WIDER Working Paper 2017/130

The donor footprint and gender gaps

Maria Perrotta Berlin, Evelina Bonnier, and Anders Olofsgård*

June 2017

United Nations University World Institute for Development Economics Research

wider.unu.edu

Abstract: In this paper we analyse the impact of foreign aid on gender outcomes and attitudes. We do this by matching geocoded household surveys with aid projects. This offers a middle way between project evaluations and aggregated cross-country comparisons, measuring an average community effect around projects. We find increased opportunities for women to work outside the household, which could strengthen their bargaining power. However, we find mixed results in terms of the impact on women's control over other key areas of their lives. We argue this is related to differences in what is required for change to happen at the community level.

Keywords: foreign aid, gender, development, geocoded data, matching, impact analysis
JEL classification: P45, O12, O19, J16, C55

Tables and figures: All authors' own work.

* All authors: Stockholm Institute of Transition Economics (SITE), Stockholm School of Economics, Stockholm, Sweden; corresponding author: anders.olofsgard@hhs.se.

This study has been prepared within the UNU-WIDER project on 'Gender and development'.

Information and requests: publications@wider.unu.edu

ISSN 1798-7237 ISBN 978-92-9256-356-1 https://doi.org/10.35188/UNU-WIDER/2017/356-1

Typescript prepared by Luke Finley.

The United Nations University World Institute for Development Economics Research provides economic analysis and policy advice with the aim of promoting sustainable and equitable development. The Institute began operations in 1985 in Helsinki, Finland, as the first research and training centre of the United Nations University. Today it is a unique blend of think tank, research institute, and UN agency—providing a range of services from policy advice to governments as well as freely available original research.

The Institute is funded through income from an endowment fund with additional contributions to its work programme from Denmark, Finland, Sweden, and the United Kingdom.

Katajanokanlaituri 6 B, 00160 Helsinki, Finland

1 Introduction

In this paper we look at how individual- and household-level outcomes and attitudes with regards to women's and girls' rights and opportunities co-vary with the presence of aid projects in the geographical neighborhood of the household. We consider both aid projects in general and aid projects with a specific gender component. Our analysis relies on combining two distinct sets of data. First, using new geocoded data on the history of aid projects in Uganda and Malawi (the only two countries for which there exists a nearly complete recent history of geocoded aid projects), we create what we refer to as the donor footprint, a map of the location of all aid projects dating back to 1999 in Malawi and 2003 in Uganda. Second, we collect information on a set of outcomes related to women's empowerment and opportunities from geocoded household-level survey data. We match these datasets by creating a 15-kilometer (km) circle around each household, and checking for the presence of aid-financed projects in this immediate vicinity of the household. This creates two groups of households, those with at least one aid project in their vicinity and those without.

Gender equality spans many dimensions and we analyze the effect of aid across six categories of outcomes: workforce participation (including child work), attitudes towards domestic violence and women's sexual empowerment, experience of domestic violence, household expenditures and decision-making, time spent on household chores, and schooling. We base our expectations on the logic of the literature on intra-household bargaining, discussed in more detail in the next section. Our starting point is that aid can influence a woman's bargaining position through offering a better outside option, in the form of access to financial revenues from job opportunities or control over social transfers (such as conditional or unconditional cash transfers). We do not have data on incomes separated by gender, but we do have data on jobs and work on the farm (typically the main source of resources of rural families). We then argue that through an improved bargaining position, women can gain stronger control over areas such as their own sexuality, freedom from domestic violence, family budget decisions, and their own time (with respect to household chores). The extent to which this happens, however, and is picked up by our data depends on the details. Questions that focus on attitudes rather than actual outcomes or actions are probably easier to affect (and may be prone to response bias). Moreover, changes in attitudes might be easier to observe community-wide if they relate to norms of behavior that may spread through networks beyond immediate aid beneficiaries. On the other hand, changes in some behavior that is public and goes against expectations and norms may be the least likely to be affected.

Our results suggest some positive community impact in the neighborhood of aid projects. Women are more likely to work and are, at least in relative terms, more engaged in farm work. In our framework, the former suggests that there is potential for more control rights also over the other areas. However, results in this respect are at best mixed. We do find some positive impact on attitudes towards female sexual empowerment but no impact on attitudes towards wife beating. We do, however, find positive effects on reported experience of physical violence and, for gender projects, also experienced sexual violence.

In terms of control over the family budget, we find no impact on women's participation in financial decisions. Even more puzzling in terms of budget priorities, we find a negative effect on spending on children

and, for gender projects, a positive effect on traditional male goods such as alcohol and tobacco. This is actually consistent with some recent research questioning the assumption that fathers and mothers have different priorities over spending on children (and finding that fathers may even prioritize children's education more than mothers) and the fact that gender equality is typically associated with smaller gender gaps in consumption of alcohol and tobacco (i.e. preferences may converge within the household), but we do not have enough evidence to strongly argue for such an interpretation.

In terms of control over time, i.e. the allocation of time spent on key household chores such as fetching water and collecting firewood, constraints against community-wide impacts are presumably quite strong given that these activities are traditionally a female prerogative, often done in groups, and breaking the norm is publicly visible. Here, we find that adult women take responsibility for a greater share of the household's time spent on these activities in the presence of aid projects, while girls' share decreases (as does boys' and men's). Hence, at least in relative terms, we see no improvements from living in the vicinity of aid projects.

Finally, looking at education we find no significant differences for either boys or girls when it comes to primary and secondary school attendance, but given the lack of an average difference in attendance across gender prior to our aid projects (see Section 3), this is not surprising. On the other hand, we do find that both boys and girls are less likely to work outside the household and on the farm, and effects are somewhat stronger for girls.

In terms of gender-specific versus general aid projects, results are typically similar in qualitative terms, but often stronger with general aid. It is not obvious how to interpret this, though. It could be interpreted as support for the effect of gender mainstreaming, or that activities affecting the welfare of communities more broadly have at least as much effect as gender-specific interventions. However, our estimates also rely on fewer observations and a different selection when we look at gender projects, therefore we cannot rule out just statistical reasons for the differences.

All in all, these results suggest that aid projects may have some positive gender-related effects in the communities surrounding them, but also that the effects are quite limited. In particular, in areas that require men to give up something (expenditure priorities) and where change would go against established behavior and norms and would be publicly observable (household chores), it is perhaps not surprising that the average effects at the community level are not (yet) observed. Gender relations depend on a host of factors, including slow-moving institutions such as social norms or culture more generally. New opportunities and expectations on women may initially be met by resistance, and even if aid projects may succeed in bringing short-term opportunities and affect attitudes among targeted beneficiaries, the effects may not be strong enough to change communities, at least not within the time frame that we are able to study.

With this paper, we offer three contributions to the existing literature. First, we provide new evidence on a question that has gotten surprisingly little attention in the academic literature: the general impact of foreign aid on women and gender equality. In so doing we distinguish between the average impact of projects specifically designed for women and the average impact of all types of projects. The latter approach is motivated by the explicit ambition of Western donors (at least since the Fourth World Conference on

Women, Beijing 1995) to "gender mainstream" all foreign aid.

Second, studies of aid effectiveness tend to either focus on the very micro level, the impact on direct participants in a specific project, or the very macro level, the effect of total aid on GDP growth or some other country-level aggregate. The first approach suffers the risk of problems of external validity due to site selection bias (Allcott, 2015), and positive or negative community-level externalities are often neglected. The second approach, in turn, faces challenges with internal validity and measurement, giving rise to a much divided literature (e.g. Rajan and Subramanian, 2008; Galiani et al., 2017). Recently, due to new data on aid project placement within countries, an intermediate approach, sometimes referred to as geospatial impact evaluation, has emerged focusing on the average impact on communities at a more disaggregated level. We contribute to this literature not only by being the first to use this approach in order to look at gender outcomes, but also by taking a more disaggregated approach than almost all previous studies and incorporating aid from all donors, while the existing literature tends to focus on only one donor.

Finally, to tackle selection problems in the placement of aid projects we apply machine learning and matching techniques as summarized in recent contributions by Imbens (2014) and others to create more comparable treatment and control units and derive more robust treatment effects. The details of the method are presented in Section 4. But first, in Section 2, we outline the foundations for our approach. In Section 3 we present the data used in our analysis and some descriptive statistics of the pre-existing gender gaps in the two countries. After the empirical strategy and specification, Section 5 reports our results. We conclude in Section 6.

