Dictating Development?

The Effects of Local Institutions under Upward Accountability*

Maximilian Linek†

May 15, 2020

Abstract

This paper studies the effects of a local institution with and without upward accountability on individual behavior change in Rwanda. Throughout Rwanda, local leaders organize mandatory community meetings on Saturdays to discuss and resolve issues of community concern. I analyze the effects of meetings on contraceptive adoption and bed net acquisition, and leverage a reform that introduced performance incentives for local leaders. Both outcomes were unpopular among the population, but desired by the central government. For identification, I exploit quasi-experimental variation in meetings' attendance over time induced by exogeneous weather fluctuations. After the reform, I find that a rainy Saturday reduces the probability of contraceptive adoption by 18% and of bed net acquisition by 10% in the same month. Before the reform, rainfall on every weekday, including Saturdays, has no effect. This pattern for two incentivized, but otherwise unrelated behavior changes points to the reform as the common underlying shift. Finally, I present evidence suggesting that behavior change is involuntary. Overall, my findings challenge the presumed downward accountability of local institutions and indicate an interdependence with performance incentives.

JEL Classification: D71, D73, I12.
Keywords: Local Institutions; Performance Incentives; Family Planning; Mosquito Bed Nets.

*I thank Francesco Amodio, Marianna Battaglia, Konrad Burchardi, Ruben Durante, Andrea Guariso, Selim Gulesci, Leander Heldring, Macartan Humphreys, Leonardo Iacovone, Rajshri Jayaraman, Eliana La Ferrara, Julien Labonne, Kate Orkin, Dina Pomeranz, Pia Raffler, Johanna Rickne, Miri Stryjan, Christian Traxler, Diego Ubfal, Marijke Verpoorten, Christopher Woodruff and Ekaterina Zhuravskaya for many helpful comments and suggestions. I would also like to thank seminar and conference participants at AEL Zurich, Bocconi University, CSAE Oxford, EEA Cologne, and the Hertie School.

†The Hertie School; m.linek@phd.hertie-school.org

1 Introduction

Local institutions are a key to development because they regulate individual behavior. After rampant corruption in many centralized states, governments and aid donors have embraced local institutions as instruments for managing decentralized development (Mansuri and Rao 2013). The underlying tenet is that citizen participation in local institutions regulates the behavior of leaders through social capital. Recent research does not verify this mechanism.[1] It commonly attributes development to downward accountability and the absence thereof to elite capture. However, if elites control local institutions, development may just as well arise from leaders' ability to mandate targeted behaviors from community members. To understand how decentralized approaches to development achieve outcomes, it is important to investigate the effects of local institutions jointly with accountability. I do so in the East African country Rwanda.

Rwanda is an interesting case for studying local institutions and accountability. After the 1994 genocide, a new government took power and rebuilt the country. During the late 2000s, this government achieved spectacular progress toward the Millenium Development Goals, many of which require behavior change. This success is widely attributed to citizen participation in local institutions (UNDP 2014). However, Rwanda is also an authoritarian state in which, by definition, leaders control local institutions. By exploring the contribution of a local institution to development progress in Rwanda, this paper challenges the assumption of downward accountability in decentralized development. The findings arguably are relevant for understanding and designing development in many countries around the world as local institutions often have strong authoritarian elements even when the national state is considered a democracy.

In this paper, I study a local institution that is used for development in Rwanda and link its effects on behavior change to accountability.[2] The local institution is a community meeting that takes place on Saturdays and forms part of a traditional, mandatory community program called *Umuganda*. I trace accountability of and in this institution through its effects on behavior change in three ways. First, I analyze behavior changes that were desired by the central government, but unpopular at the local level. Second, I compare the effects when local leaders are upward accountable to when they are not. And third, I study factors related to enforcement. My results show that community meetings only change behavior of ordinary people toward development desired by the government if local leaders are upward accountable. They also suggest that behavior change is largely involuntary, which confirms upward accountability also within communities.

1. Research on local institutions in development has commonly attempted to strengthen bottom-up accountability through either providing additional information (Banerjee et al. 2010; Björkman Nyqvist et al. 2017; Björkman and Svensson 2009) or altering composition, involvement or capacity of people in the institution (Casey et al. 2012; Olken 2007; Pradhan et al. 2014). All of these studies essentially assume downward accountability of local institutions and track, at best, civic participation as a proxy.

2. 'Accountability' describes a relationship in which a principle delegates a task to an agent and aligns the agent's incentives through threat of enforcement. As a such, accountability is unobservable, but manifests in behavior.

In the analysis, I pursue the following three objectives. My first objective is to identify the causal effects of *Umuganda* meetings on two unrelated changes in individual behavior. The behavior changes, which were desired by the central government, but arguably unpopular, are modern contraceptive adoption and mosquito bed net acquisition. For identification, I exploit exogeneous weather fluctuations on meeting days. My second objective is to relate the effects to the accountability of local leaders. To do so, I compare the effects just before and after a reform that introduced performance contracts and, thus, increased local leaders' upward accountability to the central government. My third objective is to classify the mechanism of *Umuganda* meetings by whether it generates voluntary or involuntary behavior change. For this purpose, I analyze conception, i.e. women's timing of pregnancy, as a behavior change to evade modern contraceptive adoption and heterogeneity in popular support for mosquito bed nets.

I identify the causal effects of *Umuganda* meetings on behavior change through exogeneous variation in rainfall over time. The underlying idea is that rainfall on a meeting day reduces attendance. As a consequence, the meeting is impaired or canceled and cannot assume its function as a forum for discussion and problem solving, which can generate behavior change. Although, I lack data on *Umuganda* meetings, it is well known that these take place on Saturdays. That means, I can isolate their effect from general rainfall effects in estimations of the reduced-form.[3] My explanatory variables are the number of rainy Saturdays, Sundays, Mondays, etc. in a village and month that match my panel data on behavior change. The estimate on Saturday rainfall can be interpreted as the effect of a "failed" *Umuganda* meeting that has low or zero attendance. Rainfall on each of the other six weekdays serve as placebo tests and control for potential, general rainfall effects (e.g. related to income from agriculture).

The results show that in the first year after the introduction of performance contracts one failed *Umuganda* meeting in a month significantly reduces the probability of behavior change in the direction desired by the government in the same month. The relative effects are -18% for contraceptive adoption and -10% for bed net acquisition. These effects are large, considering that there could be up to 5 *Umuganda* meetings in a month. Rainy day counts for other weekdays are consistently insignificant. The results are very robust and hold for different rainfall thresholds that are used to define a rainy day. One concern may be that some other regular event is affected by Saturday rainfall. I will discuss this possibility and present evidence that rules out access as an alternative explanation for Saturday rainfall effects at least for modern contraceptive methods.

In the year before performance contracts, rainfall on Saturdays and any other weekday have no significant effect on behavior change. The difference in effects of Saturday rainfall with and without performance contracts is statistically significant for both contraceptive adoption

3. A related paper is Bonnier et al. (2019). These authors use the same identification strategy in cross-sectional data to estimate the effect of *Umuganda* on civilian participation in the 1994 genocide. Other studies in economics document and exploit the effect of rainfall on attendance at events (see e.g. Madestam et al. 2013; Fujiwara et al. 2016; Collins and Margo 2007).

and bed net acquisition. This same effect pattern suggests the introduction of performance contracts as the common underlying shift because both changes in behavior were incentivized, but are otherwise completely unrelated. Any alternative explanation must plausibly affect both outcomes, work through *Umuganda*, be nationwide, and have taken place around the same time. I am unaware of any other such alternative reform or shift. Put into context, the findings show that performance contracts, which strengthened upward accountability, turned *Umuganda* meetings into an effective tool for implementing behavior changes that were desired by the central government.

Finally, I find evidence suggesting that *Umuganda* meetings generate involuntary behavior change after the introduction of performance contracts. A failed *Umuganda* meeting significantly reduces the probability of conception by 8%, which is an effect in the same direction as that for contraceptive adoption. To yield this same sign result, the function of conception as an evasive behavior must outweigh the mechanical reduction from meetings increasing contraceptive adoption. As will be explained in more detail, this finding strongly suggests that contraceptive adoptions require enforcement and, thus, are to a large extent involuntary. The analysis of heterogeneity in popular support for bed net acquisitions exploits the fact that the incidence of mosquitoes and altitude are inversely related. The results indicates that the effect of *Umuganda* meetings is larger in communities at high altitudes with few or no mosquitoes and low popular support for bed nets. This finding is consistent with the notion that low popular support requires more enforcement to achieve behavior change, further corroborating the claim that meetings generate involuntary behavior change.

The paper contributes to the literature in several ways. First, it speaks to the literature on local institutions as instruments for development. Research in this field is predominantly experimental and attempts to empower citizens in project implementation through information and training (Olken 2007; Duflo et al. 2015; Björkman Nyqvist et al. 2017; Björkman and Svensson 2009). However, it is well known that elites often control local institutions (Reinikka and Svensson 2004; Anderson et al. 2015). Especially successful interventions may actually have supported leaders to monitor citizens rather than the other way around. Against this backdrop, Casey et al. (2012) unsuccessfully attempt to change local institutions, using participation as a proxy for downward accountability. While Banerjee et al. (2010) suggest the use of new institutions for development, Pradhan et al. (2014) indicate that success of existing institutions is linked to elite support and power. I contribute to this literature by challenging the commonly assumed downward accountability of local institutions. I document strong development through a local institution, community meetings, under upward accountability. Similar to Acemoglu et al. (2014), I also show that participation in local institutions fosters leaders' control over community behavior, invalidating it as an indicator for downward accountability. My findings are novel among microeconomic studies of local institutions, but consistent with cross-country evidence on development after decentralization (see e.g. Enikolopov and Zhuravskaya 2007).

The paper also speaks to the literature on performance incentives in public administration (Finan et al. 2017). Performance incentives generally strengthen upward accountability. Similar to the literature on local institutions, some research documents strong improvements in outcomes under performance incentives (Ashraf et al. 2014; Duflo et al. 2012), while other research finds only small, temporary effects (Celhay et al. 2018; Rasul and Rogger 2018; Olken et al. 2014). However, performance incentives can also have negative effects. For example, Khan et al. (2015) finds that performance incentives increase bribes to tax-collectors as it strengthens their bargaining power over taxpayers, and Dhaliwal and Hanna (2017) indicate that monitoring lowers job satisfaction and leads to evasion. I contribute to this literature in a similar way. My results show that leaders implement central government targets only when incentivized, but they also indicate that very strong incentives may fuel perverse measures, in this case severe restrictions of personal freedom and rights. By jointly studying accountability and local institutions, I connect the literature on performance incentives with that on local institutions. Both of these strands attempt to solve the same problem of accountability in delegated tasks. My findings indicate that local institutions and performance incentives are complements rather than substitutes.

Finally, the paper relates to research on the "dark side" of local institutions.[4] It is most closely related to Bonnier et al. (2019) who study the same institution, *Umuganda* meetings in Rwanda, in a different context. They find that during the time leading up to the 1994 genocide meetings were used by the old government for propaganda and to agitate people, resulting in larger civic participation in violence. Similarly, Satyanath et al. (2017) show that social clubs in Germany after World War I have spurred recruitment into the Nazi party. I contribute to this literature by providing evidence on involuntary behavior change through *Umuganda* meetings. In addition to negative consequences for outsiders, local institutions can be detrimental for the people in them. This finding speaks to Acemoglu et al. (2014) who suggest that local institutions are used by leaders for social control of civil society in Sierra Leone and relates to research on social sanctions (La Ferrara 2003; Karlan 2007; Miguel and Gugerty 2005).

The rest of the paper is structured as follows. Section 2 provides background information on *Umuganda* as a local institution, on performance contracts, and on development and popular support of targets. Section 3 describes the data and its construction. Section 4 explains and discusses the empirical strategy. Section 5 presents the main results, and section 6 studies the mechanism of *Umuganda*. Finally, section 7 concludes.

4. I use "dark side" as an attribute for local institutions because participation in them is widely considered to represent social capital. This attribute was originally proposed by Putnam (2000) for social capital.

2 Background

2.1 *Umuganda*

Umuganda is a traditional, local institution in Rwanda that dates back to pre-colonial times.[5] Having been used for political mobilization before the 1994 genocide (Bonnier et al. 2019), the new government of Rwanda suspended it thereafter. In 1998, however, this government reintroduced *Umuganda* nationwide to create socio-economic development (MINALOC 2011; RGB 2020). Since then, *Umuganda* was formalized in three stages. In November 2001, it was integrated into the government's Community Development Policy. In June 2005, its organization was harmonized by the National Umuganda Policy (MINALOC 2008). And finally, on November 17, 2007, *Umuganda* became a law (Organic Law N° 53/2007). The purpose of these policies was to embed the existing practice of *Umuganda* as a tool for policy-making into the public administrative structure. For my analysis, it is only important that no policy changed *Umuganda* in 2006.

