effects by application download (overestimators)

	Posterior	Intentions		App Data			Index	App: Week 7/8	
	Belief	Any	Days	Any	Days	Door	Overall	Days	Doors
Panel A: No app download									
Treatment	-16.408***	-0.111	-1.019	-0.006	-0.022	-1.429	-0.106	-0.005	0.198
	(2.350)	(0.069)	(1.295)	(0.032)	(0.067)	(2.086)	(0.082)	(0.022)	(0.312)
Control mean	45.432	0.421	3.737	0.032	0.074	2.105	-0.270	0.021	0.095
Observations	172	172	172	173	173	173	172	173	173
Panel B: App download									
Treatment	-15.388***	0.011	0.438	-0.032	-0.250	-22.830	-0.069	0.025	-10.017
	(3.552)	(0.088)	(3.329)	(0.112)	(0.604)	(39.036)	(0.247)	(0.184)	(16.135)
Control mean	45.026	0.846	10.667	0.436	1.333	72.231	0.775	0.359	24.359
Observations	75	75	75	75	75	75	75	75	75

Notes: Table A3.20 presents treatment effects for respondents who overestimate the share of fellow party members who plan to participate in the party's door-to-door canvassing campaign. Panel A contains the sample of supporters who had not downloaded the application before the treatment, Panel B contains the sample of supporters who had downloaded the application before the treatment. Outcomes and pre-specified controls are as defined in Table 3.2.

Results without pre-specified controls

Table A3.21: Main effects (no controls)

	Posterior	Intentions		App Data			Index	App: Week 7/8	
	Belief	Any	Days	Any	Days	Door	Overall	Days	Doors
Treatment	4.982***	0.013	-1.081***	-0.009	-0.166	-14.784*	-0.082	-0.144**	-6.764**
	(0.547)	(0.030)	(0.390)	(0.019)	(0.168)	(8.456)	(0.057)	(0.067)	(2.944)
Control mean	9.083	0.473	4.028	0.123	0.701	38.348	-0.029	0.297	13.318
Observations	1150	1148	1148	1163	1163	1163	1148	1163	1163

Notes: Table A3.21 presents treatment effects without pre-specified controls for respondents who underestimate the share of fellow party members who plan to participate in the party's door-to-door canvassing campaign. Outcomes are as defined in Table 3.2. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A3.22: Treatment effects and strength of connection to the party (no controls)

	Posterior	Intentions		App Data			Index	App: Week 7/8	
	Belief	Any	Days	Any	Days	Door	Overall	Days	Doors
Treatment	4.857***	0.014	-1.131***	-0.008	-0.174	-16.116*	-0.087	-0.147**	-7.155**
	(0.563)	(0.029)	(0.397)	(0.019)	(0.167)	(8.667)	(0.058)	(0.066)	(2.992)
Treatment × Strength of connection to party (PCA)	0.769	0.035	0.833**	0.002	0.141	16.600**	0.097*	0.059	5.336**
	(0.666)	(0.030)	(0.380)	(0.015)	(0.153)	(7.314)	(0.051)	(0.066)	(2.364)
Strength of connection to party (PCA)	-1.423***	0.034	-0.225	0.006	-0.041	-10.916*	-0.011	-0.017	-2.883
	(0.504)	(0.021)	(0.334)	(0.012)	(0.139)	(6.563)	(0.043)	(0.063)	(2.116)
Control mean	9.083	0.473	4.028	0.123	0.701	38.348	-0.029	0.297	13.318
Observations	1150	1148	1148	1163	1163	1163	1148	1163	1163

Notes: Table A3.22 presents treatment effects interacted with a summary measure of connectedness to the party without pre-specified controls for respondents who underestimate the share of fellow party members who plan to participate in the party's door-to-door canvassing campaign. The measure for strength of connection to the party is the first principal component of three pre-specified dimensions of heterogeneity (previous canvassing experience, party membership and party membership duration) obtained from a principal component analysis. This summary measure is standardized to have mean zero and standard deviation one. Outcomes are as defined in Table 3.2. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A3.23: Treatment effects and career concerns (no controls)

	Posterior	Intentions		App Data			Index	App: Week 7/8	
	Belief	Any	Days	Any	Days	Door	Overall	Days	Doors
Treatment	4.961***	0.011	-1.108***	-0.010	-0.176	-15.632*	-0.088	-0.149**	-7.031**
	(0.550)	(0.030)	(0.391)	(0.019)	(0.169)	(8.556)	(0.057)	(0.067)	(2.974)
Treatment × Predicted career concerns	-0.626	-0.039	-0.759*	0.004	-0.122	-15.948**	-0.088*	-0.107**	-5.941***
	(0.546)	(0.030)	(0.388)	(0.018)	(0.145)	(7.334)	(0.053)	(0.052)	(2.242)
Predicted career concerns	0.468	0.042**	0.684**	0.034***	0.229**	20.431***	0.136***	0.115***	6.512***
	(0.338)	(0.021)	(0.330)	(0.013)	(0.098)	(5.699)	(0.039)	(0.042)	(1.809)
Control mean	9.083	0.473	4.028	0.123	0.701	38.348	-0.029	0.297	13.318
Observations	1150	1148	1148	1163	1163	1163	1148	1163	1163

Notes: Table A3.23 presents treatment effects interacted with a supporter's career concerns without pre-specified controls for respondents who underestimate the share of fellow party members who plan to participate in the party's door-to-door canvassing campaign. We do not directly observe career concerns for participants in our experiment. We therefore predict career concerns using a post-election survey administered two months after the election to the same population of party supporters (excluding individuals who participated in the main experiment). Career concerns are measured as choosing "It is possible" as an answer to at least one the following questions "Can you imagine yourself taking on a political position within the party?" and "Can you imagine yourself running for office for this party?" (the answer options were "No", "It is possible", and "I am already doing this"). We use a post-LASSO algorithm to predict career concerns for our experimental sample based on pre-determined covariates (Belloni and Chernozhukov, 2013). The predicted variable is standardized to have mean zero and standard deviation one. The results of the post-LASSO regression are displayed in Table A3.8. Outcomes are as defined in Table 3.2. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A3.24: IV estimates: post-treatment beliefs and effort choices (no controls)