2 Foreign Aid and Gender Outcomes

Western donors have for many years singled out women and girls in partner countries as important targets for foreign aid. This is partly driven by the perception that women and girls are discriminated against and are particularly vulnerable in situations of poverty and conflict. However, it is also emphasized that empowered and educated women are seen as more likely to make decisions that promote development. For instance, micro-credit is often targeted toward women because they are seen as more reliable borrowers (Chakravarty et al., 2013), female education is negatively correlated with fertility (McCrary and Royer, 2011), and better-off and educated women are more likely to make sound decisions when it comes to their children's education and health (Gakidou et al., 2010). Many aid projects therefore have an explicit gender component, which can be anything from targeting increased school enrolment of girls, to programs empowering women within the household or the community (see OECD, 2016, for a breakdown of different activities of the OECD Development Assistance Committee (DAC) members). However, donors also agreed to gender mainstreaming already in the Beijing Platform of Action in 1995, i.e. that gender concerns should be integrated into all policy and program cycles and governments should engage in a dialogue on gender and development. For instance, even though only 6% of interventions funded by the Swedish International Development Cooperation Agency (Sida) had gender equality as the prime objective in 2008–10, as many as 71% had gender equality

as a significant objective (Nanivazo and Scott, 2012). Gender mainstreaming thus implies that also projects and programs not primarily targeting gender equality are often designed with an eye towards that objective. This ambition is also shown in the development of a Gender Equality Policy Marker by the OECD DAC, an indicator that measures how strongly aid-financed projects support gender.[1]

What do we then know from the literature about the impact of aid on gender outcomes at a more aggregated level? Foreign aid is partly invested in public goods and services that have been known to particularly benefit women in different ways, such as health spending on maternity care and investments in water and sanitation (Agénor et al., 2014; Seguino, 2008). Whether this aggregates up to a positive effect of foreign aid at the macro level has, however, not been much explored. Pickbourn and Ndikumana (2016) use cross-country data to estimate the correlations between foreign aid inflows and the UNDP's Gender Inequality Index (GII), but find no robust results. At this level of aggregation, this is perhaps not so surprising. The well-known micro–macro paradox from the aid effectiveness literature may very well apply here as well, since public investments tend to be fungible and the additional resources that aid provides may also be used to strengthen an existing patriarchal system (Richey, 2003). The authors do, however, find some positive effects from aid to the health sector on maternal mortality rates, and from aid to the education sector on female relative to male youth literacy rates. It should also be emphasized that the GII leaves out many important aspects that we typically associate with gender equality.[2] An approach in between the country aggregate and the targeted beneficiaries may therefore be more useful, looking at the influence of aid on different aspects of gender equality separately.

2.1 Our Approach

Given the focus on gender mainstreaming, the first cornerstone of our approach is to look not only at projects with a defined gender component, but also at aid generally. Just as the aid effectiveness literature typically does not separate between aid specifically targeted towards economic growth and other purposes (peace and security or human rights and democracy, for instance), our baseline model looks at an average effect of all types of projects, with the ambition to estimate the effectiveness of gender mainstreaming. This implies that these results can only be interpreted at this level of aggregation, i.e. a null result can only speak to the success of mainstreaming, not to the impact of projects explicitly targeting gender issues. We do, however, also look separately at what can broadly be classified as "gender projects"[3] to see if there is a differential impact from projects with a specific gender component.

The second cornerstone of our analysis is that we take a community-level approach to measure the impact.

[1] http://www.oecd.org/dac/stats/37461060.pdf.

[2] "Like all composite measures, the GII has some limitations. First, it does not capture the length and breadth of gender inequality. For example, the use of national parliamentary representation excludes participation at the local government level and elsewhere in the community and public life. The labor market dimension lacks information on employment, having an adequate job, and unpaid work that is mostly done by women. The index misses other important dimensions, such as time use – the fact that many women have the additional burden of care giving and housekeeping cuts into their leisure time and increases stress and physical exhaustion. Asset ownership, child-care support, gender-based violence and participation in community decision-making are also not captured in the GII, mainly due to limited data availability" (http://hdr.undp.org/en/content/what-are-strengths-and-limitations-gii).

[3] The exact definition of this classification is presented in Subsection 3.1.

Aid-financed projects are intended to offer citizens in poor communities additional economic and/or social opportunities. Projects have directly targeted beneficiaries, such as school children in the case of investments in school buildings, mothers and small children in the case of investments in maternity care, and farmers in the case of irrigation projects. However, this does not mean that the benefits to the community stop there. Aid programs are typically intended to benefit whole communities, and there may be positive (or negative) externalities that extend beyond the immediately targeted beneficiaries. Think for instance of road projects connecting villages and markets, and immunization drives against infectious diseases (Miguel and Kremer, 2004). Projects targeting livelihood and jobs can for example increase aggregate demand in the community, benefiting also those not directly involved in the projects, and research shows that gender relations across most (but not all) dimensions tend to become more equal with economic development and that women tend to gain more than men (Duflo, 2012). Finally, aid projects can affect norms through increased exposure to individuals with different cultures and norm systems, be it foreigners or just more-educated NGO workers from urban areas. As illustrated in Andrabi and Das (2017), exposure to aid workers can increase trust in foreigners, presumably making individuals more responsive to new opportunities introduced by them. This impact can then travel beyond the direct participants through network effects in the geographical vicinity of projects (Cai et al., 2015; Conley and Udry, 2010).[4]

We therefore follow the recent literature on geospatial impact evaluation, focusing on the impact of aid projects in their geographical vicinity. Beyond the studies focused on questions of aid-allocation decisions, several outcomes have been considered in this literature to capture the impact of aid. Dreher and Lohmann (2015) study the regional impact of World Bank projects on nighttime light as a proxy for economic growth, finding a positive correlation but no convincing support for a causal impact. Findley et al. (2011), Wood and Sullivan (2015), and Strandow et al. (2014) find that aid project location is correlated with more local violence and conflict. De and Becker (2015) find that health and education aid have positive impact on some health and education indicators in Malawi. Finally, some studies have focused on Chinese aid. Isaksson and Kotsadam (2016) find that perceptions of corruption are higher in the vicinity of Chinese aid projects, whereas BenYishay and Mobarak (2015) find that the impact of Chinese-funded infrastructure projects on forest loss depends on the quality of domestic environmental governance. Relative to this literature, and as mentioned in the Introduction, our paper not only looks at a different outcome but also includes all donors and at a level of disaggregation rarely used.

The third and final cornerstone is that, since gender (in)equality can manifest itself in many different ways, we look at a broad range of outcome variables. We take our starting point from models of intra-household bargaining, emphasizing how bargaining power relies on jobs and income streams as they reflect outside options (e.g. Thomas, 1990). We do not have data on income separated by gender, but we do have gender-specific information on *work for a salary and work on the farm* (typically a rural household's main source

[4]How development policy more generally can change mental models is also evidenced by evaluations of the Indian law to have mandatory female leaders in one-third of village governments. Not only did males' prejudice against female leaders fall and women get elected even after reservations ended, but parents' aspirations for their daughters, in general, also increased (Beaman et al., 2009; Beaman et al., 2012). The reform also increased the reporting of and police responsiveness to crimes against women, suggesting a change in attitudes towards what is acceptable among both women and men (Iyer et al., 2012).

of revenues or resources). Our first hypothesis, then, is that aid presence offers opportunities for women to earn money outside the household and gain control over the income-generating resources of the family. Micro-credit programs for instance often target women, and in recognition of women's often traditionally limited role in rural livelihood, programs increasingly target women or at least explicitly include them. A positive impact on women's possibility of raising the revenues of the household would thus suggest also the potential of aid to give women some more control over their own lives in other areas.

One such area would be power over their own *sexuality and freedom from domestic violence*. The lifetime opportunity cost of teen pregnancy for Uganda (calculated based on lost income) could be as high as 30% of annual GDP (Chaaban and Cunningham, 2011), and Koenig et al. (2003) show that in the Rakai district of Uganda, 70% of men and 90% of women viewed beating one's wife or female partner as justifiable in some circumstances. In our data both men and women are asked about their opinions on wife beating and opinions regarding use of contraceptive and a woman's right to say no to sex. The surveys also include questions on actual exposure to physical, sexual, and emotional violence. Aid-financed projects can provide information and possibly change social norms for both men and women in these areas through information and discussions groups (Kim et al., 2007; Pulerwitz et al., 2015). Furthermore, a study on sexual risk-taking intentions among school-going AIDS-orphaned adolescents in rural Uganda (Ssewamala et al., 2010) found that having access to economic assets plays an important role in influencing adolescents' sexual risk-taking attitudes. Munshi and Myaux (2006) emphasize how norms with regards to contraception changed through networks based on religious affiliation in the Matlab region of Bangladesh. Women who came into contact with the relatively well-educated health workers of the Matlab program were affected by new norms and information around sexuality and contraception, and these women, in turn, transmitted information through their networks. This illustrates how whole communities, beyond direct beneficiaries, can be influenced by aid-financed development projects. More generally, many studies suggest a negative correlation between female earnings and domestic violence, in line with an intra-household bargaining framework (e.g. Aizer, 2010; Bobonis et al., 2013; Anderberg et al., 2016). The hypotheses in this regard would thus be that aid has a positive impact on attitudes around domestic violence and female sexual empowerment and that it reduces exposure to domestic violence.