Umuganda is and has most of the time in the past been a mandatory community program for all Rwandan adults on Saturdays. It is organized by a committee of village chiefs and consists of outdoor, physical labor (e.g. clearing bushes or cleaning roads) followed by a meeting (Uwimbabazi 2012). The local leaders typically announce *Umuganda* on the same day through word of mouth and loudspeakers mounted on cars (RGB 2014). During *Umuganda*, all shops must close and public transport stops. To enforce participation, local leaders have the discretion to fine absence by up to 5,000 Rwandan Francs, roughly 9 USD in 2007 and corresponding to half the monthly median wage (MINALOC 2007). Evidence suggests that many Rwandans participate involuntarily at *Umuganda* (Mukarubuga 2004; Uwimbabazi 2012; Purdeková 2011).

I argue that the meetings after physical labor during *Umuganda* affect behavior change. Local leaders use these meetings to mobilize, sensitize and support the population to collectively define and resolve their economic and social problems (MINALOC 2008). In practice, meetings amount to local leaders communicating top-down information about government programs and policies (Uwimbabazi 2012). They are also officially acknowledged by the government as a tool to implement development targets from performance contracts (RGB 2014). With respect to my outcomes, annual reports of Rwanda's Ministry of Health document that both family planning and mosquito bed nets were regularly discussed and promoted at *Umuganda* meetings (MoH 2009, 2008).

Some dissent exists about the frequence of *Umuganda* after the genocide. Recent government documents and research suggest different numbers of Saturdays with *Umuganda* in

5. Similar local institutions are common in countries of the African Great Lakes Region, notably Burundi, Ethiopia and South Sudan, and have also have been proposed for other countries. In addition, mandatory community programs also existed in many Soviet countries. In Russia, an equivalent institution was called *Subbotnik*, derived from the word 'subbota' meaning 'Saturday'.

a month. Organic Law N° 53/2007 states that it takes place only on the last Saturday (MINALOC 2007), and most official documents follow this representation. In contrast, research suggests that *Umuganda* is, in fact, held every week, as it was before the genocide (Purdeková 2011; Uwimbabazi 2012; NAR and Interpeace 2016). This frequency of *Umuganda* is also reported in the 2008 revised Community Development Policy, which is one of the few deviating government documents (MINALOC 2008). In my main specification, I exploit variation from rainfall on all Saturdays in a month. However, I also isolate and explore the effects of specific Saturdays, such as the last Saturday of every month.

2.2 Performance Contracts

In 2006, Rwanda's president, Paul Kagame, introduced Performance Contracts (*Imihigo*) in the public sector and beyond. On April 4, he signed the first contracts with all 30 district executives (the Mayors) to retie local government to central authority after decentralization. Immediately after, targets were passed down through cascading contracts to all levels of public administration and even further to the individual household (MINALOC 2010; Purdeková 2011). Contrasting the official portrayal as arrangements that reflect local priorities, performance contracts set targets top-down. In the first year after their introduction, three quarters of districts' targets were national policies and programs (GoR 2008), 71% were quantifiable and a majority was set at 100% (OSSREA 2007). Similar evidence exists at the household level. For example, in the government's 2010 Citizen Report Card survey, 78% of respondents state that they have not participated in formulating their own targets (Munyandamutsa 2011).

Two features of Rwandan performance contracts make them particularly effective for fast-track policy implementation. First, comparable units are regularly ranked against each other. And second, contracts set very strong social and material incentives for relative performance.[6] The consequence is a rat race in which leaders try to outperform one another. Local leaders implement village targets by letting household heads vow contributions in front of the community during *Umuganda* meetings (Bugingo and Interayamahanga 2010). Pledges are then recorded in a household's *Imihigo Booklet* and stamped upon completion. Stamped booklets serves as proof of "good standing" and are necessary to access certain government services like registering a marriage or birth (Sommers 2012; Uwimbabazi 2012). In addition, there are reports of fines, destruction of property and corporal punishment for refusing to contribute (Thomson 2008; Huggins 2009). Along this line, quantitative evidence suggests a large degree of compulsion in the implementation of performance targets at the local level (OSSREA 2007).

6. Within administration, rewards are commonly financial bonuses and promotions, whereas sanctions consist of removals from office and public shaming (Murray-Zmijewski and Gasana 2010). At district level, for example, approximately 75% of Mayors were removed from office between 2007 and 2009 due to poor performance (Scher and MacAulay 2010).

2.3 Development Targets and Popular Support

Rwanda's government targeted modern contraception and mosquitio bed nets as key performance indicators for development. In April 2004, two years before the introduction of performance contracts, it set ambitious goals in both. The aim was to raise modern contraceptive prevalence among women of reproductive age from 4% to 20% until 2010 and the percentage of children sleeping under bed nets from a baseline of 18% to 70% (MoH 2004). Rwanda reached both of these targets. Between 2005 and 2010, modern contraceptive prevalence among women 15 to 49 years old increased from 5.6% to 25.2%, and the percentage of households with at least one mosquito bed net increased from 18% to 83%. During the same time fertility dropped from 6 to 4 children per woman (NISR et al. 2012; NISR and Macro 2006).[7] Rwandan policy-makers attribute this fast-track development to performance contracts (Scher and MacAulay 2010).

Since the beginning, new users of modern contraception and households owning mosquito bed nets were explicit targets in performance contracts.[8] However, evidence suggests that many Rwandans did not approve of these behaviors. With respect to modern contraception, strong pro-natalist social norms prevailed after the 1994 genocide, and women using contraception were stigmatized as prostitutes (Kraehnert et al. 2019; Berry 2015; USAID and MoH 2002; Farmer et al. 2015). In addition, access does not seem to have been holding back development. The 2005 Rwandan Demography and Health Survey, for example, documents that only 3% of women reported knowledge, access or cost as a reason for not using modern contraception, in contrast to 33% quoting pro-natalist and opposing attitudes (NISR and Macro 2006). In fact, the government's own assessment in March 2006, one month before performance contracts, was that 'Up to now there have been very few achievements in part due to a lack of advocacy at all levels of Government and civil society' (MoH 2006, p.16). I argue that performance contracts ensured this advocacy from leaders.

With respect to mosquito bed nets, the claim of low popular support is based on the fact that Rwanda is a high altitude country. Its lowest point is already at 995 meters above sea level and much of the population lives at altitudes where the risk of contracting malaria should be very close to 0% because mosquitoes cannot survive (Bodker et al. 2003). I will explore heterogeneity in altitude as a proxy for popular support in the analysis. One potential reason why Rwanda's government implemented mosquito bed nets, despite low demand in many parts of the country, may be the fact that it received large financial support for this task from international aid donors through the U.S. President's Malaria Initiative and the Roll Back Malaria Partnership.

7. Rwanda's fertility transition between 2005 and 2010 is one of the fastest in history and comparable to that in China. Similar to China, also Rwanda considered to limit the number of children per family by law, but legislation was never passed (News 2007). For more information on the Chinese transition and how it was achieved see Zhang (2017).

8. OSSREA (2007) compiles district level targets and Sommers (2012, Appendix) presents a village leader's contract for the first year of performance contracts (2006-07). Further information on district targets between 2009 and 2013 can be found in RGB (2014).

3 Data

In this section, I describe the panel data on outcomes and rainfall. For the analysis, this data is matched through GPS coordinates and time. The panel data on outcomes of behavior change are constructed from retrospectively collected information, using dates and times that have been reported in cross-sectional Demography and Health Surveys (DHS) from Rwanda.

3.1 Family Planning

I use information from the 2010 Rwandan DHS to study women's adoptions of modern contraceptive methods and conceptions, i.e. when they become pregnant. The 2010 DHS interviewed 13,413 women who were between 15 and 49 years old and usual residents in 492 different communities. An integral part of each woman's questionnaire was a monthly calendar stretching from January 2005 to the date of the interview. In this calendar, interviewers recorded times of pregnancy and modern contraceptive use through retrospection. To ensure accuracy of the information, interviewers were required to ask a set of different questions in a recursive routine for each entry.

I construct my panel data on family planning outcomes from this retrospective calendar data. First, I build a panel data set indicating whether a woman is pregnant, using modern contraception or neither. Second, I define the two outcomes of behavior change in family planning, contraceptive adoption and conception. These outcomes are indicators that take the value 100 for behavior change in a given month and 0 otherwise. This coding produces estimates in percentage points later in the analysis. Third, following DHS sampling rules, I set all information to missing for times when a woman is below age 15.

In an interview, months of pregnancy are the first entries recorded in the calendar. Hence, I first explain the coding of conceptions before that of contraceptive adoptions. To record a pregnancy spell in the calendar, interviewers mark the monthly date of birth or termination and write back the status of pregnancy until the sum of marked months equals the number of completed months reported by the woman. As a consequence, pregnancies that end, for example, with birth generally consist of spells of 9 months in the calendar. However, these entries and also those of pregnancies that end with termination are inaccurate in the date of conception by two month and for two reasons.

First, recording the month of birth or termination as a full month of pregnancy implicitly assumes that births or terminations occur on the last day of the month. This assumption is extremely unlikely and passes through to the start of pregnancy. As a consequence, almost all women should already be pregnant one month earlier, i.e. in the month before the start of pregnancy in the calendar. Second, recording only completed months misses month 0 of a pregnancy, which is the month of conception. Around the world, pregnancy is counted to last 40 to 41 weeks, starting on the first day of the last menstrual period. This duration translates to 10

months with the fertilization occurring within 1-3 weeks after starting to count. In this respect, conception as the time of sexual intercourse and decision-making actually takes place another month before the start of pregnancy as recorded in the calendar. Based on these two reasons, I code conception as an indicator that takes the value 100 two months before the start of every pregnancy in the calendar and 0 otherwise. For pregnancies that end in birth, this definition is largely identical to lagging an analogue date-of-birth-indicator by 10 months.

The coding of contraceptive adoption of a modern method is straightforward. Contraceptive adoption is an indicator that takes the value 100 on each start date of modern contraceptive use in the calendar and is 0 otherwise. For the definition of modern contraceptive methods, I follow the standard DHS classification.

For the main analysis, I split the data along the time dimension and restrict the two panel lengths. First, I use April 2006, which is the introduction date of performance contracts, to split the data into a Before and an After panel data set. The reason for this measure is that it simplifies the regression equation, used later to separately estimate the before and after effects of *Umuganda* meetings. Second, I restrict the two panel lengths to 12 months before and 12 months after the introduction of performance contracts. The purpose of this measure is to narrow the time window around the introduction of performance contracts, which supports attribution of a change in effects to performance contracts.

Figure 1 presents the number of contraceptive adoptions and conceptions on each monthly date from February 2005 until July 2010. The solid black, vertical line marks the introduction of performance contracts at the beginning of April 2006. The grey shaded areas left and right of that line mark the lengths of the two panel data sets. No suspiciously high concentrations on certain dates can be observed for any of the two outcomes, suggesting that the calendar data is indeed accurate.[9] Table 1 presents summary statistics of the outcomes for the before and after panel data. Only women with at least two observations are kept in each data set because the inclusion of women fixed effects in my regressions drops women with only one observation.

3.2 Mosquito Bed Nets

I use information from two Rwandan DHS to study households' acquisitions of mosquito bed nets before and after the introduction of performance contracts. The 2005 DHS interviewed 10,146 households living in 456 communities with available GPS coordinates. The 2007-08 DHS interviewed 7,287 households in 246 geo-coded communities.[10] Both surveys collect information about mosquito bed nets in households. For each bed net acquired in the past three years, the data records the number of months before the interview when a household obtained

9. This observation is also confirmed when plotting the number of contraceptive adoptions and conceptions over months before the interview (see Figure A1 in the Appendix).

10. The analysis samples exclude 126 households from 6 communities without GPS coordinates in DHS 2005 and 90 households from 3 communities without GPS coordinates in DHS 2007-08.

the net. In addition, information about the source of the bed net, i.e. from where it was obtained, is available for nets acquired within the previous six months before the interview.

Figure 2 presents the raw data of the total number of mosquito bed nets that were acquired in each month before the interview. The figure shows high numbers of acquisitions on months 12, 18 and 24 in both surveys, which indicates that reporting precision deteriorates for bed nets acquired 12 months and more before the interview. Without a routine of questions to ensure data accuracy, similar to that used for 2010 DHS calendar entries, the concentrations likely are due to rounding and imprecise recall. For this reason, I restrict my analysis to bed nets acquired in months 0-11 before the interview.

Based on the raw data, I construct for each DHS a separate, retrospective, household level panel data set spanning 0-11 months before the interview. The 2005 DHS provides data before the introduction of performance contracts and the 2007-08 DHS provides data for the time thereafter. My main outcome is an indicator that takes the value 100 if a household acquired one or more mosquito bed nets in a given month before the interview and 0 otherwise. To later explore heterogeneity in the source of bed nets, I create two additional outcome indicators of panel length 0-6 months for the 2007-08 data. The first indicator takes the value 100 if at least one bed net in a month before the interview was acquired at a Health Center and is 0 otherwise. The second indicator takes the value 100 if at least one bed net in a month was acquired from Other Sources (e.g. a shop or market) and is 0 otherwise. Table 2 presents summary statistics of the panel data on acquisitions of mosquito bed nets.