	Intentions		App Data			Index	App: Week 7/8	
	Any	Days	Any	Days	Door	Overall	Days	Doors
Panel A: Main Effect (IV)								
Posterior	0.003	-0.218***	-0.002	-0.033	-2.971*	-0.017	-0.029**	-1.364**
	(0.006)	(0.083)	(0.004)	(0.034)	(1.760)	(0.012)	(0.014)	(0.619)
Observations	1148	1148	1163	1163	1163	1148	1163	1163
Panel B: Weak connection to party (IV)								
Posterior	-0.002	-0.350**	-0.001	-0.066	-5.788**	-0.031*	-0.043**	-2.438**
	(0.009)	(0.142)	(0.006)	(0.051)	(2.921)	(0.019)	(0.022)	(1.013)
Observations	547	547	554	554	554	547	554	554
Panel C: Strong connection to party (IV)								
Posterior	0.007	-0.109	-0.003	-0.006	-0.665	-0.004	-0.017	-0.483
	(0.008)	(0.103)	(0.005)	(0.047)	(2.234)	(0.015)	(0.019)	(0.805)
Observations	601	601	609	609	609	601	609	609
Panel D: Low career concerns (IV)								
Posterior	0.010	-0.061	-0.001	0.006	0.703	0.004	0.004	0.219
	(0.008)	(0.106)	(0.005)	(0.040)	(1.476)	(0.013)	(0.016)	(0.595)
Observations	547	547	554	554	554	547	554	554
Panel E: High career concerns (IV)								
Posterior	-0.005	-0.382***	-0.002	-0.074	-6.857**	-0.038*	-0.064**	-3.036**
	(0.009)	(0.134)	(0.006)	(0.058)	(3.410)	(0.020)	(0.025)	(1.184)
Observations	601	601	609	609	609	601	609	609

Notes: Table A3.24 presents IV estimates for the impact of post-treatment beliefs without pre-specified controls for all underestimators (Panel A). Post-treatment beliefs are instrumented with the treatment indicator. Parallel to Table 3.5, Panels (B) and (C) distinguish between supporters with weak and strong connections to the party, whereas panels (D) and (E) distinguish between supporters with low and high career concerns. Outcomes are as defined in Table 3.2.

Table A3.25: Treatment effects on application download (no controls)

	Survey	App Data		
	click on app-link	within 24h	within one week	any time after survey
Treatment	-0.009	0.004	-0.006	0.009
	(0.011)	(0.008)	(0.009)	(0.014)
Observations	1163	1163	1163	1163
Control group mean	0.04	0.02	0.03	0.05

Notes: Table A3.25 presents treatment effects on application download without pre-specified controls for respondents who underestimate the share of fellow party members who plan to participate in the party's door-to-door canvassing campaign.

Table A3.26: Treatment effects by canvassing experience (no controls)

	Posterior	Intentions		App Data			Index	App: Week 7/8	
	Belief	Any	Days	Any	Days	Door	Overall	Days	Doors
Panel A: Inexperienced supporters									
Treatment	4.641***	0.002	-1.259***	0.001	-0.258	-23.253**	-0.111	-0.190**	-10.656***
	(0.736)	(0.037)	(0.482)	(0.023)	(0.190)	(10.407)	(0.070)	(0.082)	(3.679)
Control mean	9.872	0.376	3.382	0.101	0.657	39.110	-0.126	0.300	14.187
Observations	701	700	700	710	710	710	700	710	710
Panel B: Experienced supporters									
Treatment	5.517***	0.029	-0.810	-0.023	-0.021	-1.482	-0.037	-0.072	-0.656
	(0.794)	(0.045)	(0.642)	(0.033)	(0.312)	(14.327)	(0.096)	(0.114)	(4.883)
Control mean	7.850	0.626	5.041	0.158	0.770	37.158	0.122	0.293	11.959
Observations	449	448	448	453	453	453	448	453	453

Notes: Table A3.26 presents treatment effects without pre-specified controls for respondents who underestimate the share of fellow party members who plan to participate in the party's door-to-door canvassing campaign. Panel A contains the sample of supporters without previous canvassing experience, Panel B contains the sample of supporters with previous canvassing experience. Outcomes are as defined in Table 3.2.

Table A3.27: Treatment effects by party membership (no controls)

	Posterior	Intentions		App Data			Index	App: Week 7/8	
	Belief	Any	Days	Any	Days	Door	Overall	Days	Doors
Panel A: No party member									
Treatment	5.025***	0.010	-1.082**	-0.021	-0.317*	-17.831**	-0.113*	-0.175**	-7.713**
	(0.607)	(0.033)	(0.420)	(0.021)	(0.178)	(8.819)	(0.061)	(0.077)	(3.342)
Control mean	8.658	0.473	3.950	0.126	0.724	35.719	-0.033	0.308	13.519
Observations	910	908	908	920	920	920	908	920	920
Panel B: Party member									
Treatment	4.953***	0.025	-1.049	0.038	0.429	-1.564	0.044	-0.023	-3.059
	(1.229)	(0.065)	(0.976)	(0.044)	(0.454)	(23.001)	(0.146)	(0.135)	(6.342)
Control mean	10.623	0.475	4.311	0.113	0.621	47.782	-0.016	0.258	12.597
Observations	240	240	240	243	243	243	240	243	243

Notes: Table A3.27 presents treatment effects without pre-specified controls for respondents who underestimate the share of fellow party members who plan to participate in the party's door-to-door canvassing campaign. Panel A contains the sample of non-party members, Panel B contains the sample of party members. Outcomes are as defined in Table 3.2.