Another possible bargaining outcome concerns women's *control over the household budget*, both in terms of revenue streams and in terms of expenditure priorities. It has become increasingly common to target transfer schemes, such as old-age pension and conditional or unconditional cash transfers, to women (Attanasio and Lechene, 2002; Duflo, 2012). This is partly done to correct a perceived imbalance between men and women in terms of control of household resources, but also because women often are thought of as making spending decisions more in line with the welfare of the family, not least the children (Thomas, 1990; Case and Deaton, 1998; Lundberg et al., 1997). Our data contains two sets of relevant questions. The first asks women whether they participate in the household's financial decisions. The second asks about household expenditures, and there we focus on relative spending on food, clothing, health, and education for children and what are referred to as "male goods", such as alcohol and tobacco. The first question measures empowerment most directly,

whereas the other question builds on an assumption about diverging preferences over expenditures that has long been the norm in the literature but has been challenged in recent studies using randomization to identify causal effects (Benhassine et al., 2015; Akresh et al., 2016). These later results are also confirmed in lab experiments by Ringdal and Hoem Sjursen (2017), who find that more bargaining power in the hands of mothers does not increase spending on children's education; if anything it is the other way around. They also emphasize that other characteristics, such as time preferences, are more important than gender in determining the share of a given budget allocated to children's education. So, to the extent that presence of aid projects strengthens women's outside options (through for instance work opportunities or targeted transfers) we should expect aid to increase women's say in financial decisions, whereas the effect on expenditure patterns of the family is unclear. Changes in this respect may also be more difficult to pick up at the community level compared to changes in attitudes towards domestic violence and sexuality, as they may be less prone to social contagion through norms and expectations.

Bargaining power can also influence women's *control over their own time*, in particular their share of time devoted to household chores. Here we focus on survey questions asking about time spent on fetching water and collecting firewood, as time-use data suggest these two activities account for the biggest gap between men and women in time on household activities (Blackden and Wodon, 2006). Research from the US suggests that increasing female relative wages reduces time spent on household chores, though effects are quite limited in size and smaller for couples with children (Friedberg and Webb, 2005). On the other hand, a gender-targeted conditional cash transfer program in Pakistan reduced time spent by women on children's needs by 100 minutes on average per day but also increased the time spent on housework by 120 minutes, suggesting a net reduction in time for work or leisure (Hasan et al., 2010). Generally, research suggests that time use is highly dependent on norms (Maxwell and Wozny, 2017), and men fetching water and collecting firewood is a very visible way of breeching a norm. Our prior is therefore that it is unlikely that we would find significant impact of foreign aid presence on the division of time in the household towards these particular activities.

Finally, we also look at *schooling*. Aid projects can influence schooling for both girls and boys through many different channels such as directly by new school buildings, conditional cash transfers, information about returns to education, and increased economic opportunities (income effects). Here we focus specifically on girls' education. Girls have made substantial relative progress when it comes to primary education (which is now nearly universal), but less progress when it comes to secondary education. Gross secondary enrolment rates for girls in low- and middle-income countries increased from 22 to 34% between 1991 and 2010, but the gap to boys only shrank from 8 to 7% (Duflo, 2012). The question is to what extent projects contribute specifically to female education. Aid projects can influence these imbalances in many ways: targeting conditional cash transfers to mothers under the assumption that they care relatively more about girls' education; provision of information about economic returns to girls' education; and indirectly by contributing to reduced poverty and changing social norms (Duflo, 2012; Benhassine et al., 2015). Whether these effects are strong enough to appear at the community level is an open question, though.

3 Data and Descriptives

As mentioned earlier, our approach to understanding how gender outcomes and attitudes co-vary with the presence of aid projects relies on the possibility of linking the precise location of aid projects to geocoded household-level data as well as predetermined covariates from the same locations. Given that we are interested in understanding the average effect of any type of aid project from any type of donor, we focus on the countries for which a mapping of at least 80% of total aid flows in recent years is available. This leaves us with two suitable countries: Uganda and Malawi. These two countries will therefore be the focus of this study.

3.1 Treatment Variables

The newly released datasets on Uganda and Malawi from the AidData consortium (2016) provide geocoded information on all aid projects reported to these countries' respective Aid Information Management Systems (AIMS).[5] More precisely, we use the Uganda AIMS, level 1, version 1.4.1, which includes all geocoded projects from Uganda's Aid Management Platform (AMP; Peratsakis et al., 2012), corresponding to a total of 2,426 project locations over the years 1988–2014. For Malawi, we use the Geocoded Activity-Level Data from the Government of Malawi's AMP. This dataset includes 80% of all externally funded projects initiated in the period 1996–2011, which corresponds to a total of 2,523 project locations.

The project locations are provided with different levels of precision. We limit our sample to only include projects where the geographic coordinates correspond to the exact location (precision code 1) or a location known to be within 25km of the reported coordinates (precision code 2). In other words, we exclude projects formally allocated to the whole administrative area of the district or above. Our rationale for doing so is that it is the projects that are precisely geocoded that affect nearby households on the margin, while we assume that the influence of projects with larger coverage is uniformly spread across households in the corresponding administrative unit. Moreover, in order to match these project locations to both predetermined covariates and post-intervention outcomes, we further restrict the sample to projects that were initiated in 2003–2011 for Uganda and 1999–2010 for Malawi. Applying these restrictions leaves us with a dataset including 598 project locations from the Uganda AIMS and 1,001 project locations from Malawi's AMP. The distribution of these project locations can be seen in Figures 1 to 3.[6] However, we will use the information on the number of project-years at precision level 3 (district level) as a control variable in our empirical estimation.

Our first treatment, the presence of at least one aid project in the immediate vicinity of a household, is constructed by drawing a circle with a 15km radius around each household's location and checking whether

[5] The AIMS was developed and introduced by the State Committee on Investments and State Property Management (SCISPM) with technical support from the United Nations Development Programme (UNDP) and funded by Department for International Development (DFID). It is a web-based Aid Management Platform (AMP) that allows governments of developing countries and their donors to share and analyze aid information. The data and information collected include the number of implemented projects and agreements, their cost, terms, and duration, and executing and implementing agencies.

[6] Most of the excluded projects are dropped because their level of geographic precision is above prevision level 2. For Uganda, 1,588 project locations have a precision level above 2, of which 1,300 are at the district level (precision level 3), 33 at the province level (precision level 4), 5 at the level of larger geographic features such as rivers or national parks (precision level 5), and 255 at the country level (precision level 6 and 8). Another 239 project locations are dropped because their implementation year was prior to 2003 or after 2011. For Malawi, 1,516 observations are dropped because their precision level is above 2. Of these, 1,061 project locations have precision level 3, 34 have precision level 4, 25 have precision level 5, and 396 have precision level 6 or 8. Another six project locations are dropped because their implementation year was either before 1999 or after 2010.

any type of aid project was implemented there during the relevant time period. Our second treatment is the presence of at least one "gender project" and it is constructed in the same way as our first treatment variable. We define a project as a gender project if its title, project description, or activities list include any of the following words: *women*, *girl*, *bride*, *maternal*, *gender*, *genital*, and *child*. There are 77 projects that meet this definition in Uganda and 277 in Malawi.