3.3 Rainfall

I construct my rainfall measures from CMORPH rainfall estimates of the US National Oceanic and Atmospheric Administration's Climate Prediction Center. This data starts in 1998 and has two advantages. First, it has very high spatial and temporal resolution that captures the rainfall variations of Rwanda's many different micro-climates. A tile (data point) in the gridded map has a side length of approximately 8 km (0.073 degrees) and measures rainfall for a 30 minutes interval (Joyce et al. 2004). This resolution facilitates the confinement of rainfall to local communities and single days, with Saturday being the day of *Umuganda* meetings. Second, validation studies suggest that CMORPH rainfall estimates are particularly precise over complex terrain like Rwanda due to the morphing of satellite images and the exploitation of both infrared and microwave electromagnetic radiation (see e.g. Abera et al. 2016). While there will always be some measurement error in satellite rainfall data, this error should work against my findings as long as it is uncorrelated with the outcomes.

I construct my rainfall measures in two steps. First, I aggregate the data to daily estimates and extract rainfall in each community based on its GPS coordinate. Second, I create rainfall measures for each weekday (Mondays, Tuesdays, Wednesdays, etc.) that count the number of

"rainy" days on that weekday in a month. A rainy day is defined as a day with rainfall above a certain threshold, and a month is either a calendar month or a month before the interview, depending on the time structure of the outcome data to be matched with. I use round number thresholds from 1 mm up to 10 mm rainfall. My preferred threshold choice is 3 mm rainfall, which I discuss in section 4.1. Measures with other thresholds are used to evaluate the robustness of my results to that choice.

For the analysis, rainfall data are matched to outcome data using community and month identifiers. Table 3 presents summary statistics of the number of rainy Saturdays in a calendar month for the 492 communities of the 2010 Rwandan DHS between April 2005 and March 2007. The statistics for rainfall on other weekdays and time periods are very similar.

4 Empirical Strategy

To identify the effect of *Umuganda* meetings on behavior change, I use variation of rainfall over time. Without data on meetings, I estimate the reduced-form relationship. Rainfall on Saturdays proxies for low participation and cancellation (zero participation). This identification strategy rests on two assumptions. First, Saturday rainfall affects participation at *Umuganda* within a community over time (first stage). Second, the reduced-form effect of Saturday rainfall on behavior change operates only through this channel (exclusion restriction).

4.1 Rainfall and *Umuganda*

Saturday rainfall should strongly affect *Umuganda* because meetings and physical labor happen outside under the open sky and are usually only communicated on the same day through word of mouth or loudspeakers mounted on cars driving around (RGB 2014). Hence, both the program and the spread of information about it (e.g. the meeting point and time) are likely inhibited by rainfall. Ideally, I would like to test this first-stage assumption, but data on meetings does not exist. Bonnier et al. (2019) face this same obstacle and make two efforts to substantiate an existing first stage.

First, Bonnier et al. (2019) collect anecdotal evidence in the form of government and media reports on low participation at and cancellations of different *Umuganda* meetings and other events due to rainfall. Second, they estimate an effect of Saturday rainfall on participation at community meetings in neighboring Burundi, using Afrobarometer data. Burundi shares with Rwanda the same colonial history and is comparable in many socio-economic characteristics. Most importantly, community service with a meeting, very similar to *Umuganda*, is also held every Saturday in Burundi. Its name there is *Ibikorwa rusangi*. Bonnier et al. (2019) find a statistically significant, negative relationship between self-reported frequency of attendance at these meetings and the number of rainy Saturdays in the year leading up to the interview. Both,

the anecdotal evidence from Rwanda and the quantitative evidence from Burundi, directly suggest the existence of a first stage, i.e. an effect of Saturday rainfall on participation at *Umuganda* in Rwanda.

Several other studies document and exploit a negative relationship between rainfall and participation at different events. An important example is Madestam et al. (2013). They are the first to use a single, binary rainfall indicator to instrument for participation in Tea Party rallies in the US.[11] Similar to Bonnier et al. (2019), I follow this approach because I expect threshold effects, meaning that *Umuganda* is over proportionally affected or fails due to small dips in participation. Uwimbabazi (2012, p.216), for example, suggests this type of effect when she writes that at *Umuganda* 'successful implementation of any policy can be affected by the absence of the full participation of those especially who should benefit from these policies'. Moreover, the expectation of threshold effects is supported by theory and evidence of collective decision-making and action (see e.g. Olken 2010; Dal Bó et al. 2010; Faillo et al. 2013).

As I cannot empirically determine the relationship between rainfall and participation at *Umuganda*, I choose the threshold that defines a rainy day based on established standards and reasoning. According to the American Meterological Society, rainfall above 2.5 mm is classified as "moderate" and above 7.5 mm as "heavy" rain (AMS 2012). Madestam et al. (2013) use the first mark and exploit both, moderate and heavy rain, by defining a rainy day to count more than 2.5 mm (0.1 inches) rainfall in their main specification. Bonnier et al. (2019) deviate from this practice and use only heavy rain above a threshold of 10 mm. They motivate this choice with their anecdotal evidence on low participation and cancellations due to rainfall. For these cases, they find daily rainfall to have ranged between 1 mm and 18 mm with a median of 8 mm. However, all events with rainfall of 6 mm and more, except one, are reported as canceled. Hence, *Umuganda* very likely experiences reduced participation already at lower thresholds.

Two simple arguments with respect to the size of the complier group also suggest a threshold that is lower than that used by Bonnier et al. (2019). First, the size of the complier group, i.e. the number of people that do not attend *Umuganda* due to rainfall, may actually be already large at low levels of rainfall because most people do not like to attend *Umuganda*. Hence, rainfall reduces the cost of remaining absent rather than increasing the cost of attending. Specifically, I expect rainfall to protect against sanctions for absence because it is a 'good reason' in the sense that it is verifiable and because also other people will be absent. Both conditions make enforcement difficult. Second, as stated above, small dips in participation can let *Umuganda* fail in terms of being effective for policy implementation. Consequently, a comparably small complier group may in fact be already sufficiently large to cause strong or even complete im-

11. Most other studies use continuous measures of rainfall. For example, Collins and Margo (2007) use rainfall in April 1968 to instrument for participation in riots in the US. A large set of studies use rainfall to instrument for voter turnout on election day (see e.g. Fujiwara et al. 2016; Gomez et al. 2007; Hansford and Gomez 2010; Lind 2019; Fraga and Hersh 2010).

pairment of *Umuganda* (e.g. because leaders or the community cannot make binding decisions). In my analysis, this claim should result in similar effect sizes when using different thresholds.

In my preferred specification, I use a threshold of 3 mm to define a rainy day because it is the closest integer number to the standard of 2.5 mm (0.1 inches). However, I show robustness of my results to thresholds between 2 mm and 10 mm rainfall.

4.2 Exclusion Restriction

My empirical strategy relies on two exclusion restrictions, one with respect to Saturday rainfall affecting outcomes only through *Umuganda* meetings and the other with respect to the introduction of performance contracts being the only change that altered the objectives of these meetings at the time.

The first restriction assumes that without *Umuganda* meetings rainfall on Saturdays does not affect behavior change. Two characteristics of my analysis limit the scope for a different channel other than *Umuganda*. First, any such channel would need to be time varying due to the inclusion of two-way fixed effects in all my regressions. Second, it would need to be specific to Saturday rainfall because rainfall regressors for every single other day of the week control for general rainfall effects and serve as natural placebo tests. Under these two constraints, an effect of Saturday rainfall must be generated by a reoccurring event on that day. In my analysis, I address this issue and can rule out any channel related to accessing modern contraception (e.g. market days, but also distribution of methods during *Umuganda*), which leaves very little scope for a channel unrelated to the social interactions during *Umuganda* meetings.

The second restriction builds upon the first and assumes that without the introduction of performance contracts *Umuganda* meetings do not affect my outcomes of behavior change. With my panel data, I can attribute the change in effects of meetings to the time when performance contracts were introduced. Hence, some other nationwide policy or change would need to have altered the practice or objectives of *Umuganda* meetings with respect to my outcomes and coincided in timing with the introduction of performance contracts. I am not aware of any such change. It is certain, however, that both my outcomes of behavior change were targeted under performance contracts and that *Umuganda* was used to implement targets.

4.3 Specification

To estimate the effect of *Umuganda* meetings on behavior change through OLS, I run variations of the following reduced-form regression:

$$y_{it} = \sum_{d=1}^{7} \beta_d \, rain_{ctd} + \alpha_i + \tau_t + \varepsilon_{it}. \tag{1}$$

y_{it} is a binary indicator of behavior change of the observational unit i during month t. The unit of observation, i, depends on the outcome and is either an individual woman or household. Similarly, t may be either a monthly date or a month before the interview, depending on the panel structure of the outcome. $rain_{ctd}$ is the number of days with rainfall above a specific threshold on weekday d in observational unit i's community c during month t. Hence, the regression includes seven rainfall variables that count the number of rainy Mondays, Tuesday, Wednesday etc. in each community and month. In my preferred specification, a rainy day is defined by rainfall above 3 mm. α_i and τ_t are observational unit and monthly time fixed effects. In all estimations, I cluster standard errors at the community level because the community (or village) is the entity of *Umuganda* meetings and local leaders' performance targets. This allows the error term, ε_{it}, to be correlated both within communities and over time.

The coefficients, β_d, capture percentage point changes in the probability of behavior change in any given month following from an additional rainy day on the different weekdays in the same month. The interpretation as percentage point changes follows from the dependent variable, the indicator of behavior change, taking values of either 0 (no change) or 100 (change). Most interesting is the coefficient on Saturday rainfall, which can be interpreted as the effect of a failed *Umuganda* meeting. Rainfall on the other weekdays control for general rainfall effects and are placebo tests. As I will show later, their inclusion in the regressions is unimportant for the results. With unit fixed effects, the coefficients β_d are identified from temporal variation in rainfall and behavior change.

To identify the effect of performance contracts on creating behavior change through *Umuganda* meetings, I estimate equation 1 with panel date before and after the introduction of performance contracts. Subsequently, I test whether the corresponding coefficient estimates from both regressions are statistically different. This test essentially evaluates the significance of the Differences-in-Differences.[12] I present estimates from separate regressions with before and after data to interpret each of the two point estimates on Saturday rainfall as the effect of a failed *Umuganda* meeting. Their difference, the Differences-in-Differences, are rather uninteresting. Only their statistical significance is relevant to show that performance contracts led to a change in the practice or objectives of meetings. For this reason, I directly present p-values of the Differences-in-Differences.

12. I conduct this test by including interactions of all regressors with an after-performance-contracts-dummy, $I(t >= Apr.2006)$, and estimating this expanded equation jointly with before and after data.

5 Results

5.1 Main Reduced-form Effects

Table 4 presents the relationship between the two outcomes, contraceptive adoption and bed net acquisition, and the total number of days with rainfall above 3 mm for each weekday in a month. Both outcomes were targeted by the government for development. The point estimate on Saturday rainfall can be interpreted as the effects of a failed *Umuganda* meeting, which is a meeting that is canceled or has too low attendance for effective decision-making.

Regression 1 uses a 12-months panel of women over the first year after the introduction of performance contracts, i.e. from April 2006 until March 2007. The reduced-form estimate on Saturday rainfall is strongly statistically significant at the 99% confidence level. It suggests that a failed *Umuganda* meeting reduces the probability that a woman adopts modern contraception in a given month by 0.071%. While this absolute effect seems small, the relative effect compared to the unconditional probability is -18%. Reassuringly, rainfall on any other weekday is insignificant.

Regression 2 estimates the same relationship for the year before the introduction of performance contracts, from April 2005 until March 2006. None of the coefficient estimates, including that on Saturday rainfall, is statistically significant at any conventional level. P-values for the differences in corresponding coefficient estimates between regressions 1 and 2 are presented one column to the right in Table 4. The difference in estimates on Saturday rainfall is statistically significant at the 99% confidence level. No other difference is statistically significant at the 90% level or below. This finding suggests that *Umuganda* meetings became an effective tool for the implementation of national family planning policy after the introduction of performance contracts.

Regressions 3 and 4 estimate an analogue relationships for the second government target, acquisitions of mosquito bed nets. The data structure in these two regressions is slightly different. Acquisitions are observed at the household level and the time dimension of the panel data are months-before-the-interview. Regression 3 uses 12-months panel data on bed net acquisitions 0 to 11 months before DHS 2007-08 interviews, which are data after the introduction of performance contracts. Regression 4 uses equivalent data from DHS 2005, before the introduction of performance contracts. In regression 3, the reduced-form estimate on Saturday rainfall is highly significant at the 99% confidence level (similar to regression 1). None of the other weekdays is statistically significant at 95% confidence level or high. The relative effect of the point estimate on Saturday rainfall is -10%.