Table A3.28: Treatment effects by party membership duration (no controls)

	Posterior	Intentions		App Data			Index	App: Week 7/8	
	Belief	Any	Days	Any	Days	Door	Overall	Days	Doors
Panel A: Below med. membership dur.									
Treatment	4.697***	-0.022	-1.687***	0.003	-0.187	-23.340*	-0.133	-0.202**	-10.191**
	(0.816)	(0.042)	(0.580)	(0.029)	(0.248)	(14.146)	(0.088)	(0.093)	(4.557)
Control mean	9.900	0.523	4.674	0.141	0.827	52.676	0.076	0.363	17.359
Observations	575	574	574	582	582	582	574	582	582
Panel B: Above med. membership dur.									
Treatment	5.248***	0.047	-0.485	-0.021	-0.146	-6.317	-0.033	-0.086	-3.355
	(0.728)	(0.042)	(0.520)	(0.024)	(0.226)	(9.210)	(0.072)	(0.095)	(3.726)
Control mean	8.278	0.424	3.392	0.105	0.575	24.070	-0.133	0.232	9.291
Observations	575	574	574	581	581	581	574	581	581

Notes: Table A3.28 presents treatment effects without pre-specified controls for respondents who underestimate the share of fellow party members who plan to participate in the party's door-to-door canvassing campaign. Panel A contains the sample of supporters who have a below median party membership duration (including non-members), Panel B contains the sample of supporters who have above median membership duration. Outcomes are as defined in Table 3.2.

Table A3.29: Treatment effects by application download (no controls)

	Posterior	Intentions		App Data			Index	App: Week 7/8	
	Belief	Any	Days	Any	Days	Door	Overall	Days	Doors
Panel A: No app download									
Treatment	4.238***	0.011	-0.719*	-0.002	-0.142	-8.305*	-0.051	-0.091*	-3.107*
	(0.669)	(0.033)	(0.415)	(0.011)	(0.095)	(4.548)	(0.040)	(0.050)	(1.715)
Control mean	9.438	0.366	2.908	0.026	0.206	10.848	-0.304	0.114	4.241
Observations	844	842	842	855	855	855	842	855	855
Panel B: App download									
Treatment	7.042***	-0.038	-2.589***	-0.071	-0.463	-44.511	-0.307*	-0.366*	-20.459**
	(0.882)	(0.048)	(0.854)	(0.056)	(0.546)	(28.324)	(0.158)	(0.205)	(9.929)
Control mean	8.000	0.799	7.439	0.418	2.206	121.823	0.808	0.851	40.872
Observations	306	306	306	308	308	308	306	308	308

Notes: Table A3.29 presents treatment effects without pre-specified controls for respondents who underestimate the share of fellow party members who plan to participate in the party's door-to-door canvassing campaign. Panel A contains the sample of supporters who had not downloaded the application before the treatment, Panel B contains the sample of supporters who had downloaded the application before the treatment. Outcomes are as defined in Table 3.2.

Conclusion

The research presented in the first two chapters of this thesis considers the role of economic shocks in developing countries. The first chapter demonstrates that agricultural labor markets in Indonesia exhibit frictions in the wage-adjustment process in response to productivity shocks. However, the nominal wage rigidities I find are weaker than those found in other contexts (Kaur, 2019). I propose a theoretical framework that provides a potential explanation for these different findings. While the empirical results presented in this thesis are in line with my proposed explanation, I cannot draw firm conclusions about the underlying mechanisms. As mentioned in Chapter 1, fairness norms around both equality and equity are found to be an important factor contributing to worker effort (Fehr and Gächter, 2000b; Breza et al., 2017; Abeler et al., 2010). Fairness norms of workers are a plausible explanation of the results, however, the friction could also be caused by firms' suboptimal wage setting behavior.

However, there is little direct evidence on the structure of wage-setting processes at the micro-level.[24]. This is particularly relevant for developing countries with a large number of owner-operated SMEs because they are less likely to be perfect profit maximizers. The literature on behavioral firms is just emerging and, I think, has the potential to significantly broaden our understanding of labor markets in developing countries (Kremer et al., 2019).[25]

In future research, I plan to investigate the structure of all stages of the hiring processes, including wage-setting, from a firm's perspective. The aim of this research agenda is twofold. First, I want to identify inefficient decision making in SMEs that limits their productivity and growth. Second, I want to explore biases in the design of the hiring process and decisions taken throughout the hiring process that lead to the economic exclusion of specific populations (Beaman et al., 2018). This could be directly through discrimination or indirectly through the design of the process (eg Abebe et al., 2019). A better understanding of these processes would help

[24] There is a large literature on workers' reaction to different compensation schemes (see Bandiera et al., 2011, for an overview) However, exogenously changing workers compensation abstracts from firms endogenous wage-setting behavior.

[25] Behavioral decision making in firms also affects product markets and investment decisions (e.g. Shlain, 2019; Hanna et al., 2014).

policymakers to create more efficient and more inclusive labor market institutions, helping both workers and businesses.

My second chapter explores the relationship between economic shocks, conditional cash transfers, and suicides. The link between economic conditions and mental health has been relatively well researched. However, there is relatively little causal research on the impact of mental health on economic behavior and outcomes in developing countries, despite its potentially large effects (Lund et al., 2019). In particular, there is little evidence on how the availability of psychological treatment affects individuals' personal and economic resilience when facing negative economic shocks. Studying these effects of psychotherapy would be an important academic contribution and be relevant to policymakers globally.

The third chapter of my thesis is part of a larger agenda exploring political behavior in democratic systems. We document strategic substitutability between an activists' own effort choice and the effort choices of her peers. However, we also document a large degree of heterogeneity suggesting that multiple, competing motives might be at play. This is in line with evidence on differences in strategic interactions across movements in the same context (Hager et al., 2019b). This emphasizes the need to model different types of agents, or multiple motives simultaneously as recently conceptualized by Jia and Persson (2017). My coauthors and I plan to conduct further field experiments in diverse contexts to understand the distribution of motives for political activism across contexts and individuals.