The appropriate radius to define the treatment is debatable. On the one hand, there are purely statistical considerations such as the size of the treatment and control groups and the noise implied by potential misclassification of individual units. On the other hand, the very idea of impact of a development project is a question of how far the presence of a facility (school, clinic, other service delivery point, infrastructure) is noticeable and impacting the individual's conditions, and how far is too far. Obviously this will depend on the type of project. To give an idea, the 2009 Uganda Malaria Indicator Survey found that 96% of respondents lived within 9km of a health-care facility, and the 2009 Uganda National Household Survey reported that the average distance of a household from a government hospital was 20km. Less than 2% of the respondents in the Malawi Demographic and Health Survey (DHS) 2004 lived farther away than 15km from a health facility. In half of the school districts in Malawi the average distance to a school is more than 4km, which is considered too far by parents, according to a survey cited in Ravishankar et al. (2016). Kabunga et al. (2016) report that the median distance to a primary school is 2km in Uganda, and twice as far to a secondary school. The average distance to a water source is 1km and to an all-weather road 2km. We try to strike a balance between these very different acceptable distances and take into account the substantial variation in the nature of the aid projects we study. We have also tried a 10km radius and the results are very similar.

3.2 Outcome Variables

The World Bank's Living Standards Measurement Survey (LSMS) program serves as a convenient starting point for retrieving some of the outcome variables of interest. The LSMS program collects nationally representative surveys of households with information on their GPS-based locations, and links these locations to geospatial variables such as distances to other features, climatology, and terrain from various other databases. The latter are important for our empirical strategy, as will be described further in Subsection 3.3 and Section 4.

From the LSMS program, we use the Uganda National Panel Survey (UNPS) from 2013/2014 and the Malawi Integrated Household Panel Survey (IHPS) from 2013. The UNPS 2013/2014 was carried out on 3,119 households from 317 enumeration areas (EA). The IHPS 2013 comprises 4,000 households from 204 EAs. Based on information from these two datasets, we construct outcome variables related to schooling, time spent on household chores, workforce participation, and household expenditures.

In terms of *schooling*, we look at three different outcomes for children (respondents 18 years or younger): namely, the share of girls and boys in the relevant cohorts who (i) have never attended school, (ii) are currently attending primary school, and (iii) are currently attending secondary school. In terms of *time*

spent on household chores, we look at the share of total time that women, men, girls, and boys within the same household respectively spend on (i) collecting firewood and (ii) fetching water. In terms of *workforce participation*, we look at the likelihood of performing paid work and working on the household farm for girls, boys, women, and men, respectively. Finally, in terms of *household expenditures*, we look at the share of total household expenditures spent on food, clothing, and child expenses (related to health and education), and on typical male goods (alcohol and tobacco).

The outcomes related to household decision-making, domestic violence, and women's sexual rights are retrieved from Uganda's and Malawi's most recent Standard DHS. These datasets include, similarly to the LSMS data, geocoded information on the location of the household clusters that participated in the surveys. The Uganda DHS was collected in 2011 and includes information on 11,055 individuals (8,774 women in the age range 15–49 and 2,281 men in the age range 15–54) from 400 different clusters. The Malawi DHS was carried out in 2010 and includes information on 29,050 individuals (22,153 women in the age range 15–49 and 6,897 men in the age range 15–54) from 827 clusters.

Based on the DHS data, we construct three indicator variables for women's participation in *household decision-making*. The first variable takes the value 1 if the woman has responded that she, by herself or jointly with her husband/partner, decides how to spend her own earnings (if she has any). The second variable takes the value 1 if the woman, by herself or jointly with her husband/partner, decides how to spend the husband's/partner's earnings (if he has any), and the third takes the value 1 if the woman, by herself or jointly with her husband/partner, makes decisions about large household purchases. We then combine these three variables into a measure that counts the number of financial decisions the woman participates in. Thus, our measure of participation in household decision-making takes a value between 0 and 3.

In terms of *domestic violence*, we are interested in two sets of outcomes. First, we are interested in the share of women who have experienced domestic violence in the past 12 months, separating between three types of violence: emotional, physical, and sexual violence. Thus, three indicator variables are constructed, one for each type of violence, which take the value 1 if the woman has experienced that type of violence in the past 12 months, and 0 otherwise. This set of questions was only asked to a subset of the women and hence this sample is smaller than for the other DHS variables. The second set of outcomes we are interested in is women's and men's attitudes toward domestic violence. We construct a measure for women and men, respectively counting the number of reasons for which the respondent agrees that a husband is justified in beating his wife. This variable takes a value between 0 and 5 depending on how many of the following reasons for beating one's wife the respondent agrees with: wife burns the food, wife neglects the children, wife goes out without telling the husband, wife argues with the husband, and wife refuses to have sex with husband.

Lastly, we construct two measures reflecting women's views on their own *sexual rights* and two measures of men's views on women's sexual rights. These are based on two questions, which are phrased slightly differently for men and women. For women, the two questions are whether they feel that they can (i) refuse to have sex, and (ii) ask their partner to use contraception. Our measures will take the value 1 if the woman agrees to the statement, and 0 otherwise. For men, the questions are whether they agree that (i) a woman is

justified to refuse sex if her husband has other women, and (ii) a woman is justified to ask for a condom if her husband has any sexually transmitted infections (STI). The two variables based on these questions are also constructed as indicator variables taking the value 1 if the man agrees with the statement, and 0 otherwise.

3.3 Matching Variables

Our empirical strategy, which we will describe in detail in Section 4, is based on matching on observables. These predetermined observables are primarily retrieved from the LSMS data, which provide a wide set of geospatial variables for each LSMS household location, such as average rainfall, onset of greenness, and landcover.[7] Many of these variables are provided as district averages, which means that they can easily be matched also to the DHS cluster locations. For those that are more tied to the exact household location, such as the distance to the closest border post, administrative center, and market place, we collect the original data from the referenced databases in the LSMS and construct the equivalent measures for the DHS clusters.

In addition to the geospatial variables, we construct predetermined socio-economic and needs indicators, such as district averages in educational attainment, women's employment rate, and access to piped water, based on information from the Uganda 2002 census and the Malawi 1998 census.[8] For the outcomes measured at the individual level, we include additional control variables such as ethnicity, religion, and parent's education that are retrieved from the relevant outcome dataset: LSMS or DHS. Notice, in passing, that not all of these variables are actually used as matching factors. The specification of the matching equation is left to a data-driven algorithm that will be described in detail in Section 4, which will select, among all of these exogenous and predetermined variables, the ones with the best predicting power for our two treatments.

3.4 Descriptive Statistics

To understand what the situation looked like before the implementation of the aid projects in our dataset, we have collected data on some of the outcome variables of interest from the Uganda 2002 census, the Malawi 1998 census, and the Uganda DHS and Malawi DHS from 2000. More precisely, we examine the pre-existing gender gaps in primary and secondary schooling at the district level in Uganda, 2002, and Malawi, 1998, for adults (Figure 1) and children aged 6 to 18 (Figure 2). We also show women's and men's pre-existing attitudes toward domestic violence (Figure 3) in each of the countries in 2000. All maps also show the distribution of aid projects upon which our first treatment indicator is based. The purpose of this is to show that there does not appear to be any evident relationship between the pre-existing gender gaps and where aid projects were later implemented.

Figure 1 shows the women-to-men ratio in the share of adults with at least some primary or secondary schooling in Uganda and Malawi. A darker color indicates a higher women-to-men ratio (smaller gap). From

[7] The complete list of variables available in the LSMS datasets and their sources can be found in the LSMS information material, and in particular for Malawi at
http://siteresources.worldbank.org/INTLSMS/Resources/3358986-1233781970982/5800988-1271185595871/6964312-1404828635943/IHPS.Geovariables.Description.pdf and for Uganda at
http://siteresources.worldbank.org/INTLSMS/Resources/3358986-1233781970982/5800988-1265043582346/UNPS_2009_10_BID_rev_2014.pdf.

[8] We use a 10% sample of the Uganda 2002 census and the Malawi 1998 censes, provided from the Integrated Public Use Microdata Series (IPUMS), International.