In regression 4, rainfall on all weekdays, including Saturdays, is statistically insignificant at any conventional level (as in regression 2). The p-values for the differences in corresponding coefficient estimates from regressions 3 and 4 demonstrate that only the effect of Saturday rainfall

changed significantly at the 95% confidence level. Finding the same pattern of coefficient estimates for another targeted, but otherwise unrelated outcome suggests that in fact performance contracts are responsible for aligning *Umuganda* meetings with national policy. Over the observation period, the practice of *Umuganda* arguably did not change. However, performance contracts allowed the central government to suddenly set the agenda of meetings.

Magnitude The above coefficient estimates can be considered lower bound estimates of a failed *Umuganda* meeting. The estimates should be biased toward zero because the number of rainy Saturdays is an imprecise measure of the number of failed meetings in a month. Hence, relative effects of -18% and -10% suggest that meetings have a very strong effect on the timing of behavior change, especially when also considering that there may be up to 5 meetings in a month. The relative effect size is slightly larger compared to those found by Bonnier et al. (2019) who estimate the relationship between Saturday rainfall leading up to the Rwandan genocide and civilian participation rates in violence using cross-sectional data. For the period from October 1993 until March 1994, which is driving their results, they find that a rainy Saturday (defined by rainfall above 10 mm) reduced civilian participation by 10% compared to the unconditional mean.

5.2 Tracing the Effects

Table 5 traces the effects of Saturday rainfall on contraceptive adoption and bed net acquisition using the two 12-months panel data sets after the introduction of performance contracts. Regressions 1 and 3 show that only the coefficient estimates on the number of Saturdays with rainfall above 3 mm in the same month are highly significant at the 99% confidence level. All estimates on lagged Saturday rainfall can be considered placebo tests and are insignificant with one exception. The coefficient on the first lag in regression 1 is positive and statistically significant at the 90% confidence level. This finding may indicate that Saturday rainfall dams up policy implementation and causes a catching up in the following month.

Regressions 2 and 4 evaluate the relationship between the two outcomes and four binary indicators that respectively take the value 1 if rainfall on the first, second, third or last Saturday of a monthly date is above 3 mm and 0 otherwise. The effects of these specific Saturdays are similar to one another within the same regression. If at all, regression 2 suggests that central Saturdays of a monthly date are slightly more important for generating the overall effect of Saturday rainfall in a month, and regression 4 suggests that Saturdays toward the end of a monthly date may be more important. However, none of the differences of Saturday rainfall coefficients in the same regression is statistically significant and all estimates have the same sign. This finding rather supports the evidence discussed in section 2 that *Umuganda* was held on multiple Saturdays every month.

Finally, the long panel data of contraceptive adoption allows me to study the effect of Saturday rainfall over time. Figure 3 presents estimates from rolling regressions over the 7th month of a 12-months rolling window. The solid black line connects the coefficient estimates and the dashed curves mark 95% confidence intervals. It can be observed that Saturday rainfall effects become negative and statistically significant upon or shortly after the introduction of performance contracts (vertical line on April 2006). The effect continues to persist for roughly 1.5 years after the introduction of performance contracts and then seems to fade away. This finding may mechanically arise if the fraction of women who are both not using modern contraception and being affected by *Umuganda* meetings decreases over time, which is very likely the case.

5.3 Robustness Checks

In this section, I present additional robustness checks. Figure 4 shows robustness of the results after the introduction of performance contracts with respect to two dimensions. First, the effects of Saturday rainfall are robust to using different thresholds to define a rainy day. And second, they are largely unaffected by potential multicollinearity in the weekly rainfall variables. In Figure 4, each coefficient estimate (diamond) and 95% confidence interval (capped bar) is obtained from a separate regression of the outcome on the number of Saturdays with rainfall above a certain threshold, controlling only for unit of observation and time fixed effects, but not rainfall on other weekdays.

Panel (A) shows that the effect of Saturdays rainfall on contraceptive adoption is statistically significant when using thresholds between 2 mm and 10 mm rainfall. Panel (B) shows that also the effect on bed net acquisition is statistically significant for 9 out of 10 thresholds at the 95% confidence level. The coefficient estimates with a 3 mm threshold are very similar to the estimates in Table 4, suggesting that multicollinearity between the rainfall regressors does not affect the results. In addition, the similar effect sizes across the different definitions of a rainy day are consistent with threshold effects in participation at *Umuganda*, as proposed in section 4.1.

Figures A2 and A3 in the appendix present results from equivalent regressions of the two outcomes on rainfall on each of the other weekdays. In this battery of 120 regressions only one coefficient estimate is marginally statistically significant at the 95% confidence level, which can be expected to occur by chance from this large number of multiple hypothesis testing.

5.4 Alternative Channels

Strong evidence suggests that access to contraception and bed nets does not generate my results. With respect to contraception, two complementary pieces of information rule out this explanation. First, hormonal contraceptives (injections, pills, IUDs and implants), as the most commonly used class of methods in Rwanda, were only available at health facilities, i.e. health

centers and hospitals (USAID et al. 2011). And second, at the time of the analysis, health centers were closed and hospitals had high surcharges on weekends (Ueberschär 2018), preventing access on Saturdays.

My data supports this argument. In the 2010 DHS, 95% of hormonal method users (and 91% of any modern method users) report that their first source for the method was a health facility. Adoptions of these methods make up 90% of all adoptions and drive my results. Evidence that health facilities are closed on weekends comes from 52,539 vaccinations with their exact dates copied from children's health cards in the 2010 DHS. Only 3% of vaccinations took place on a Saturday or Sunday. Further support with respect to family planning is provided in the 2007 Rwandan DHS Service Provision Assessment. It documents (and these numbers are likely overreported) that most health facilities provided family planning services on five or less days a week (NISR et al. 2008), which most likely excludes weekends.

For mosquito bed nets, information on the source is available if the net is obtained up to 6 months before the interview. Table 6 presents results for the time after the introduction of performance contracts. The three outcomes are binary, monthly indicators that take the value 100 if a bed net was acquired from a specific source and 0 otherwise. Regression 1 estimates the relationship between rainfall and bed net acquisition from any source (as in Table 4) on this short panel. All results hold and are very similar. Regressions 2 and 3 only use acquisitions from health facilities and other sources, respectively. The results clearly show that acquisitions from other sources, mostly shops and pharmacies, are generating the effect of Saturday rainfall. Hence, access to health facilities cannot be the channel.

However, health facilities are the only source for the in Rwanda commonly used hormonal contraceptive methods. Consequently, it is reasonable to conclude that the same relationship pattern between the two targeted outcomes and Saturday rainfall must be generated by something else than access. I claim *Umuganda* meetings are generating the effects as they are known to regularly take place on Saturdays and the two analyzed behavior changes were explicitly discussed there.

6 Mechanism

Having documented a strong, robust effect of Saturday rainfall on two targeted behaviors, I now provide evidence indicating a mechanisms of *Umuganda* based on pressure. First, I explore conception as an evasive behavior that protects against pressure to adopt modern contraception. Second, I study spatial heterogeneity in the prevalence of mosquitoes as a proxy for popular support for bed nets.

6.1 Evasive Behavior

Enforcement creates evasion. With respect to contraception, conception can be considered an evasive behavior because pregnant women must not use it. If the mechanism of *Umuganda* is based on pressure then meetings should also increase conception, which is tantamount to a negative correlation between Saturday rainfall (failed meetings) and conception. However, the probability of conception depends mechanically and inversely on contraceptive use. With a negative effect of Saturday rainfall on contraceptive adoption, documented above, this connection predicts a positive correlation between Saturday rainfall and conception. Hence, we can infer the mechanism of *Umuganda* from the coefficient in a regression of conception on Saturday rainfall. A negative coefficient suggests a mechanism based on pressure, whereas a positive coefficient suggests a mechanism based on free choice.

Regressions 1 in Table 7 shows the relationship between conception and rainfall on different weekdays after the introduction of performance contracts. The coefficient estimate on Saturday rainfall is negative and statistically significant at the 95% confidence level. It suggests that one failed *Umuganda* meeting reduces the probability to become pregnant in the same month by 8%. This finding of a negative effect suggests that the mechanism of *Umuganda* is based on pressure. The statistical significance further indicates that a comparably large fraction of women chose conception as a behavior to evade contraceptive adoption. Otherwise, in the displayed average effect of Saturday rainfall on conception, the negative evasion effect would not outweigh the mechanical and positive effect from reduced contraceptive adoptions. Figure A4 in the appendix presents the effects of each weekday at different rainfall thresholds from separate regressions. Panel (A) shows that the effect of Saturday rainfall is robust to using thresholds between 2 mm and 9 mm.

Regression 1 also displays an effect of Wednesday rainfall that is statistically significant at the 95% confidence level. However, Panel (E) in Figure A4 shows that this effect is only statistically significant for rainfall thresholds 3 mm and below. As the effect does not persist for larger thresholds, I believe this finding is spurious. The other Panels in Figure A4 document that all effects of other weekdays at different rainfall thresholds are statistically insignificant.

Regression 2 in Table 7 shows the relationship between conception and rainfall on different weekdays before the introduction of performance contracts. None of the coefficient estimates is statistically significant at the 95% confidence level. The column to the right displays p-values of the differences in estimates between regression 1 and 2. Only the difference for Saturday rainfall is statistically significant at the 95% confidence level. This finding further corroborates my claim performance contracts turned *Umuganda* meetings into an effective tool to implement national development targets and that targets were achieved through pressure.

6.2 Heterogeneity in Popular Support

Regressions 3 and 4 in Table 7 explore heterogeneity in the effects of rainfall on bed net acquisition with respect to altitude. The underlying motivation is the well-known fact that the incidence of mosquitoes strongly decreases with altitude.[13] This variation in the objective usefulness of mosquito bed nets should be strongly correlated with popular support because people are reluctant to invest time and money for something they do not need. With lower levels of support, more pressure (enforcement) is needed to create behavior change. Hence, if the mechanism of *Umuganda* is based on pressure, we can expect to find larger effects in high altitude areas.

In the 2007-08 DHS, the median community is located at an altitude of 1,670 meters. At this altitude, the risk of contracting malaria should be close to 0%, and we can expect extremely low support for targets in mosquito bed nets. Regression 3 and 4 in Table 7 estimate the relationship between acquisition of bed nets and rainfall on different weekdays after the introduction of performance contracts respectively using only communities located above and below median altitude. Saturday rainfall is negative and statistically significant at the 95% confidence level in regression 3 and at the 90% level in regression 4. While the difference in coefficients is not statistically significant, the results show a larger point estimate and relative effect for high altitude communities. This finding is consistent with pressure as the mechanism of *Umuganda*. In addition, it shows that Rwanda's government increased the prevalence of mosquito bed nets across the country without accounting for whether they were actually needed or not.

7 Conclusion

This paper studies the role of community meetings and performance incentives in Rwanda to achieve development targets in health. Specifically, it investigates the effects of meetings on two changes in individual behavior that were unpopular among the population, but desired by the central government. The behavior changes are adoptions of modern contraceptive methods and acquisitions of mosquito bed nets. Identification comes from exogenous variation in meeting attendance over time induced by rainfall. This setup allows me to compare the effects before and after a reform that strongly increased upward accountability through performance contracts. I show that local leaders use community meetings on Saturdays to implement targets after the reform. Before the reform, these meetings have no effects. Similar effect patterns in the two unrelated but targeted behavior changes suggest that community meetings and performance incentives are complementary and form a governance system that can be used to implement a

13. In similar climate and terrain as in Rwanda, Bodker et al. (2003) study the incidence of mosquitoes in Tanzania using mosquito light traps. Their traps caught only 4 mosquitoes a year at altitude 1,700 meters above sea level, compared to 269 mosquitoes at altitude 1,000 meters and 3,282 mosquitoes at altitude 300 meters.

wide range of development goals. However, I find evidence that suggests that behavior change is involuntary.

These findings have two important implications. First, they challenge the commonly assumed downward accountability of local institutions when generating development. By showing fast-track development through a local institution under upward accountability, my findings warn that many successful community-based development programs may, in fact, have exploited upward accountability by helping leaders control behavior in their communities. Community-based projects need to be carefully designed in order not to thwart the actual objective of making development more democratic and inclusive. Second, my findings emphasize the importance of accounting for the institutional context in development projects. Performance incentives and community meetings are widely considered good policies on their own. However, their combination can have adverse consequences as suggested by the evidence of involuntary behavior change in Rwanda.

Finally, this paper provides a new perspective on Rwanda's top-performance in many of the Millenium Development Goals and an explanation for its recent fertility transition. Its lessons may be valuable when assessing current and past developments in other countries, especially when these are authoritarian and aim at exploiting social capital for policy-making as e.g. China.