* * *

References

Abebe, Girum, Stefano Caria, and Esteban Ortiz-ospina, "The Selection of Talent," *Working Paper*, 2019.

_ , _ , Marcel Fafchamps, Paolo Falco, Simon Franklin, and Simon Quinn, "Anonymity or Distance? Job Search and Labour Market Exclusion in a Growing African City," *Working Paper*, 2018.

Abel, Martin, Rulof Burger, and Patrizio Piraino, "The Value of Reference Letters - Experimental Evidence from South Africa," *American Economic Journal: Applied Economics*, 2019, *forthcoming*.

Abeler, Johannes, Steffen Altmann, Sebastian Kube, and Matthias Wibral, "Gift Exchange and Workers' Fairness Concerns: When Equality is Unfair," *Journal of the European Economic Association*, 2010, *8* (6), 1299–1324.

Acemoglu, Daron, Tarek A. Hassan, and Ahmed Tahoun, "The Power of the Street: Evidence from Egypt's Arab Spring," *Review of Financial Studies*, 2018, *31* (1), 1–42.

Adhvaryu, Achyuta, James Fenske, and Anant Nyshadham, "Early Life Circumstance and Adult Mental Health," *Journal of Political Economy*, 2019, *127* (4), 1516–1549.

Akerlof, George A. and Janet L. Yellen, "The Fair Wage-Effort Hypothesis and Unemployment," *The Quarterly Journal of Economics*, 1990, *105* (2), 255–283.

Allcott, Hunt, "Site Selection Bias in Program Evaluation," *The Quarterly Journal of Economics*, 2015, *130* (3), 1117 – 1165.

Angelucci, Manuela and Giacomo De Giorgi, "Indirect Effects of an Aid Program: How Do Cash Transfers Affect Ineligibles' Consumption?," *American Economic Review*, 2009, *99* (1), 486–508.

Appleby, Louis, Jayne Cooper, Tim Amos, and Brian Faragher, "Psychological Autopsy Study of Suicides by People Aged Under 35," *British Journal of Psychiatry*, 1999, *175*, 168–174.

Baicker, Katherine, Sarah L. Taubman, Heidi L. Allen, Mira Bernstein, Jonathan H. Gruber, Joseph P. Newhouse, Eric C. Schneider, Bill J. Wright, Alan M. Zaslavsky, and Amy N. Finkelstein, "The Oregon Experiment — Effects of Medicaid on Clinical Outcomes," *New England Journal of Medicine*, 2013, *368* (18), 1713–1722.

Baird, Sarah, Craig McIntosh, and Berk Özler, "Cash or Condition? Evidence from a Cash Transfer Experiment," *The Quarterly Journal of Economics*, 2011, *126* (4), 1709–1753.

_ , Jacobus De Hoop, and Berk Özler, "Income Shocks and Adolescent Mental Health," *Journal of Human Resources*, 2013, *48* (2), 370–403.

Bandiera, Oriana, Iwan Barankay, and Imran Rasul, "Field Experiments with Firms," *Journal of Economic Perspectives*, 2011, *25* (3), 63–82.

Banerjee, Abhijit V., Rema Hanna, Gabriel E. Kreindler, and Benjamin A. Olken, "Debunking the Stereotype of the Lazy Welfare Recipient: Evidence from Cash Transfer Programs," *The World Bank Research Observer*, 2017, *32* (2), 155–184.

Barattieri, Alessandro, Susanto Basu, and Peter Gottschalk, "Some Evidence on the Importance of Sticky Prices," *American Economic Journal: Macroecnomics*, 2014, *6* (1), 70–101.

Bassi, Vittorio and Aisha Nansamba, "Screening and Signaling Non-Cognitive Skills: Experimental Evidence from Uganda," *Working Paper*, 2019.

Bazzi, Sam and Matthew Gudgeon, "The Political Boundaries of Ethnic Divisions," *Working Paper*, 2019.

Bazzi, Samuel, "Wealth Heterogeneity and the Income Elasticity of Migration," *American Economic Journal: Applied Economics*, April 2017, *9* (2), 219–55.

Beaman, Lori, Niall Keleher, and Jeremy Magruder, "Do Job Networks Disadvantage Women? Evidence from a Recruitment Experiment in Malawi," *Journal of Labor Economics*, 2018, *36* (1), 121–157.

Becker, Sascha O. and Ludger Woessmann, "Social Cohesion, Religious Beliefs, and the Effect of Protestantism on Suicide," *The Review of Economics and Statistics*, 2018, *100* (3), 377–391.

Beegle, Kathleen, Calogero Carletto, and Kristen Himelein, "Reliability of Recall in Agricultural Data," *Journal of Development Economics*, 2012, *98* (1), 34–41.

Belloni, Alexandre and Victor Chernozhukov, "Least Squares After Model Selection in High-Dimensional Sparse Models," *Bernoulli*, 2013, *19* (2), 521–547.

Bénabou, Roland and Jean Tirole, "Incentives and Prosocial Behavior," *American Economic Review*, December 2006, *96* (5), 1652–1678.

Benigno, Pierpaolo and Luca Antonio Ricci, "The Inflation-Output Trade-off with Downward Wage Rigidities," *American Economic Review*, 2011, *101* (4), 1436–1466.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan, "How Much Should We Trust Differences-in-Differences Estimates?," *The Quarterly Journal of Economics*, 2004, *119* (1), 249–275.

Bewley, Truman F., "A Depressed Labor Market as Explained by Participants," *American Economic Review*, 1995, *85* (2), 250–254.

Bils, Mark J., "Real Wages over the Business Cycle: Evidence from Panel Data," *Journal of Political Economy*, 1985, *93* (4), 666–689.