Figure 1: Pre-Existing Gender Gaps in Educational Attainment (Adults)

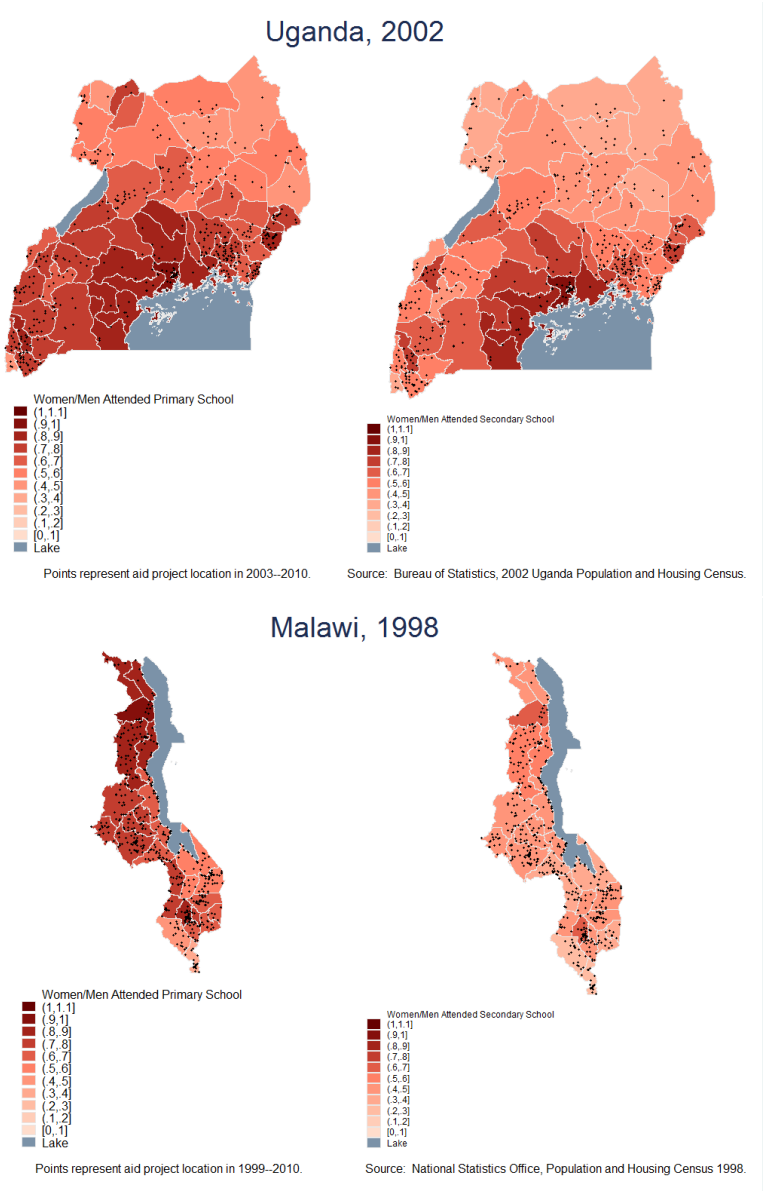
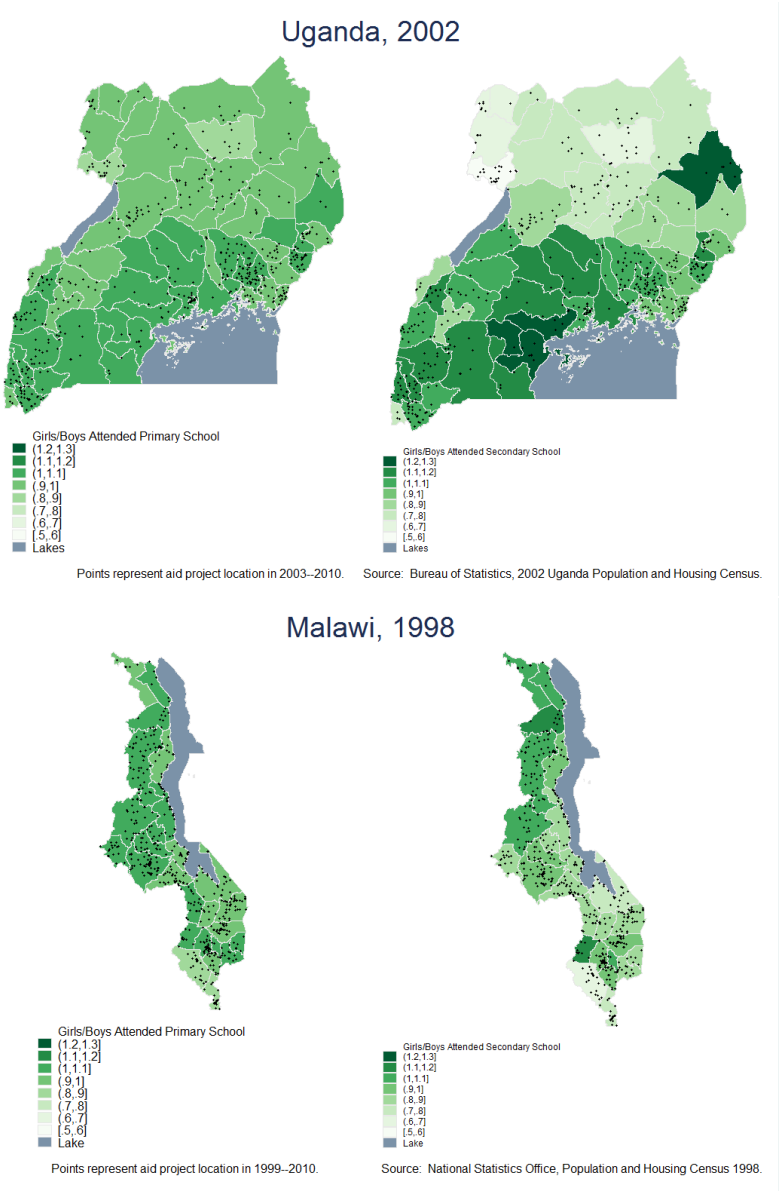


Figure 1 it is clear that gender gaps in educational attainment among adults existed in both countries. Moreover, we see that the gender gaps were larger in terms of secondary schooling rather than primary schooling, and that regional differences existed both within and between the countries. For example, the districts in the Central Region of Uganda had considerably smaller gender gaps than the districts in the Northern Region. In Malawi, regional differences in the women-to-men ratio in primary schooling also existed. However, in terms of secondary schooling, Malawi had only small regional differences: for this outcome, the situation was almost equally bad across the whole country. The overall average share with at least some primary schooling was 57.0% (52.9%) for women in Uganda (Malawi), and 78.0% (74.1%) for men. The equivalent for secondary schooling was 30.2% (15.3%) for women and 49.9% (33.1%) for men.

Figure 2 shows the girls-to-boys ratio in the share of children with at least some primary or secondary school in each district of Uganda and Malawi. The first thing to notice from these maps is that almost

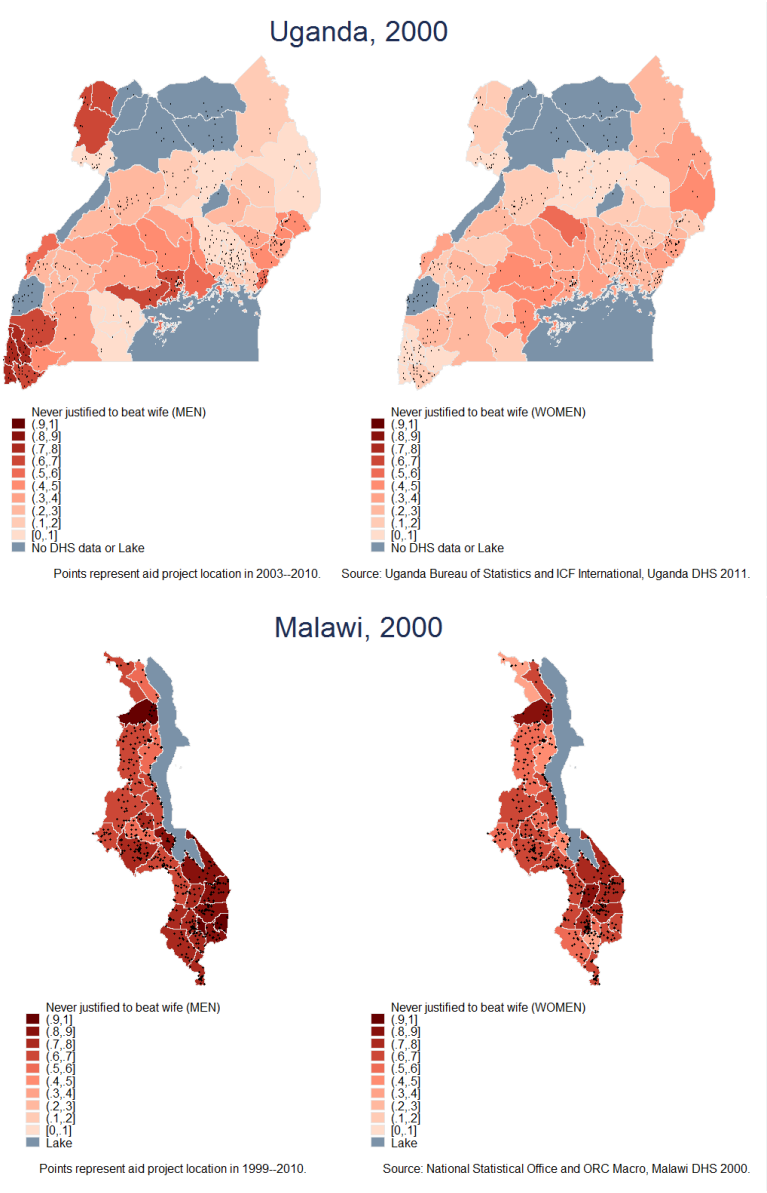
Figure 2: Pre-Existing Gender Gaps in Educational Attainment (Children)



no regional differences or overall gender gap in primary schooling remained for this generation. In some districts, girls were even more likely to attend primary school than boys. In terms of secondary schooling, gender gaps still existed and the size of the gaps varied across districts, particularly in Uganda. Moreover, while the gender gap in primary and secondary schooling was closing, there was still room for improvement in the overall attendance rate. In Uganda (Malawi) average attendance rate in primary schooling was 78.4% (79.5%) for girls and 79.6% (79.9%) for boys. The average attendance rate in secondary schooling was 42.1% (16.8%) for girls in Uganda (Malawi), and 42.6% (17.7%) for boys.