References

Abera, Wuletawu, Luca Brocca, and Riccardo Rigon. 2016. "Comparative Evaluation of Different Satellite Rainfall Estimation Products and Bias Correction in the Upper Blue Nile (UBN) Basin." *Atmospheric research* 178:471–483.

Acemoglu, Daron, Tristan Reed, and James A Robinson. 2014. "Chiefs: Economic Development and Elite Control of Civil Society in Sierra Leone." *Journal of Political Economy* 122 (2): 319–368.

AMS. 2012. *Rain.* American Meteorological Society Glossary, April. Accessed March 6, 2020. http://glossary.ametsoc.org/wiki/Rain.

Anderson, Siwan, Patrick Francois, and Ashok Kotwal. 2015. "Clientelism in Indian Villages." *American Economic Review* 105 (6): 1780–1816.

Ashraf, Nava, Oriana Bandiera, and B Kelsey Jack. 2014. "No Margin, No Mission? A Field Experiment on Incentives for Public Service Delivery." *Journal of Public Economics* 120:1–17.

Banerjee, Abhijit V, Rukmini Banerji, Esther Duflo, Rachel Glennerster, and Stuti Khemani. 2010. "Pitfalls of Participatory Programs: Evidence from a Randomized Evaluation in Education in India." *American Economic Journal: Economic Policy* 2 (1): 1–30.

Berry, Marie E. 2015. "When "Bright Futures" Fade: Paradoxes of Women's Empowerment in Rwanda." *Signs: Journal of Women in Culture and Society* 41 (1): 1–27.

Björkman, Martina, and Jakob Svensson. 2009. "Power to the People: Evidence from a Randomized Field Experiment on Community-based Monitoring in Uganda." *The Quarterly Journal of Economics* 124 (2): 735–769.

Björkman Nyqvist, Martina, Damien de Walque, and Jakob Svensson. 2017. "Experimental Evidence on the Long-run Impact of Community-based Monitoring." *American Economic Journal: Applied Economics* 9 (1): 33–69.

Bodker, Rene, J Akida, D Shayo, W Kisinza, HA Msangeni, EM Pedersen, and SW Lindsay. 2003. "Relationship Between Altitude and Intensity of Malaria Transmission in the Usambara Mountains, Tanzania." *Journal of Medical Entomology* 40 (5): 706–717.

Bonnier, Evelina, Jonas Poulsen, Thorsten Rogall, and Miri Stryjan. 2019. "Preparing for Genocide: Quasi-Experimental Evidence from Rwanda."

Bugingo, Irenee, and Reverien Interayamahanga. 2010. *A Study On The Development And Use Of Governance Indicators in Rwanda.* Technical report. Kigali: Institute of Research and Dialogue for Peace (IRDP).

Casey, Katherine, Rachel Glennerster, and Edward Miguel. 2012. "Reshaping Institutions: Evidence on Aid Impacts Using a Preanalysis Plan." *The Quarterly Journal of Economics* 127 (4): 1755–1812.

Celhay, Pablo, Paul Gertler, Paula Giovagnoli, and Christel Vermeersch. 2018. "Long Run Effects of Temporary Incentives on Medical Care Productivity." *American Economic Journal: Applied Economics,* no. forthcoming.

Collins, William J, and Robert A Margo. 2007. "The Economic Aftermath of the 1960s Riots in American Cities: Evidence From Property Values." *The Journal of Economic History* 67 (4): 849–883.

Dal Bó, Pedro, Andrew Foster, and Louis Putterman. 2010. "Institutions and Behavior: Experimental Evidence on the Effects of Democracy." *American Economic Review* 100 (5): 2205–29.

Dhaliwal, Iqbal, and Rema Hanna. 2017. "The Devil Is in the Details: The Successes and Limitations of Bureaucratic Reform in India." *Journal of Development Economics* 124:1–21.

Duflo, Esther, Pascaline Dupas, and Michael Kremer. 2015. "School Governance, Teacher Incentives, and Pupil–teacher Ratios: Experimental Evidence from Kenyan Primary Schools." Publisher: Elsevier, *Journal of Public Economics* 123:92–110.

Duflo, Esther, Rema Hanna, and Stephen P Ryan. 2012. "Incentives Work: Getting Teachers to Come to School." *American Economic Review* 102 (4): 1241–78.

Enikolopov, Ruben, and Ekaterina Zhuravskaya. 2007. "Decentralization and Political Institutions." *Journal of Public Economics* 91 (11-12): 2261–2290.

Faillo, Marco, Daniela Grieco, and Luca Zarri. 2013. "Legitimate Punishment, Feedback, and the Enforcement of Cooperation." *Games and Economic Behavior* 77 (1): 271–283.

Farmer, Didi Bertrand, Leslie Berman, Grace Ryan, Lameck Habumugisha, Paulin Basinga, Cameron Nutt, Francois Kamali, et al. 2015. "Motivations and Constraints to Family Planning: A Qualitative Study in Rwanda's Southern Kayonza District." *Global Health: Science and Practice* 3 (2): 242–254.

Finan, Frederico, Benjamin A Olken, and Rohini Pande. 2017. "The Personnel Economics of the Developing State." In *Handbook of Economic Field Experiments,* 2:467–514. Elsevier.

Fraga, Bernard, and Eitan Hersh. 2010. "Voting Costs and Voter Turnout in Competitive Elections." *Quarterly Journal of Political Science* 5:339–356.

Fujiwara, Thomas, Kyle Meng, and Tom Vogl. 2016. "Habit Formation in Voting: Evidence from Rainy Elections." *American Economic Journal: Applied Economics* 8 (4): 160–188.

Gomez, Brad T, Thomas G Hansford, and George A Krause. 2007. "The Republicans Should Pray for Rain: Weather, Turnout, and Voting in US Presidential Elections." *Journal of Politics* 69 (3): 649–663.

GoR. 2008. *Rwanda Joint Governance Assessment Report.* Technical report. Kigali: Government of Rwanda, October. Accessed March 20, 2020. https://www.afdb.org/fileadmin/uploads/afdb/Documents/Project-and-Operations/ADF-BD-IF-2008-220-EN-RWANDA-JOINT-GOVERNANCE-ASSESSMENT-JGA.PDF.

Hansford, Thomas G, and Brad T Gomez. 2010. "Estimating the Electoral Effects of Voter Turnout." *American Political Science Review* 104 (2): 268–288.

Huggins, Chris. 2009. "Agricultural Policies and Local Grievances in Rural Rwanda." *Peace Review* 21 (3): 296–303.

Joyce, Robert J, John E Janowiak, Phillip A Arkin, and Pingping Xie. 2004. "CMORPH: A Method that Produces Global Precipitation Estimates from Passive Microwave and Infrared Data at High Spatial and Temporal Resolution." *Journal of Hydrometeorology* 5 (3): 487–503.

Karlan, Dean S. 2007. "Social Connections and Group Banking." *The Economic Journal* 117 (517): F52–F84.

Khan, Adnan Q, Asim I Khwaja, and Benjamin A Olken. 2015. "Tax farming redux: Experimental evidence on performance pay for tax collectors." *The Quarterly Journal of Economics* 131 (1): 219–271.

Kraehnert, Kati, Tilman Brück, Michele Di Maio, and Roberto Nisticò. 2019. "The Effects of Conflict on Fertility: Evidence from the Genocide in Rwanda." *Demography:* 1–34.

La Ferrara, Eliana. 2003. "Kin Groups and Reciprocity: A Model of Credit Transactions in Ghana." *American Economic Review* 93 (5): 1730–1751.

Lind, Jo Thori. 2019. "Spurious Weather Effects." *Journal of Regional Science* 59 (2): 322–354.

Madestam, Andreas, Daniel Shoag, Stan Veuger, and David Yanagizawa-Drott. 2013. "Do Political Protests Matter? Evidence from the Tea Party Movement." *The Quarterly Journal of Economics* 128 (4): 1633–1685.

Mansuri, Ghazala, and Vijayendra Rao. 2013. *Localizing Development: Does Participation Work?* Policy Research Report. Washington, DC: World Bank. Accessed March 25, 2020. https://openknowledge.worldbank.org/handle/10986/11859.

Miguel, Edward, and Mary Kay Gugerty. 2005. "Ethnic Diversity, Social Sanctions, and Public Goods in Kenya." *Journal of Public Economics* 89 (11): 2325–2368.

MINALOC. 2007. *Law Establishing Community Works in Rwanda* [in English], November. Accessed September 26, 2017. http://lip.alfa-xp.com/lip/AmategekoDB.aspx?Mode=r&pid=8729&iid=2289&rid=30700136.

———. 2008. *Community Development Policy*. Revised version. Ministry of Local Government, Rwanda, April. Accessed September 17, 2019. https://www.ilo.org/dyn/natlex/docs/ELECTRONIC/98538/117303/F187363017/RWA-98538.pdf.

———. 2010. *Concept Paper on Imihigo Planning and Evaluation*. Technical report. Kigali: Ministry of Local Government, Rwanda, February. Accessed September 27, 2017. https://www.minaloc.gov.rw/IMG/pdf_Imihigo_Concept_Note_22_feb_2010_1_.pdf.

———. 2011. *Concept Note on Umuganda Current Status*. Technical report. Ministry of Local Government, Rwanda. Accessed October 2, 2017. http://www.rwandapedia.rw/cmis/views/workspace%253A%252F%252FSpacesStore%252Ff8d00bc6-4b7b-4d92-bf64-c02373e61f25.

MoH. 2004. *Health Sector Strategic Plan 2005-2009*. Technical report. Kigali: Ministry of Health, Rwanda, April. Accessed April 23, 2020. http://www.healthyfutures.eu/images/healthy/deliverables/D5.1/Rwanda/rwanda%20health%20sector%20strategic%20plan%202005%20-%202009%202004.pdf.

———. 2006. *National Family Planning Policy and Its Five-Year Strategies (2006-2010)*. Technical report. Ministry of Health, Rwanda, March. Accessed March 16, 2020. https://www.gfmer.ch/SRH-Course-2010/national-guidelines/pdf/National-FP-policy-MOH-Rwanda-2006.pdf.

———. 2008. *Annual Report 2007*. Technical report. Ministry of Health, Rwanda. Accessed May 20, 2016. http://www.moh.gov.rw/fileadmin/templates/MOH-Reports/RAPPORT-ANNUEL-2007.pdf.

———. 2009. *Annual Report 2008*. Technical report. Ministry of Health, Rwanda, April. Accessed May 20, 2016. http://www.moh.gov.rw/fileadmin/templates/MOH-Reports/Final-MoH-annual-report-2008.pdf.

Mukarubuga, Cecile. 2004. *Social Movements in Rwanda: Experience of Social Forums Against Poverty*. Technical report. Agency for Co-operation and Research in Development.

Munyandamutsa, Jean Paul. 2011. *2010 Citizen Report Card Survey: Major Findings*. Kigali, December. Accessed March 23, 2020. https://opendocs.ids.ac.uk/opendocs/bitstream/handle/20.500.12413/10076/2010%20Citizen%20report%20card%20survey.pdf?sequence=1.

Murray-Zmijewski, Alexandra, and Charles Gasana. 2010. *Monitoring and Evaluation in Rwanda Country Case Study,* October.

NAR and Interpeace. 2016. *Governing With and For Citizens.* Technical report. Never Again Rwanda, Interpeace, August.

News, BBC. 2007. *Rwanda Moves to Limit Family Size,* February. Accessed May 13, 2020. http://news.bbc.co.uk/2/hi/africa/6358381.stm.

NISR and ORC Macro. 2006. *Rwanda Demographic and Health Survey 2005.* Technical report. Calverton, Maryland, U.S.A.: National Institute of Statistics of Rwanda (NISR) and ORC Macro.

NISR, MoH, and Macro International Inc. 2008. *Rwanda Service Provision Assessment Survey 2007.* Technical report. Calverton, Maryland, U.S.A.: National Institute of Statistics of Rwanda (NISR), Ministry of Health of Rwanda (MoH), and Macro International Inc.

NISR, MoH, and ICF International. 2012. *Rwanda Demographic and Health Survey 2010.* Technical report. Calverton, Maryland, USA: National Institute of Statistics of Rwanda (NISR), Ministry of Health of Rwanda (MoH), and ICF International.

Olken, Benjamin A. 2007. "Monitoring Corruption: Evidence from a Field Experiment in Indonesia." *Journal of Political Economy* 115 (2): 200–249.

———. 2010. "Direct Democracy and Local Public Goods: Evidence from a Field Experiment in Indonesia." *American Political Science Review* 104 (2): 243–267.

Olken, Benjamin A, Junko Onishi, and Susan Wong. 2014. "Should Aid Reward Performance? Evidence From a Field Experiment on Health and Education in Indonesia." *American Economic Journal: Applied Economics* 6 (4): 1–34.

OSSREA. 2007. *Rapid and Extensive Assessment of Performance Management Contracts - Imihigo.* Technical report. Organisation for Social Science Research in Eastern and Southern Africa, July.