Blüml, Victor, Nestor D. Kapusta, Stephan Döring, Elmar Brähler, Birgit Wagner, and Anette Kersting, "Personality Factors and Suicide Risk in a Representative Sample of the German General Population," *PLoS One*, 2013, *8* (10), e76646.

Boldrini, Maura and J. John Mann, "Depression and Suicide," in Michael J. Zigmond, Lewis P. Rowland, and Joseph T. Coyle, eds., *Neurobiology of Brain Disorders*, London: Academic Press, 2015, chapter 43, p. 709–729.

Bound, John, Charles Brown, and Nancy Mathiowetz, "Measurement Error in Survey Data," in "Handbook of Econometrics," Vol. 5, Elsevier Masson SAS, 2001, chapter 59, pp. 3705–3843.

Breza, Emily, Supreet Kaur, and Yogita Shamdasani, "The Morale Effects of Pay Inequality," *The Quarterly Journal of Economics*, 10 2017, *133* (2), 611–663.

Bryan, Gharad and Melanie Morten, "The Aggregate Productivity Effects of Internal Migration: Evidence from Indonesia," *Journal of Political Economy*, 2019, *forthcoming*.

Bursztyn, Leonardo, Davide Cantoni, Patricia Funk, and Noam Yuchtman, "Polls, the Press, and Political Participation: The Effects of Anticipated Election Closeness on Voter Turnout," *Working Paper*, 2018.

Cahyadi, Nur, Rema Hanna, Benjamin A. Olken, Rizal Adi Prima, Elan Satriawan, and Ekki Syamsulhakim, "Cumulative Impacts of Conditional Cash Transfer Programs: Experimental Evidence from Indonesia," *NBER Working Paper No. 24670*, 2018.

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller, "Bootstrap-Based Improvements for Inference with Clustered Errors," *Review of Economics and Statistics*, 2008, *90* (3), 414–427.

Campaniello, Nadia, Theodoros M. Diasakos, and Giovanni Mastrobuoni, "Rationalizable Suicides: Evidence from Changes in Inmates' Expected Length of Sentence," *Journal of the European Economic Association*, 2017, *15* (2), 388–428.

Cantoni, Davide, David Y Yang, Noam Yuchtman, and Y Jane Zhang, "The Fundamental Determinants of Anti-Authoritarianism," *Working Paper*, 2016.

_ , _ , _ , and _ , "Protests as Strategic Games: Experimental Evidence from Hong Kong's Antiauthoritarian Movement," *The Quarterly Journal of Economics*, 01 2019, *134* (2), 1021–1077.

Card, David, "Unexpected Inflation, Real Wages, and Employment Determination in Union Contracts," *American Economic Review*, 1990, *80* (4), 669–688.

Carleton, Tamma, "Crop-Damaging Temperatures Increase Suicide Rates in India," *Proceedings of the National Academy of Sciences*, 2017, *114* (33), 8746–8751.

Carranza, Eliana, Rob Garlick, Kate Orkin, and Neil Rankin, "Overcoming Barriers to Firms Hiring Young, Inexperienced Work-Seekers," *Working Paper*, 2019.

Cavanagh, Jonathan T., Alan J. Carson, Michael Sharpe, and Stephen M. Lawrie, "Psychological Autopsy Studies of Suicide: A Systematic Review," *Psychological Medicine*, 2003, *33* (3), 395–405.

Cesarini, David, Erik Lindqvist, Robert Östling, and Björn Wallace, "Wealth, Health, and Child Development: Evidence from Administrative Data on Swedish Lottery Players," *The Quarterly Journal of Economics*, 2015, *131* (2), 687–738.

Chang, Shu-Sen, David Stuckler, Paul Yip, and David Gunnell, "Impact of 2008 Global Economic Crisis on Suicide: Time Trend Study in 54 Countries," *British Medical Journal*, 2013, *347* (7925), f5239.

Chen, By Yan, F Maxwell Harper, Joseph Konstan, and Sherry Xin Li, "Social Comparisons and Contributions to Online Communities: A Field Experiment on MovieLens," *American Economic Review*, 2010, *100* (4), 1358–1398.

Clarida, Richard H., Jordi Gali, and Mark Gertler, "The Science of Monetary Policy: A New Keynesian Perspective," *Journal of Economic Literature*, 1999, *37* (4), 1661–1707.

Coate, Stephen and Michael Conlin, "A Group Rule-Utilitarian Approach to Voter Turnout: Theory and Evidence," *American Economic Review*, 2004, *94* (5), 1476–1504.

Colmer, Jonathan, "Weather, Labour Reallocation, and Industrial Production: Evidence from India," *Working Paper*, 2018.

Conley, Timothy G., "GMM Estimation with Cross Sectional Dependence," *Journal of Econometrics*, 1999, *92* (1), 1–45.

Cruz, Cesi, Philip Keefer, and Julien Labonne, "Buying Informed Voters: New Effects of Information on Voters and Candidates," *Working Paper*, 2019.

Cutler, David M., Edward L. Glaeser, and Karen E. Norberg, "Explaining the Rise in Youth Suicide," in Jonathan Gruber, ed., *Risky Behavior Among Youths: An Economic Analysis*, Chicago, Ill: University of Chicago Press, 2001, pp. 219–270.

Daly, Mary C, Daniel J. Wilson, and Norman J. Johnson, "Relative Status and Well-Being: Evidence from US Suicide Deaths," *The Review of Economics and Statistics*, 2013, *95* (5), 1480–1500.

Das, Jishnu, Quy-Toan Do, Jed Friedman, David McKenzie, and Kinnon Scott, "Mental Health and Poverty in Developing Countries: Revisiting the Relationship," *Social Science & Medicine*, 2007, *65* (3), 467–480.

de Quidt, Jonathan, Johannes Haushofer, and Christopher Roth, "Measuring and Bounding Experimenter Demand," *American Economic Review*, 2018, *108* (11), 3266–3302.