Figure 3 shows the share of men and women who agree with the statement that it is not justified for a husband to beat his wife for any of the following five reasons: burning the food, neglecting the children, going out without telling the husband, arguing with the husband, and refusing to have sex with husband. First, we see that the Uganda DHS 2000 was not carried out in all districts, hence the missing data. Second, we

Figure 3: Pre-Existing Gender Gaps in Attitudes toward Wife Beating (Adults)



notice that the share of men and women who agree that beating one's wife is never justified is much lower in Uganda than in Malawi, especially among women. Thus, the tolerance of domestic violence was much higher in Uganda compared to Malawi. The average share of women in Uganda agreeing that it is never justified to beat one's wife was only 23.4% compared to 75.0% in Malawi. Similarly for men, the average in Uganda was 34.7% compared to 61.3% in Malawi.

The main take-aways from these figures are the following. First, given the lack of large pre-existing gender gaps in the primary and secondary school attendance rates for children, we do not expect to see any strong differential effect of aid on girls compared to boys. In the area of attitudes towards domestic violence, there is room for improvements, especially in Uganda, where the tolerance was in general high, and particularly among women.

Lastly, from the map it emerges that the aid projects implemented in 2003–2011 in Uganda and 1999–2010

in Malawi do not appear to have been specifically targeted to areas with larger pre-existing gender gaps. We test this hypothesis by applying t-tests testing for equality the district averages in districts with a below-median number of aid project-years against districts with an above-median number of aid project-years. The results can be found in Table 8 in the Appendix. The only statistically significant difference is found in the number of reasons for beating one's wife that women agree with. Aid intensity appears to have been higher in the Ugandan districts where women had initially a higher tolerance for domestic violence. In Malawi, on the other hand, the opposite appears to have been the case. Overall, we find no strong relationship between pre-existing gender gaps and later aid project locations.

4 Method

The main methodological challenge lies in the identification of the causal effect of aid. It is possible that donors seek out areas where the local population have certain attitudes or are more responsive to new opportunities, either deliberately to increase the chances of a successful project or because this correlates with other factors that go into project placement decisions. The opposite may be true if donors are targeting particularly poor or isolated communities. Our proposed strategy draws on recent developments in the literature on matching (Imbens and Rubin, 2015) to offer an estimate of the treatment effect as robust as possible, although based on a conditional independence assumption. This strategy has four key ingredients. First, we exploit the rich set of geospatial variables in the LSMS and add several predetermined "need" indicators from the 2002 and 1998 censuses, described in the previous section, in order to provide a large set of presumably exogenous covariates. Second, we carefully assess the overlap in the distribution of covariates between treated and control units and improve it as much as possible through iterative trimming of the sample. Third, we let the data determine which covariates provide the best model for the conditional expectations through a lasso algorithm (Tibshirani, 1996; Belloni et al., 2012). Fourth, we choose a blocking estimator based on averaging the OLS coefficients separately estimated in a number of data-defined homogeneous subsamples, or *blocks*. The within-block estimate does not rely as much on extrapolation as it would in the full sample since the covariates are well balanced within the block. This whole process reduces one type of bias in the OLS estimation, related to functional form assumptions and extrapolation across potentially very different covariate distributions. The resulting estimator has better small-sample properties and is robust to the difference in covariates distribution by treatment status (Imbens, 2015).

The first part of the analysis only uses the treatment indicators and the covariates. The aim is to assess the overlap in the distribution of covariates between treated and control units. Following Imbens and Rubin (2015) and Imbens (2015), the extent of overlap is assessed using normalized differences, rather than t-statistics,[9] defined as follows:

$$\Delta_X = \frac{\overline{X}_T - \overline{X}_C}{\sqrt{\frac{(S_T^2 + S_C^2)}{2}}} \tag{1}$$

[9] T-test p-values are used for indicator variables.

where X is an element of the covariates vector and $\overline{X}_i$ and S_i^2 are the mean and variance of X in the treatment and control groups, for $i = T, C$ respectively. Large values of these normalized differences indicate substantial differences between the treatment and control groups. In this case, it might be very difficult or impossible to obtain estimates that are robust to different ways of estimating the counterfactual outcome, and one might wish to focus instead on a subset of observations that are more similar, excluding observations that have no good counterpart in the alternate treatment status.

The covariates to be used in the matching will be determined by a data-driven algorithm.[10] We therefore start off, as mentioned above, using all the geographic and climate-related variables from the LSMS and some predetermined "need" indicators from the 2002 and 1998 census for, respectively, Uganda and Malawi. These are the 44 variables shown in Table 1. We also add some contemporaneous covariates from the LSMS that we think might affect the outcomes, but are not very likely to be affected by the treatment in the short run, such as the household size, the number of daughters and sons, household poverty status, and the education level of the mother and father in the household. As we see, the normalized differences are quite large for a substantial number of covariates, indicating that the overlap between treated and control households is not very good.

Trimming the sample is a way to improve the overlap between the covariates distribution for the treated and control units, i.e. to make the two groups more similar. A trimming rule that works well with multiple covariates, and is not sensitive to outliers, is based on propensity scores, and suggests the exclusion of observations with extreme values, i.e. close to 0 or 1 (Imbens, 2015). This requires the estimation of the propensity scores. As mentioned above, we take the most agnostic approach and let the data determine which covariates provide the best model for the conditional expectations. The selection of covariates to be included in the estimation of the propensity scores is left to a lasso[11] algorithm (Tibshirani, 1996; Belloni et al., 2012). The list of chosen factors and the corresponding parameter estimates for the propensity score equation are presented in Table 2. The first two columns refer to the treatment defined by exposure to any aid project, while the third and fourth columns refer to exposure to gender projects.

Propensity scores are estimated with two alternative linking functions. The correlation is very close to 1. The probit model gives a lower log-likelihood value and slightly less density at the extremes of the propensity scores distribution compared to the logit model, so we proceed with this model. The observations that are closer to the extremes than threshold distance of 0.1,[12] i.e. with a propensity score lower than 0.1 or higher than 0.9, are trimmed from the sample, in order to improve the covariate overlap. We repeat the lasso in the smaller sample to insure the same factors are selected for inclusion, and re-estimate the propensity

[10] This stage has no structural or causal interpretation. Moreover, as already emphasized, data on outcomes are not used at this point. Specification search is therefore acceptable, as noted also in Imbens and Rubin (2015), as "the role of these covariates is purely mechanical in balancing the two samples and approximate accurately the conditional expectation". Potential role for theory at this stage is purely supportive, if the researcher has relevant information guiding the choice of factors.

[11] Lasso is one of the best-known techniques to shrink regression coefficients and by this means select among a large number of potential predictors. Other related methods are ridge regression and subset selection through stepwise least squares regressions. The advantages of lasso are, compared to the first one, that it actually restricts some of the coefficients all the way to 0, and compared to the second one, that it is more likely to identify the global, rather than a local, optimum.