Pradhan, Menno, Daniel Suryadarma, Amanda Beatty, Maisy Wong, Arya Gaduh, Armida Alisjahbana, and Rima Prama Artha. 2014. "Improving Educational Quality through Enhancing Community Participation: Results from a Randomized Field Experiment in Indonesia." *American Economic Journal: Applied Economics* 6 (2): 105.

Purdeková, Andrea. 2011. "'Even if I am not here, there are so many eyes': Surveillance and State Reach in Rwanda." *The Journal of Modern African Studies* 49 (3): 475–497.

Putnam, Robert D. 2000. *Bowling Alone: The Collapse and Revival of American Community.* New York: Simon & Schuster.

Rasul, Imran, and Daniel Rogger. 2018. "Management of Bureaucrats and Public Service Delivery: Evidence from the Nigerian Civil Service." *The Economic Journal* 128 (608): 413–446.

Reinikka, Ritva, and Jakob Svensson. 2004. "Local Capture: Evidence from a Central Government Transfer Program in Uganda." *The Quarterly Journal of Economics* 119 (2): 679–705.

RGB. 2014. *The Assessment of the Impact of the Home Grown Initiatives.* Special Issue 4. Kigali: Rwanda Governance Board.

———. 2020. *Umuganda,* March. http://www.rgb.rw/index.php?id=37.

Satyanath, Shanker, Nico Voigtländer, and Hans-Joachim Voth. 2017. "Bowling for Fascism: Social Capital and the Rise of the Nazi Party." *Journal of Political Economy* 125 (2): 478–526.

Scher, Daniel, and Christine MacAulay. 2010. "The Promise of Imihigo: Decentralized Service Delivery in Rwanda, 2006-2010." *Innovations for Successful Societies: Policy Note* 133.

Sommers, Marc. 2012. *Stuck: Rwandan Youth and the Struggle for Adulthood.* Athens, Georgia: University of Georgia Press.

Thomson, Susan M. 2008. "Rwanda." In *Africa Yearbook Online,* edited by Andreas Mehler and Klaas Van Walraven, vol. 5.

Ueberschär, Nicole. 2018. *Access to and Utilization of Health Care: Spatial Disparities of Health Center Utilization in Huye District (Rwanda).* 1st ed. Vol. 3. Geographische Gesundheitsforschung. Düren, Germany: Shaker Verlag.

UNDP. 2014. *Millenium Development Goals Rwanda: Final Progress Report 2013.* Technical report. Kigali: United Nations Development Program, December. Accessed April 8, 2020. https://www.undp.org/content/dam/rwanda/docs/Research%20and%20publications/Millenium%20development%20goals/UNDP_RW_MDGR%20Rwanda_31_03_2015.pdf.

USAID, MCHIP, and TRP. 2011. *Assessment of Rwanda's National Family Planning Policy And Its Five-Year Strategies (2005-2010).* Technical report. USAID, Maternal and Child Health Integrated Program, The Respond Project, June. Accessed February 14, 2020. http://www.respond-project.org/archive/files/4/4.3/4.3.2/Report-2011-Rwanda-Assessment.pdf.

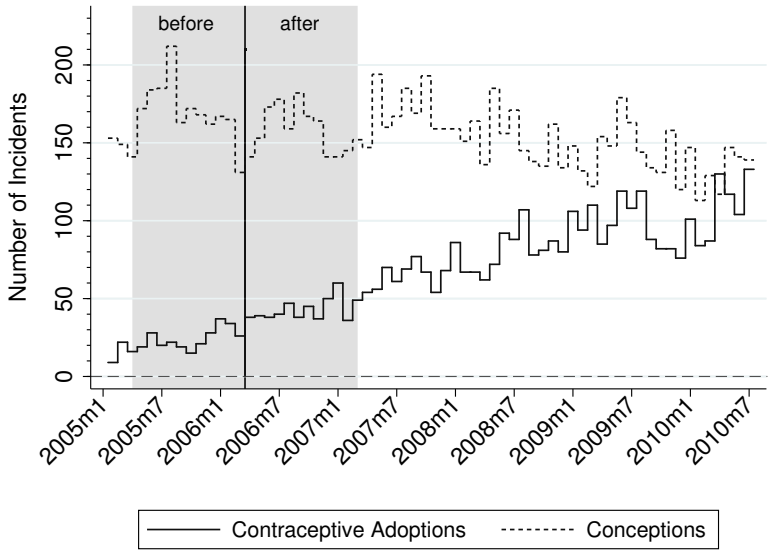
USAID and MoH. 2002. *Report of a Qualitative Assessment of Family Planning in Rwanda.* Technical report. USAID and Ministry of Health of Rwanda.

Uwimbabazi, Penine. 2012. “An Analysis of Umuganda: the Policy and Practice of Community Work in Rwanda.” PhD diss., University of KwaZulu-Natal.

Zhang, Junsen. 2017. “The Evolution of China’s One-Child Policy and its Effects on Family Outcomes.” *The Journal of Economic Perspectives* 31 (1): 141–159.

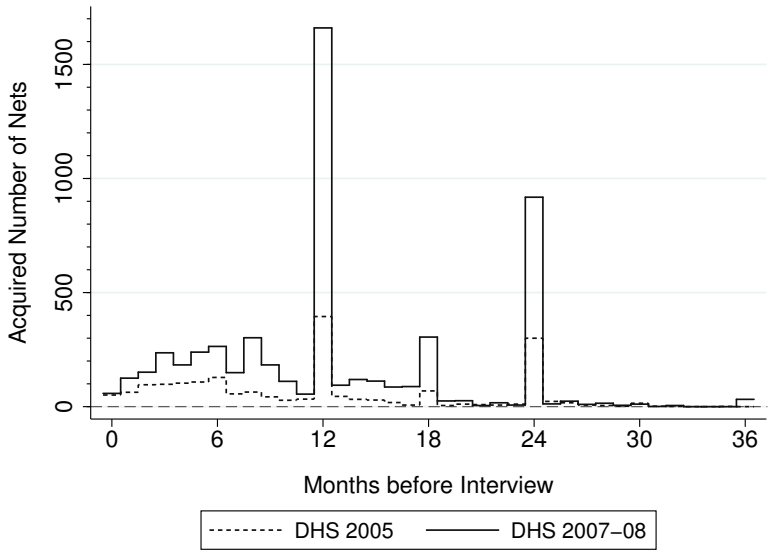
Figures and Tables

Figure 1: Number of Contraceptive Adoptions and Conceptions over Time



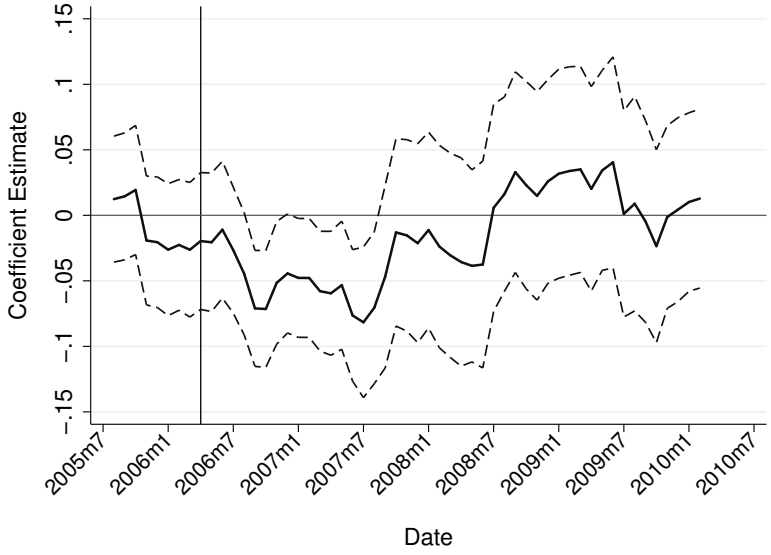
Notes: Based on 13,413 women between 15 and 49 years old and who are usual residents of interviewed households in the 2010 Rwandan DHS. The solid, vertical line marks the introduction of performance contracts in April 2006.

Figure 2: Number of Mosquito Bed Nets Acquired in Different Months



Notes: Based on 10,146 and 7,287 households with GPS coordinates in DHS 2005 and DHS 2007-08 data.

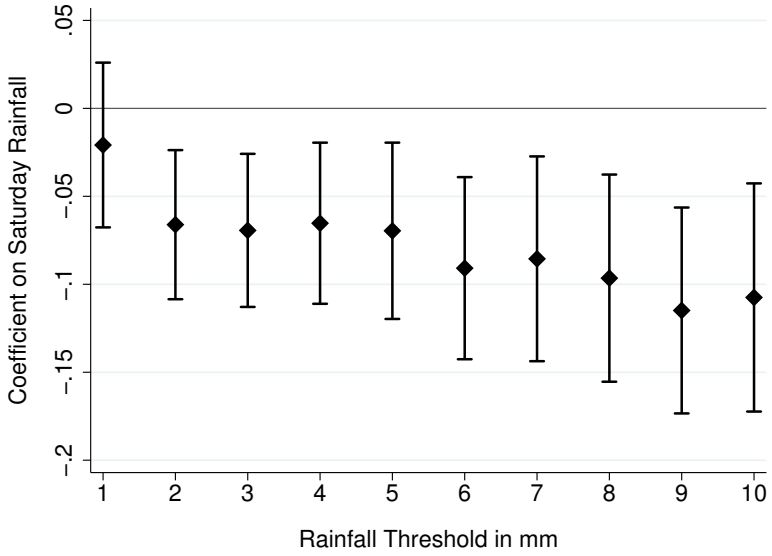
Figure 3: Effect of Saturday Rainfall on Contraceptive Adoption over Time



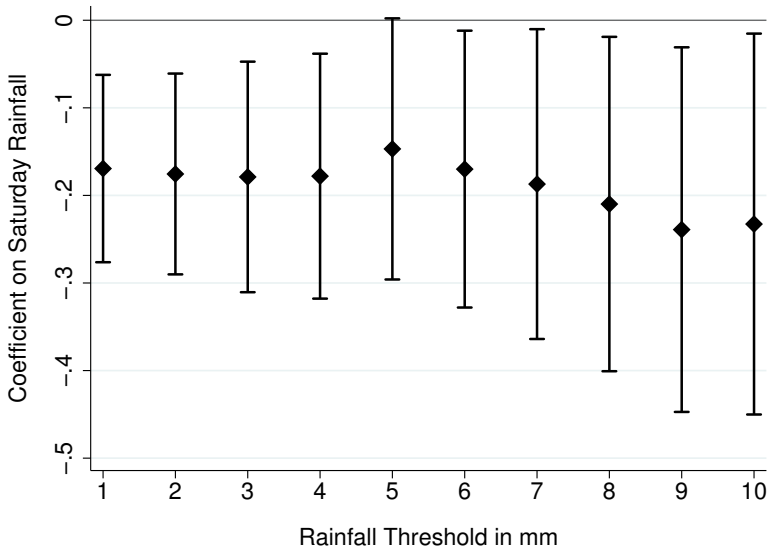
Notes: The figure presents rolling window coefficient estimates on # Sat.(Rainfall>3mm) (solid line) and 95% confidence intervals (dashed lines). The dependent variable is a monthly, binary indicator of contraceptive adoption. # Sat.(Rainfall>3mm) is the number of Saturdays with rainfall above 3 mm in a calendar month. All regressions include analogue rainfall regressors for the other weekdays. The rolling window size is 12 months. The regression estimates are displayed above the 7th month of the rolling window. The vertical line on April 2006 marks the introduction of performance contracts. Standard errors are clustered at the community level.

Figure 4: Effects of Saturday Rainfall at Different Thresholds Under Performance Contracts

(A) Contraceptive Adoption



(B) Bed Net Acquisition



Notes: The figures present the coefficients (diamonds) and 95% confidence intervals (capped bars) on the number of rainy Saturdays (# Sat.(Rainfall$>$ Xmm)) when varying the rainfall threshold in separate regressions. The dependent variables, Contraceptive Adoption and Bed Net Acquisition, are monthly, binary indicators. All regressions control for unit of observation and time fixed effects. Figure (A) uses monthly-date panel data of women for April 2006 to March 2007. Figure (B) uses months-before-interview panel data of households 0-11 months before DHS 2007-08. Standard errors are clustered at the community level.

Table 1: Summary Statistics of Family Planning Outcomes

A. Before (Apr. 2005 - Mar. 2006)	Mean	Min.	Max.	Obs.
Adoption	0.228	0	100	125,193
Conception	1.610	0	100	125,193
B. After (Apr. 2006 - Mar. 2007)				
Adoption	0.395	0	100	130,966
Conception	1.445	0	100	130,966

Notes: Based on 10,629 women in Before Panel and 11,103 women in After Panel. Only women with at least two panel observations are included in each data set.