Deaton, Angus, "Instruments, Randomization, and Learning about Development," *Journal of Economic Literature*, 2010, *48* (2), 424–455.

DellaVigna, Stefano, John A. List, Ulrike Malmendier, and Gautam Rao, "Voting to Tell Others," *The Review of Economic Studies*, 2017, *84* (1), 143–181.

Devoto, Florencia, Esther Duflo, Pascaline Dupas, William Parienté, and Vincent Pons, "Happiness on Tap: Piped Water Adoption in Urban Morocco," *American Economic Journal: Economic Policy*, 2012, *4* (4), 68–99.

Di Tella, Rafael, Ricardo Perez-Truglia, Andres Babino, and Mariano Sigman, "Conveniently Upset: Avoiding Altruism by Distorting Beliefs about Others' Altruism," *American Economic Review*, 2015, *105* (11), 3416–3442.

Downs, Anthony, "An Economic Theory of Political Action in a Democracy," *Journal of Political Economy*, 1957, *65* (2), 135–150.

Duncan, Greg J. and Daniel H Hill, "An Investigation of the Extent and Consequences of Measurement Error in Labor-Economic Survey Data," *Journal of Labor Economics*, 1985, *3* (4), 508–532.

Enikolopov, Ruben, Alexey Makarin, and Maria Petrova, "Social Media and Protest Participation: Evidence from Russia," *Working Paper*, 2019.

_ , _ , _ , and Leonid Polishchuk, "Social Image, Networks, and Protest Participation," *Working Paper*, 2017.

Falcon, Walter P., Rosamond L. Naylor, Whitney L. Smith, Marshall B. Burke, and Ellen B. McCullough, "Using Climate Models to Improve Indonesian Food Security," *Bulletin of Indonesian Economic Studies*, 2004, *40* (3), 355–377.

Falk, Armin and Urs Fischbacher, "A Theory of Reciprocity," *Games and Economic Behavior*, 2006, *54* (2), 293 – 315.

Feddersen, Timothy and Alvaro Sandroni, "A Theory of Participation in Elections," *American Economic Review*, 2006, *96* (4), 1271–1282.

Fehr, Ernst and Simon Gächter, "Cooperation and Punishment in Public Goods Experiments," *American Economic Review*, 2000, *90* (4), 980–994.

_ **and Simon Gächter**, "Fairness and Retaliation: The Economics of Reciprocity," *The Journal of Economic Perspectives*, 2000, *14* (3), 159–181.

_ **, Lorenz Goette, and Christian Zehnder**, "A Behavioral Account of the Labor Market: The Role of Fairness Concerns," *Annual Review of Economics*, 2009, *1* (1), 355–384.

Frankenberg, Elizabeth and James P Smith, "Economic Shocks , Wealth, and Welfare," *The Journal of Human Resources*, 2003, *38* (2), 280–321.

_ **and Karoly L**, "The 1993 Indonesian Family Life Survey: Overview and Field Report," Technical Report, RAND, Santa Monica, CA 1993.

Franklin, Simon, "Location, search costs and youth unemployment: experimental evidence from transport subsidies," *The Economic Journal*, 2017, *128* (614), 2353–2379.

_ **and Julien Labonne**, "Economic Shocks and Labor Market Flexibility," *Journal of Human Resources*, 2019, *54* (1), 171–199.

Frederick, Shane and George Loewenstein, "Hedonic Adaptation," in "The Foundations of Hedonic Psychology," New York, NY: Russell Sage Foundation, 1999, pp. 302–329.

Frey, Bruno S. and Stephan Meier, "Social Comparisons and Pro-Social Behavior: Testing Conditional Cooperation in a Field Experiment," *American Economic Review*, 2004, *94* (5), 1717–1722.

Friedman, Jed and Duncan Thomas, "Psychological Health Before, During, and After an Economic Crisis: Results from Indonesia, 1993–2000," *The World Bank Economic Review*, 2009, *23* (1), 57–76.

Galí, Jordi:, *Monetary Policy, Inflation, and the Business Cycle: An Introduction to the New Keynesian Framework*, Princeton University Press, 2009.

Galiani, Sebastian, Paul J Gertler, and Raimundo Undurraga, "The Half-Life of Happiness: Hedonic Adaptation in the Subjective Well-Being of Poor Slum Dwellers to the Satisfaction of Basic Housing Needs," *Journal of the European Economic Association*, 01 2018, *16* (4), 1189–1233.

Gallus, Jana, "Fostering Public Good Contributions with Symbolic Awards: A Large-Scale Natural Field Experiment at Wikipedia," *Management Science*, 2017, *63* (12), 3999–4015.

Gardner, Jonathan and Andrew J. Oswald, "Money and Mental Well-being: A Longitudinal Study of Medium-sized Lottery Wins," *Journal of Health Economics*, 2007, *26* (1), 49–60.

Gerber, Alan, Mitchell Hoffman, John Morgan, and Collin Raymond, "One in a Million: Field Experiments on Perceived Closeness of the Election and Voter Turnout," *American Economic Journal: Applied Economics*, 2019, *forthcoming*.

Gerber, Alan S, Daniel P Kessler, and Marc Meredith, "The Persuasive Effects of Direct Mail: A Regression Discontinuity Based Approach," *The Journal of Politics*, 2011, *73* (01), 140–155.

Gignoux, Jérémie and Marta Menéndez, "Benefit in the Wake of Disaster: Long-Run Effects of Earthquakes on Welfare in Rural Indonesia," *Journal of Development Economics*, 2015, *118*, 26–44.

González, Felipe, "Collective Action in Networks: Evidence from the Chilean Student Movement," *Working Paper*, 2018.

Green, Donald P, Mary C McGrath, and Peter M Aronow, "Field Experiments and the Study of Voter Turnout," *Journal of Elections, Public Opinion & Parties*, 2013, *23* (1), 27–48.