[12] An optimal threshold distance α in terms of efficiency of the estimator can be derived from the data (Crump et al., 2008). However, their simulations suggest that in most cases a default $\alpha = 0.1$ provides a good approximation.

scores (potentially also repeating the trimming iteratively until we reach satisfactory overlap, judged by visual inspection and normalized differences). Table 4 shows the distribution of observations falling above and below the threshold distance from the extremes of the propensity scores distribution, that are therefore trimmed. The plots in Figure 4 show the density of propensity scores by treatment status before and after trimming the sample. They reveal that our model is very good at predicting successes, i.e. correctly classifying treated households, but not as good at identifying non-treated ones. However, the probability of success is clearly lower for non-treated observations, and the interval of overlap between the two samples is substantial (after trimming), which is what matching estimation requires. Table 3 shows that the overlap between covariates improves in this smaller sample also in terms of normalized differences, as most fall now below 40% (see Column 5; Column 6 reports as a reference the last column of Table 1).

When it comes to the DHS data, we do not have all of the geographic variables corresponding to the clusters' location. For this sample we proceed as follows. We assume that the geographic predictors of aid presence are the same as for the LSMS households' locations. Starting from this restricted set of predictors, we improve as much as we can the model for the propensity scores by adding a set of individual demographic characteristics, considering that all the outcomes in this case are at the individual rather than the household level. The list of variables and their summary statistics are reported in Table 5. After trimming the sample according to the same rule as above (sample sizes reported in Table 6), the normalized differences are very small. The parameter estimates for the final specification of the propensity score equation are presented in Table 7, where again the first two columns refer to exposure to any aid projects while the third and fourth columns refer to exposure to gender-specific aid projects.

We perform the same covariate-overlap analysis and repeated trimming for the second treatment, i.e. exposure to gender-specific aid.[13] The comparison group for this treatment is the same, namely households not exposed to any aid projects. Interestingly, our model is much better at classifying both successes and failures with this treatment, likely because there is a larger number of zeroes ("untreated neighborhoods"). However, the fact that there are far fewer gender-specific projects and that these were implemented relatively more recently reduces the number of observations for these estimations.

As mentioned, the parameter estimates for the final specification of the propensity score equation are presented in Columns 3 and 4 of Table 2 for the LSMS data and Table 7 for the DHS data.

For the estimation of the treatment effect, we use a blocking estimator proposed by Imbens (2015). This consists in essence in partitioning the data into smaller subsamples, or *blocks*, that are more homogeneous in terms of covariates distribution compared to the whole sample. The treatment effect is estimated using OLS within each of these blocks, and then averaged across blocks with weights taking into account the block composition in terms of treated and control units. The division into blocks is completely data-driven, according to an algorithm developed in Imbens and Rubin (2015).[14] The linear regression in each of the blocks does not rely as much on extrapolation as it would in the full sample, since the covariates are well

[13] Tables and plots are not reported but can be requested from the authors.

[14] Researchers have often used five subsamples with an equal number of units. This rule was, however, developed for the case of a single normally distributed covariate (Cochrane, 1968) and should not be considered generally applicable.

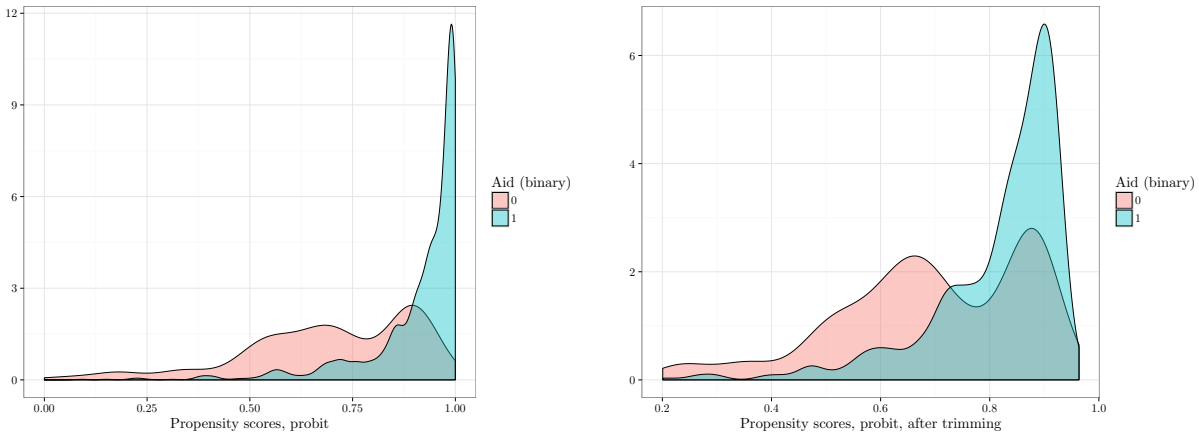


Figure 4: Covariates Overlap, LSMS Sample

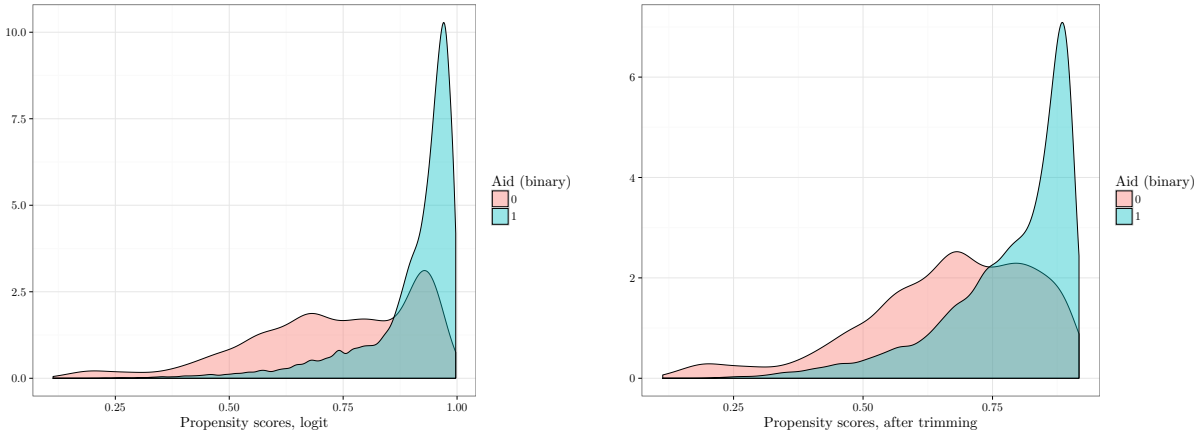


Figure 5: Covariates Overlap, DHS Sample

balanced within the block.

While the propensity score, and hence the division into blocks, is based on a binary definition of the treatment, namely exposure, or lack thereof, to any number of (gender-specific) aid projects, we know that the number and duration of these projects varies substantially in the sample. In order to capture the variation in the intensity of treatment, we estimate the following equation within each block:

$$Y = \alpha + \beta * AidPrYr + \mathbf{M}'\gamma + \delta * \overline{PrYr}_d + \epsilon \tag{2}$$

where the variable of interest is the number of (gender) aid project-years that the household or individual is exposed to. The estimated effect β can thus be interpreted as the impact of the marginal project-year. Tables 10–18 report, along with β, the impact size evaluated at the mean project-years of exposure for the treated units. The regressions also include and control for all the matching factors (the matrix $\mathbf{M}$ in equation 2) plus the number of (gender) project-years implemented at the district level.

Given our broad approach, looking for impact in a multifaceted definition of "gender gaps" through many outcomes, we face an issue of multiple hypothesis testing. When testing separately a large number of hypotheses, there is some probability of observing a few significant results just by chance even if all of the tests are in fact not significant. A standard, and quite straightforward, solution to this issue is the Bonferroni method, which tests the significance of individual coefficients viewed as part of a family of hypotheses, by simply using a critical value for statistical significance of α/n rather than α, where n is the number of hypotheses. This correction assumes, however, that the outcomes are independent. Since many of our outcomes are very much correlated or even interdependent, we follow Sankoh et al. (1997) and adjust the Bonferroni correction for correlation.[15] The families of outcomes that we adjust in this way are: (i) market and household work by different household members ($n = 12$); (ii) household expenditure decisions ($n = 4$); (iii) domestic violence ($n = 3$).[16]

5 Results

In this section we illustrate the marginal effects of aid-financed project-years with a visual presentation. All the corresponding tables are in the Appendix (Tables 9–18).

As already mentioned, we consider separately the effect of two "treatments", the exposure to any type of aid projects within 15km of the household implemented during the previous ten years, and the exposure to gender-targeted aid projects within the same geographical and temporal range. The control group is the same in both cases: namely, households that were not in the neighborhood of any aid project during this time span.

The plots report estimated effects and confidence intervals together with the number of observations for every outcome. This varies according to the survey response rate. Since these results are obtained through

[15] The confidence intervals reported in the figures and tables do not reflect this correction; however, the p-values in the tables are the correct ones.