Table 2: Summary Statistics of Acquisitions of Mosquito Bed Nets

A. Before (based on 2005 DHS)	Mean	Min.	Max.	Obs.
Bed Net Acquisition	0.636	0	100	121,752
Acquisition from Health Center	0.275	0	100	71,022
Acquisition from Other Source	0.536	0	100	71,022
B. After (based on 2007-08 DHS)				
Bed Net Acquisition	1.969	0	100	87,444
Acquisition from Health Center	1.049	0	100	51,009
Acquisition from Other Source	1.057	0	100	51,009

Notes: Based on 10,146 and 7,287 households with GPS coordinates in DHS 2005 and DHS 2007-08 data. Panel lengths of variables: 12 months (0-11) for 'Bed Net Acquisition'. 7 months (0-6) for 'Acquisition from Health Center' and 'Acquisition from Other Source'.

Table 3: Summary Statistics of Rainfall on Saturdays

Rainy Saturdays (Definition)	Mean	Std. dev.	Min.	Max.	Obs.
# Sat.(Rainfall>1mm)	1.249	1.125	0	5	11,808
# Sat.(Rainfall>2mm)	0.951	1.017	0	5	11,808
# Sat.(Rainfall>3mm)	0.764	0.938	0	5	11,808
# Sat.(Rainfall>4mm)	0.620	0.838	0	4	11,808
# Sat.(Rainfall>5mm)	0.512	0.756	0	4	11,808
# Sat.(Rainfall>6mm)	0.433	0.695	0	4	11,808
# Sat.(Rainfall>7mm)	0.374	0.641	0	4	11,808
# Sat.(Rainfall>8mm)	0.325	0.597	0	4	11,808
# Sat.(Rainfall>9mm)	0.273	0.547	0	4	11,808
# Sat.(Rainfall>10mm)	0.239	0.509	0	4	11,808

Notes: Based on 492 communities in the 2010 Rwandan DHS between April 2005 and March 2007. # Sat.(Rainfall> Xmm) is the number of Saturdays with rainfall above X mm in a calendar month.

Table 4: Main Effects

Dependent variable:	Contraceptive Adoption			Bed Net Acquisition		
Panel data:	after	before		after	before	
	(1)	(2)	p-value (1) − (2)	(3)	(4)	p-value (3) − (4)
# Sat.(Rainfall>3mm)	−0.071***	0.019	[0.008]	−0.203***	−0.019	[0.017]
	(0.023)	(0.025)		(0.070)	(0.032)	
# Sun.(Rainfall>3mm)	0.015	0.029	[0.688]	0.090	0.036	[0.512]
	(0.027)	(0.022)		(0.074)	(0.035)	
# Mon.(Rainfall>3mm)	−0.020	0.004	[0.575]	−0.115*	0.009	[0.086]
	(0.032)	(0.028)		(0.066)	(0.031)	
# Tue.(Rainfall>3mm)	−0.034	−0.009	[0.522]	0.027	0.028	[0.989]
	(0.033)	(0.021)		(0.070)	(0.038)	
# Wed.(Rainfall>3mm)	−0.036	0.012	[0.213]	−0.019	−0.033	[0.869]
	(0.032)	(0.021)		(0.074)	(0.037)	
# Thu.(Rainfall>3mm)	−0.014	−0.002	[0.750]	0.065	−0.043	[0.184]
	(0.027)	(0.027)		(0.075)	(0.033)	
# Fri.(Rainfall>3mm)	0.037	−0.020	[0.106]	0.016	−0.021	[0.611]
	(0.027)	(0.023)		(0.066)	(0.032)	
Unit FE	Yes	Yes		Yes	Yes	
Time FE	Yes	Yes		Yes	Yes	
Observations	130,966	125,193		87,444	121,752	
R-squared	0.080	0.082		0.086	0.091	
Dep. var. mean	0.395	0.228		1.969	0.636	

Notes: *Imihigo* performance contracts were signed at the beginning of April 2006. 12-months panel data for the time after (before) that date is used in regressions 1 and 3 (2 and 4). The dependent variables, Contraceptive Adoption and Bed Net Acquisition, are monthly, binary indicators. # Sat.(Rainfall>3mm) is the number of Saturdays with rainfall above 3 mm in a calendar month (and similarly for all other weekdays). The unit of observation in regressions 1 and 2 is a woman, and a time step is a monthly date. Regression 1 uses data for April 2006 to March 2007. Regression 2 uses data for April 2005 to March 2006. The unit of observation in regressions 3 and 4 is a household, and a time step is a month-before-the-interview. Regression 3 uses data for 0-11 months before the DHS 2007-08 interview. Regression 4 uses data for 0-11 months before the DHS 2005 interview. Standard errors are clustered at community level. P-value: *** p< 0.01, ** p< 0.05, ** p< 0.1.

Table 5: Tracing the Effects Under Performance Contracts

Dependent variable:	Contraceptive Adoption		Bed Net Acquisition	
	(1)	(2)	(3)	(4)
# Sat.(Rainfall>3mm)	−0.068***		−0.192***	
	(0.023)		(0.073)	
First Sat.(Rainfall>3mm)		−0.064		−0.097
		(0.057)		(0.161)
Second Sat.(Rainfall>3mm)		−0.117*		−0.125
		(0.060)		(0.144)
Third Sat.(Rainfall>3mm)		−0.085*		−0.265*
		(0.049)		(0.150)
Last Sat.(Rainfall>3mm)		−0.068		−0.285***
		(0.053)		(0.104)
L1.# Sat.(Rainfall>3mm)	0.046*		0.016	
	(0.027)		(0.065)	
L2.# Sat.(Rainfall>3mm)	−0.020		0.089	
	(0.025)		(0.062)	
L3.# Sat.(Rainfall>3mm)	0.011		0.045	
	(0.028)		(0.064)	
Other Rainfall Regressors	Yes	Yes	Yes	Yes
Unit FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
Observations	130,966	130,966	87,444	87,444
R-squared	0.080	0.080	0.086	0.086
Dep. var. mean	0.395	0.395	1.969	1.969

Notes: The dependent variables, Contraceptive Adoption and Bed Net Acquisition, are monthly, binary indicators. # Sat.(Rainfall>3mm) is the number of Saturdays in a month with rainfall above 3 mm. L1.# Sat.(Rainfall>3mm) is this variable lagged by one month (and similar for higher order lags). First Sat.(Rainfall>3mm) is a monthly, binary indicator which takes the value 100 if rainfall on the first Saturday of that monthly date is above 3 mm and 0 otherwise. Second, Third and Last Sat.(Rainfall>3mm) are corresponding indicators for rainfall on the other Saturdays in the calendar month. Other Rainfall Regressors are the numbers of days with rainfall above 3 mm for every other weekday. In regressions 1 and 2, the unit of observation is a woman, a time step is a monthly date and the data are for April 2006 to March 2007. In regressions 3 and 4, the unit of observation is a household, a time step is a month-before-the-interview, and the data are for 0-11 months before the DHS 2007-08 interview. Standard errors are clustered at community level. P-value: *** p< 0.01, ** p< 0.05, ** p< 0.1.

Table 6: Bed Net Acquisition from Different Sources Under Performance Contracts

Dependent variable:	Bed Net Acquisition from ...		
	Any source	Health Facility	Other source
	(1)	(2)	(3)
# Sat.(Rainfall>3mm)	−0.211**	−0.029	−0.177**
	(0.100)	(0.070)	(0.073)
# Sun.(Rainfall>3mm)	0.143	0.034	0.114
	(0.093)	(0.064)	(0.073)
# Mon.(Rainfall>3mm)	−0.099	−0.015	−0.072
	(0.083)	(0.052)	(0.062)
# Tue.(Rainfall>3mm)	0.031	0.010	0.024
	(0.095)	(0.065)	(0.068)
# Wed.(Rainfall>3mm)	0.063	0.103	−0.055
	(0.100)	(0.066)	(0.072)
# Thu.(Rainfall>3mm)	0.115	0.039	0.093
	(0.114)	(0.075)	(0.085)
# Fri.(Rainfall>3mm)	0.014	0.034	−0.010
	(0.097)	(0.065)	(0.072)
Household FE	Yes	Yes	Yes
Months-before-Interview FE	Yes	Yes	Yes
Observations	51,009	51,009	51,009
R-squared	0.147	0.148	0.150
Dep. var. mean	2.082	1.049	1.057

Notes: The dependent variables, bed net acquisitions from different sources, are monthly, binary indicators that take the value 100 if a bed net was acquired from that source and 0 otherwise. # Sat.(Rainfall>3mm) is the number of Saturdays with rainfall above 3 mm in a month before the interview (and similarly for all other weekdays). The data are for 0-6 months before the DHS 2007-08 interview. Standard errors are clustered at community level. P-value: *** p< 0.01, ** p< 0.05, ** p< 0.1.

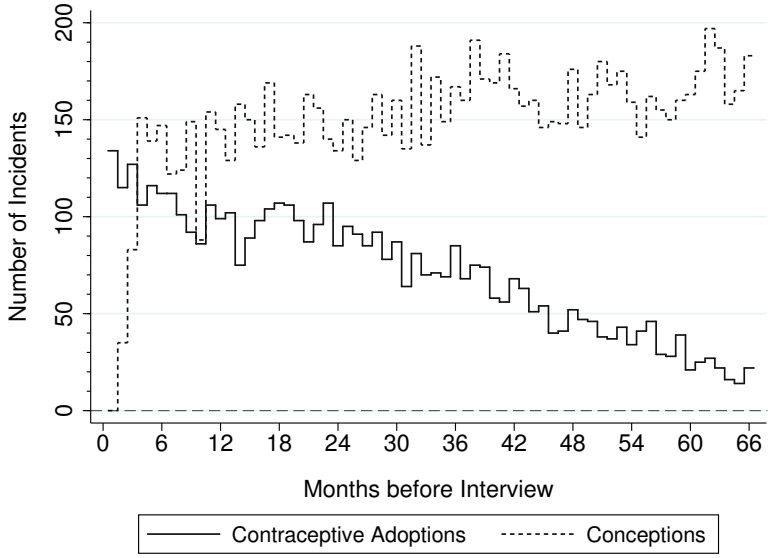
Table 7: Mechanism: Conception and Bed Net Acquisition by Altitude

Dependent variable:	Conception			Bed Net Acquisition	
Panel data:	after	before		Median split by altitude high	low
	(1)	(2)	p-value (1) − (2)	(3)	(4)
# Sat.(Rainfall>3mm)	−0.121**	0.062	[0.024]	−0.256**	−0.178*
	(0.048)	(0.066)		(0.099)	(0.100)
# Sun.(Rainfall>3mm)	0.050	−0.075	[0.143]	0.089	0.061
	(0.057)	(0.059)		(0.107)	(0.113)
# Mon.(Rainfall>3mm)	0.063	0.016	[0.614]	−0.128	−0.094
	(0.062)	(0.068)		(0.079)	(0.105)
# Tue.(Rainfall>3mm)	−0.088	−0.053	[0.705]	−0.019	0.064
	(0.061)	(0.066)		(0.099)	(0.099)
# Wed.(Rainfall>3mm)	−0.145**	−0.005	[0.096]	−0.017	−0.012
	(0.062)	(0.058)		(0.104)	(0.106)
# Thu.(Rainfall>3mm)	0.024	−0.133*	[0.079]	0.056	0.069
	(0.052)	(0.070)		(0.095)	(0.119)
# Fri.(Rainfall>3mm)	0.075	−0.005	[0.323]	−0.033	0.064
	(0.055)	(0.062)		(0.086)	(0.103)
Unit FE	Yes	Yes		Yes	Yes
Time FE	Yes	Yes		Yes	Yes
Observations	130,966	125,193		43,500	43,944
R-squared	0.072	0.070		0.083	0.087
Dep. var. mean	1.445	1.610		1.733	2.203

Notes: *Imihigo* performance contracts were signed at the beginning of April 2006. The dependent variables, Conception and Bed Net Acquisition, are monthly, binary indicators. # Sat.(Rainfall>3mm) is the number of Saturdays with rainfall above 3 mm in a calendar month (and similarly for all other weekdays). The unit of observation in regressions 1 and 2 is a woman, and a time step is a monthly date. Regression 1 uses data for April 2006 to March 2007. Regression 2 uses data for April 2005 to March 2006. The unit of observation in regressions 3 and 4 is a household, and a time step is a month-before-the-interview. Both regressions use data for 0-11 months before the DHS 2007-08 interview. Regression 3 only uses observations from villages located above median altitude, whereas regression 4 only uses observations from villages at and below median altitude. The median altitude is 1,670 meters above sea level. Standard errors are clustered at community level. P-value: *** p< 0.01, ** p< 0.05, ** p< 0.1..