Gröger, Andre and Yanos Zylberberg, "Internal Migration as a Risk-coping Strategy: Evidence from a Typhoon," *American Economic Journal: Applied Economics*, 2016, *8* (2), 123–153.

Hager, Anselm, Lukas Hensel, Johannes Hermle, and Christopher Roth, "Does Party Competition Affect Political Engagement?," *Working Paper*, 2019.

_ , _ , _ , **and** _ , "Strategic Interdependence in Political Movements and Countermovements," *Working Paper*, 2019.

Hanna, Rema, Joshua Schwartzstein, and Sendhil Mullainathan, "Learning Through Noticing: Theory and Evidence From a Field Experiment," *The Quarterly Journal of Economics*, 2014, pp. 1311–1353.

Hardin, Russell, *Collective Action*, Routledge, 2015.

Harrison, Glenn W and John A List, "Field Experiments," *Journal of Economic Literature*, 2004, *42* (4), 1009–1055.

Haushofer, Johannes and Ernst Fehr, "On the Psychology of Poverty," *Science*, 2014, *344* (6186), 862–867.

_ **and Jeremy Shapiro**, "The Short-Term Impact of Unconditional Cash Transfers to the Poor: Experimental Evidence from Kenya," *The Quarterly Journal of Economics*, 2016, *131* (4), 1973–2042.

Hawton, Keith, Carolina Casañas i Comabella, Camilla Haw, and Kate Saunders, "Risk Factors for Suicide in Individuals with Depression: a Systematic Review," *Journal of Affective Disorders*, 2013, *147* (1), 17–28.

Hedström, Peter, Ka-Yuet Liu, and Monica K. Nordvik, "Interaction Domains and Suicide: A Population-Based Panel Study of Suicides in Stockholm, 1991-1999," *Social Forces*, 2008, *87* (2), 713–740.

Jayachandran, Seema, "Selling Labor Low: Wage Responses to Productivity Shocks in Developing Countries," *Journal of Political Economy*, 2006, *114* (3), 538–575.

Jia, Ruixue and Torsten Persson, "Individual vs. Social Motives in Identity Choice: Theory and Evidence from China," *Working Paper*, 2017.

Joiner, Thomas, *Why People Die by Suicide*, Cambridge, MA: Harvard University Press, 2005.

Kahneman, Daniel, Jack L Knetsch, and Richard Thaler, "Fairness as a Constraint on Profit Seeking: Entitlements in the Market," *American Economic Review*, 1986, *76* (4), 728–741.

Kaur, Supreet, "Nominal Wage Rigidity in Village Labor Markets," *American Economic Review*, 2019, *forthcoming.*

Keane, Michael P and Eswar S Prasad, "The Employment and Wage Effects of Oil Price Changes: A Sectoral Analysis," *The Review of Economics and Statistics*, 1996, *78* (3), 389–400.

Keane, Michael, Robert Moffitt, and David Runkle, "Real Wages over the Business Cycle: Estimating the Impact of Heterogeneity with Micro Data," *Journal of Political Economy*, 1988, *96* (6), 1232–1266.

Kendall, Chad, Tommaso Nannicini, and Francesco Trebbi, "How Do Voters Respond to Information? Evidence from a Randomized Campaign," *American Economic Review*, 2014, *105* (1), 322–353.

Kirchberger, Martina, "Natural Disasters and Labour Markets," *Journal of Development Economics*, 2017, *125*, 40–58.

Kleemans, Marieke and Jeremy Magruder, "Labour Market Responses To Immigration: Evidence From Internal Migration Driven By Weather Shocks," *The Economic Journal*, 2018, *128* (613), 2032–2065.

Kling, Jeffrey R., Jeffrey B. Liebman, and Lawrence F. Katz, "Experimental Analysis of Neighborhood Effects," *Econometrica*, 2007, *75* (1), 83–119.

Kochar, Anjini, "Explaining Household Vulnerability to Idiosyncratic Income Shocks," *American Economic Review*, 1995, *85* (2), 159–164.

Kremer, Michael, Gautam Rao, and Frank Schilbach, "Chapter 5 - Behavioral development economics," *Handbook of Behavioral Economics*, 2019, *2*, 345–458.

Kuhn, Peter, Peter Kooreman, Adriaan Soetevent, and Arie Kapteyn, "The Effects of Lottery Prizes on Winners and their Neighbors: Evidence from the Dutch Postcode Lottery," *American Economic Review*, 2011, *101* (5), 2226–2247.

LaFave, Daniel and Duncan Thomas, "Farms, Families, and Markets: New Evidence on Completeness of Markets in Agricultural Settings," *Econometrica*, 2016, *84* (5), 1917–1960.

Levine, David I. and Dean Yang, "The Impact of Rainfall on Rice Output in Indonesia," *NBER Working Paper No. 20302*, 2014.

Li, Qi and Jeffrey Scott Racine, *Nonparametric Econometrics Theory and Practice*, Princeton: Princeton University Press, 2006.

Lipsey, R. G. Lancaster, K, "The General Theory of Second Best," *The Review of Economic Studies*, 1956, *24* (1), 11–32.

Ludwig, Jens, Dave E Marcotte, and Karen Norberg, "Anti-Depressants and Suicide," *Journal of Health Economics*, 2009, *28* (3), 659–676.

Lund, Crick, Kate Orkin, Marc Witte, Thandi Davies, Johannes Haushofer, Judy Bass, Paul Bolton, Sarah Murray, Laura Murray, Wietse Tol, Graham Thornicroft, and Vikram Patel, "Economic Impacts of Mental Health Interventions in Low and Middle Income Countries: A Systematic Review and Meta-Analysis," *Working Paper*, 2019.

Maccini, Sharon and Dean Yang, "Under the Weather: Health, Schooling, and Economic Consequences of Early-life Rainfall," *American Economic Review*, 2009, *99* (3), 1006–1026.