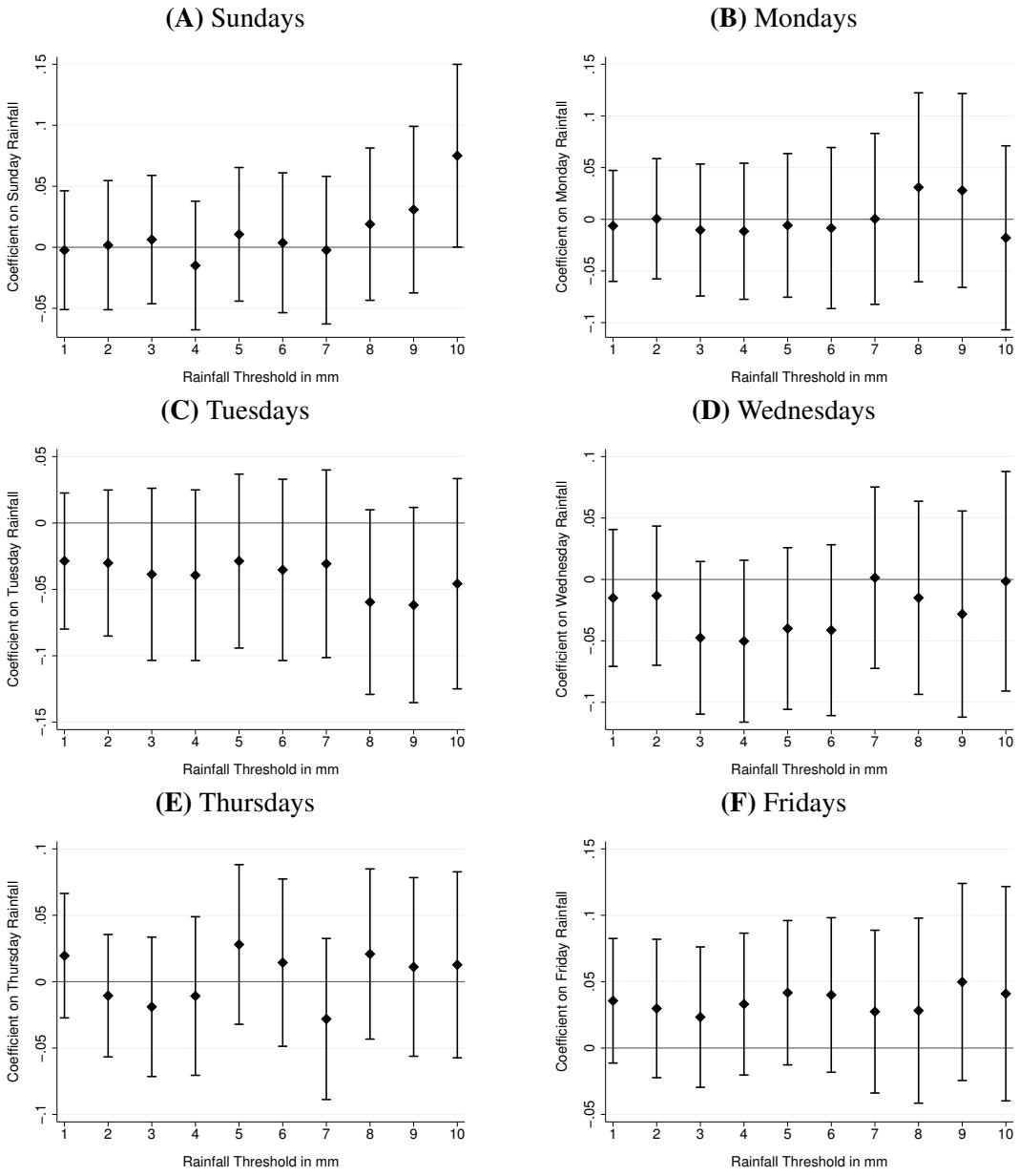
Appendix (for online publication)

Figure A1: Number of Contraceptive Adoptions and Conceptions in Different Months before the Interview



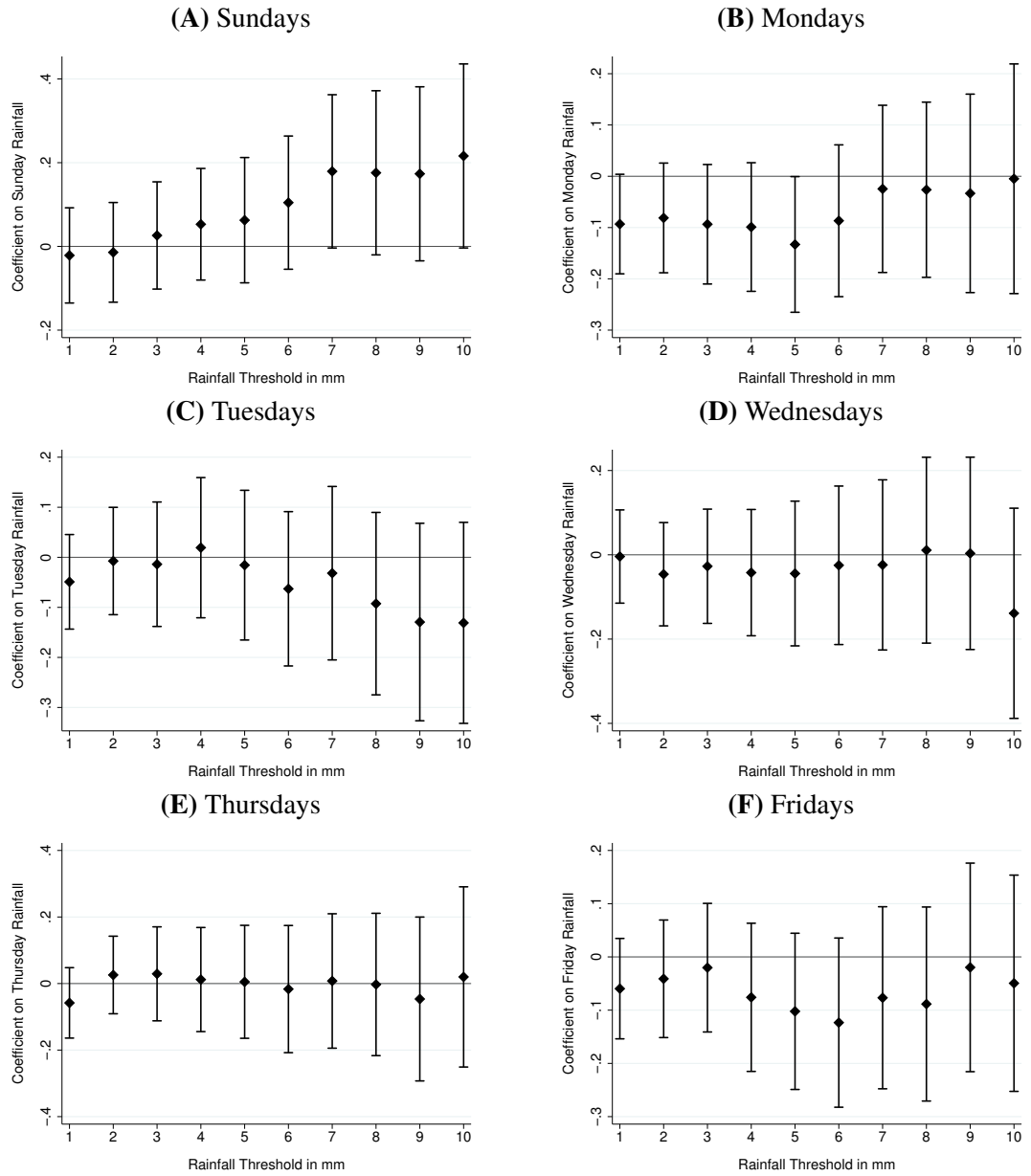
Notes: Based on 13,413 women between 15 and 49 years old and who are usual residents of interviewed households in the 2010 Rwandan DHS. The solid, vertical line marks the introduction of performance contracts in April 2006.

Figure A2: Contraceptive Adoption Under Performance Contracts: Effects at Different Thresholds



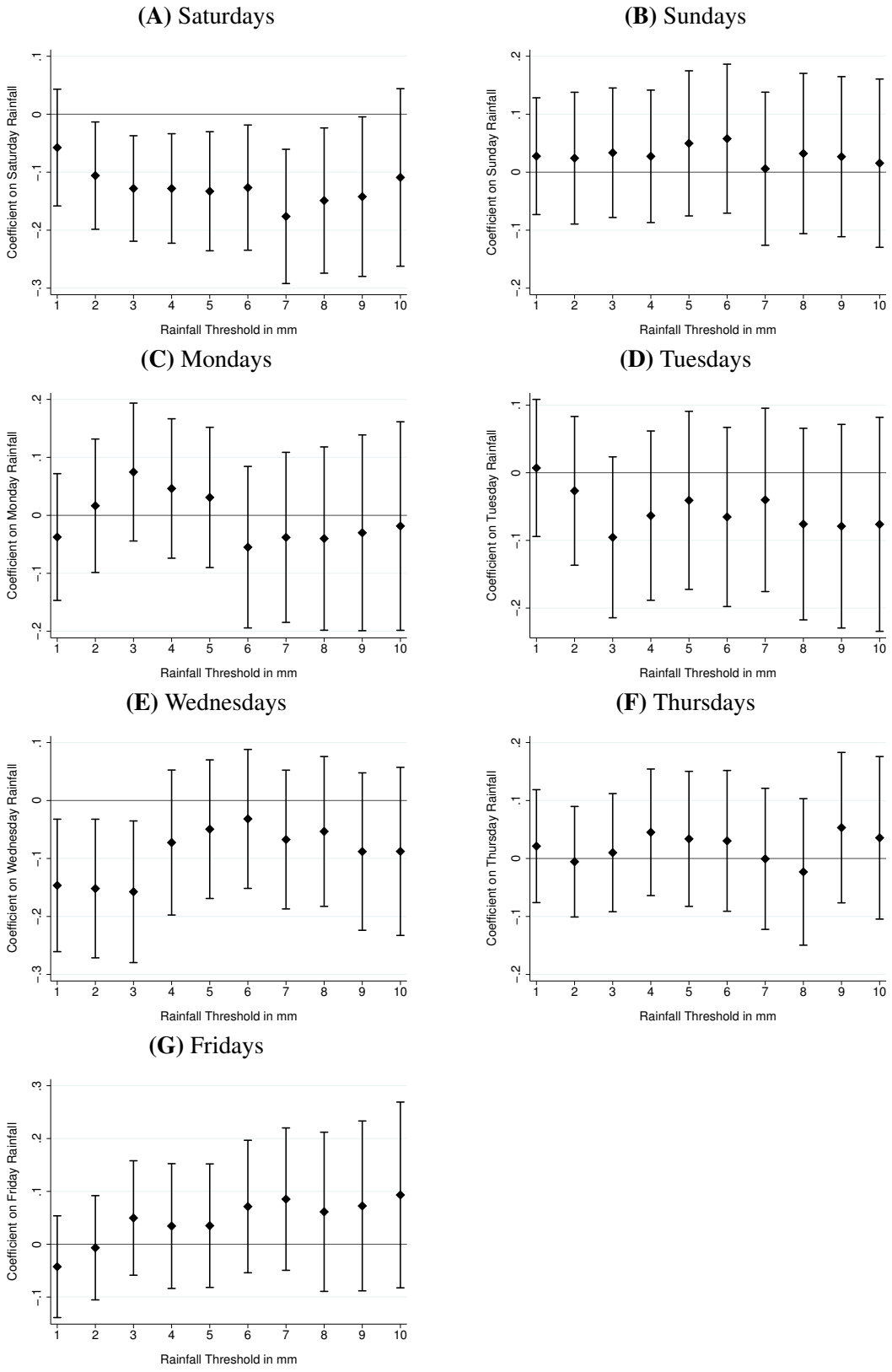
Notes: The figures present the coefficients (diamonds) and 95% confidence intervals (capped bars) on the number of rainy days on different weekdays when varying the rainfall threshold in separate regressions. The dependent variable, Contraceptive Adoption, is a monthly, binary indicator. All regressions control for women and monthly date fixed effects. The data are for April 2006 to March 2007. Standard errors are clustered at the community level.

Figure A3: Bed Net Acquisition Under Performance Contracts: Effects at Different Thresholds



Notes: The figures present the coefficients (diamonds) and 95% confidence intervals (capped bars) on the number of rainy days on different weekdays when varying the rainfall threshold in separate regressions. The dependent variable, Bed Net Acquisition, is a monthly, binary indicator. All regressions control for household and months-before-interview fixed effects. The data are for 0-11 months before DHS 2007-08 interviews. Standard errors are clustered at the community level.

Figure A4: Conception Under Performance Contracts: Effects at Different Thresholds



Notes: The figures present the coefficients (diamonds) and 95% confidence intervals (capped bars) on the number of rainy days on different weekdays when varying the rainfall threshold in separate regressions. The dependent variable, Conception, is a monthly, binary indicator. All regressions control for women and monthly date fixed effects. The data are for April 2006 to March 2007. Standard errors are clustered at the community level.