Madestam, Andreas, Daniel Shoag, Stan Veuger, and David Yanagizawa-Drott, "Do Political Protests Matter? Evidence from the Tea Party Movement," *The Quarterly Journal of Economics*, 2013, *128* (4), 1633–1685.

Manacorda, Marco and Andrea Tesei, "Liberation Technology: Mobile Phones and Political Mobilization in Africa," *Working Paper*, 2018.

Mann, J. John, Christine Waternaux, Gretchen L. Haas, and Kevein M. Malone, "Toward a Clinical Model of Suicidal Behavior in Psychiatric Patients," *American Journal of Psychiatry*, 1999, *156* (2), 181–189.

McInerney, Melissa, Jennifer M. Mellor, and Lauren Hersch Nicholas, "Recession Depression: Mental Health Effects of the 2008 Stock Market Crash.," *Journal of Health Economics*, 2013, *32*, 1090–1104.

Naylor, Rosamond L, David S Battisti, Daniel J Vimont, Walter P Falcon, and Marshall B Burke, "Assessing risks of climate variability and climate change for Indonesian

rice agriculture.," *Proceedings of the National Academy of Sciences of the United States of America*, 2007, *104* (19), 7752–7757.

O'Connor, Rory C. and Matthew K. Nock, "The Psychology of Suicidal Behaviour," *Lancet Psychiatry*, 2014, *1* (1), 73–85.

Olson, Mancur, *The Logic of Collective Action: Public Goods and the Theory of Groups*, Schocken Books, 1965.

Ostrom, Elinor, "Collective Action and the Evolution of Social Norms," *The Journal of Economic Perspectives*, 2000, *14* (3), 137–158.

Palmfrey, Thomas R and Howard Rosenthal, "Participation and the Provision of Discrete Public Goods: a Strategic Analysis," *Journal of Public Economics*, 1984, *24* (2), 171–193.

Passarelli, Francesco and Guido Tabellini, "Emotions and Political Unrest," *Journal of Political Economy*, 2017, *125* (3), 903–946.

Paxson, Christina and Norbert Schady, "Cognitive Development Among Young Children in Ecuador the Roles of Wealth, Health, and Parenting," *Journal of Human Resources*, 2007, *42* (1), 49–84.

Perez-Truglia, Ricardo and Guillermo Cruces, "Partisan Interactions," *Journal of Political Economy*, 2017, *125* (4), 1208–1243.

Persson, Petra and Maya Rossin-Slater, "Family Ruptures, Stress, and the Mental Health of the Next Generation," *American Economic Review*, 2018, *108*, 1214–52.

Plasser, Fritz and Gunda Plasser, *Global Political Campaigning: A Worldwide Analysis of Campaign Professionals and Their Practices*, Greenwood Publishing Group, 2002.

Pons, Vincent, "Does Door-to-door Canvassing Affect Vote Shares? Evidence from a Countrywide Field Experiment in France," *Working Paper*, 2014.

Radloff, Lenore Sawyer, "The CES-D Scale: A Self-Report Depression Scale for Research in the General Population," *Applied Psychological Measurement*, 1997, *1* (3), 385–401.

Reeves, Aaron, David Stuckler, Martin McKee, David Gunnell, Shu-Sen Chang, and Sanjay Basu, "Increase in State Suicide Rates in the USA During Economic Recession," *The Lancet*, 2012, *380*, 1813–1814.

Rosenzweig, Mark R. and Christopher Udry, "Rainfall Forecasts, Weather, and Wages over the Agricultural Production Cycle," *American Economic Review*, 2014, *104* (5), 278–283.

Roy, Alex, "Are There Genetic Factors in Suicide?," *International Review of Psychiatry*, 1992, *4* (2), 169–175.

Santangelo, Gabriella, "Firms and Farms: The Local Effects of Farm Income on Firms' Demand ," *Workign Paper*, 2019.

Schilbach, Frank, Heather Schofield, and Sendhil Mullainathan, "The Psychological Lives of the Poor," *American Economic Review: Papers and Proceedings*, 2016, *106* (5), 435–40.

Shang, Jen and Rachel Croson, "A Field Eexperiment in Charitable Contribution: The Impact of Social Information on the Voluntary Provision of Public Goods," *The Economic Journal*, 2009, *119* (October), 1422–1439.

Shlain, Avner S, "More than a Penny's Worth: Left-Digit Bias and Firm Pricing," *Working Paper*, 2019.

Smith, James P., Duncan Thomas, Elizabeth Frankenberg, Kathleen Beegle, and Graciela Teruel, "Wages, Employment and Economic Shocks: Evidence from Indonesia," *Journal of Population Economics*, 2002, *15* (1), 161–193.

Stevenson, Betsey and Justin Wolfers, "Bargaining in the Shadow of the Law: Divorce Laws and Family Distress," *The Quarterly Journal of Economics*, 2006, *12* (121), 267–288.

The World Bank, *Indonesia Jobs Report Towards Better Jobs and Security for All*, Washington, DC: The World Bank, 2010.

_ , *Program Keluarga Harapan - Main Findings from the Impact Evaluation of Indonesia's Pilot Household Conditional Cash Transfer Program*, Washington, DC: The World Bank, 2011.

_ , *Protecting Poor and Vulnerable Households in Indonesia*, Washington, DC: The World Bank, 2012.

_ , *Towards a Comprehensive, Integrated, and Effective Social Assistance System in Indonesia*, Washington, DC: The World Bank, 2017.

_ , "World Development Indicators," 2019.

Uhlaner, Carole Jean, ""Relational Goods" and Participation: Incorporating Sociability into a Theory of Rational Action," *Public Choice*, 1989, *62* (3), 253–285.

WHO, *Preventing Suicide: A Global Imperative*, Geneva: World Health Organization, 2014.