[16] We do not adjust the group of outcomes in which no test is significant, for example schooling outcomes.

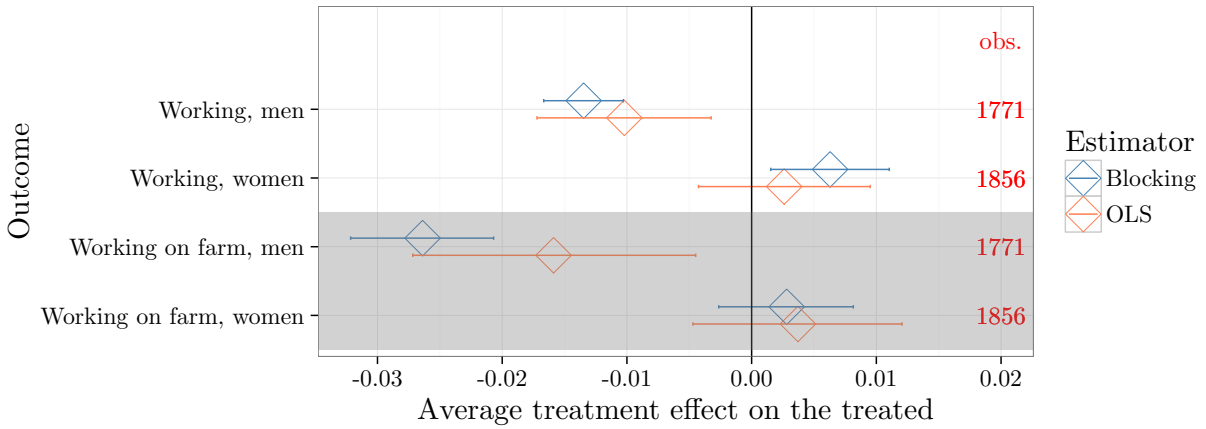


Figure 6: Adults' Outcomes: Marginal Effect of Aid Project-Years

the blocking estimator, defined in the previous section, the tables corresponding to each plot also report the number of blocks that were used in the estimation, which also varies from outcome to outcome. The tables also report the p-value of the test adjusted for the multiple hypotheses problem, the baseline mean of the outcome, and the size of the impact evaluated at the mean exposure (in project-years).

Aid presence affects labor market participation differently for women and men, increasing women's and decreasing men's probability of working. Men are also less likely to work on the family farm. However, no effects are associated with gender-specific aid (see Figures 6 and 7). The impact might look small, but evaluated at the mean project-years (which is around 30), it corresponds to an increase of more than 18% in the probability of having performed some paid work during the previous year for women (see Table 10).

Perhaps as a result of the improved relative standing, in economic terms, of the woman in the household, an overall positive pattern emerges from Figure 8 and Figure 9, investigating attitudes and experiences in the areas of sexual freedoms and domestic violence. Interviewed men and women in the areas exposed to aid are no less likely to justify beating one's wife, or are even slightly more likely when it comes to men exposed to gender-specific projects. However, the women interviewed are less likely to have experienced physical or sexual violence. Men express to a larger extent support for the idea that wives can refuse sex or demand protected sex, at least in the somewhat extreme cases of a husband's infidelity or sexual transmittable diseases. However the impact on the actual ability of women to, in practice, choose these behaviors is more modest. These findings are consistent with our expectations and with evidence from previous studies, which identify the improved bargaining position of the woman in the household as a channel, as highlighted in Section 2.

Comparing Figure 10 and Figure 11, we conclude that neither the presence of aid nor the presence of gender-specific aid have large impacts on the participation of women in the household's financial decisions. Furthermore, the clearest impact of the improved economic opportunities presented by aid seems to be an increase in the share of the household's budget spent on what are traditionally viewed as male goods, namely alcohol and tobacco. This increase is stronger but not statistically significant in the case of general aid, although the likelihood of men working is reduced in this case (or maybe because of this). The increase in

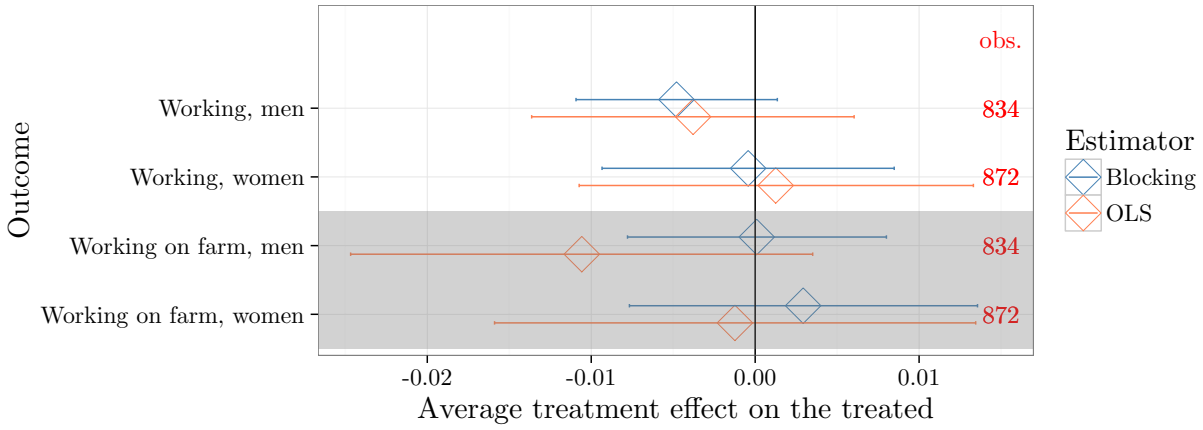


Figure 7: Adults' Outcomes: Marginal Effect of Gender Project-Years

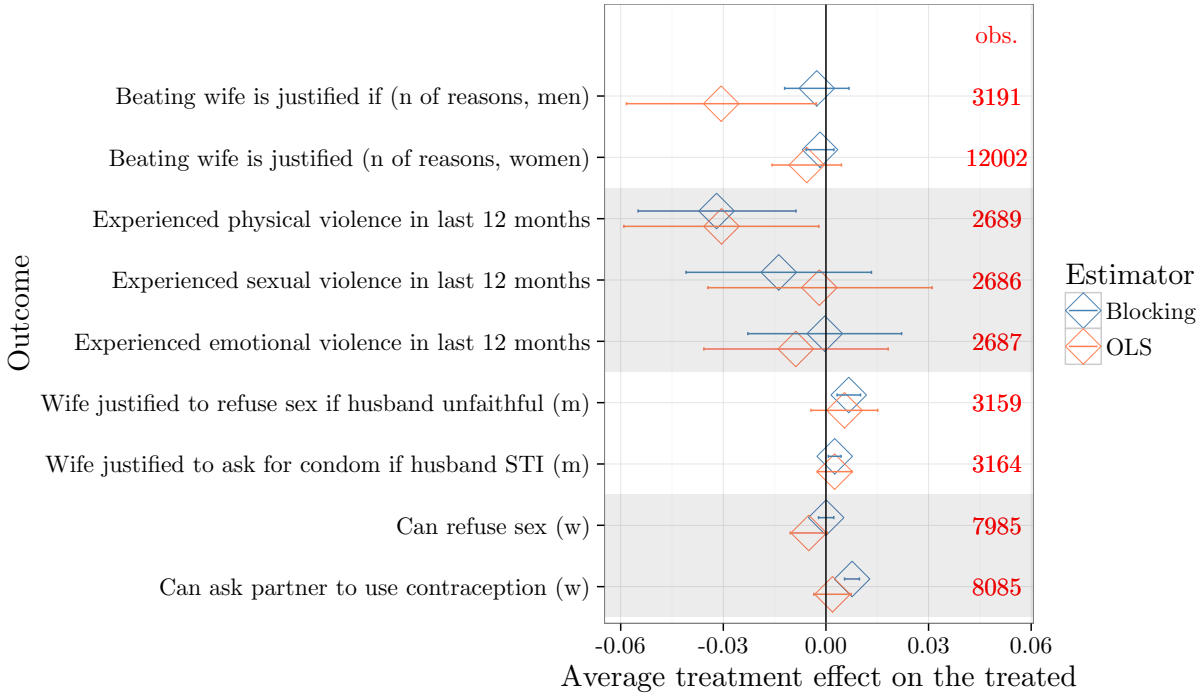


Figure 8: Domestic Violence and Sexual Rights: Marginal Effect of Aid Project-Years

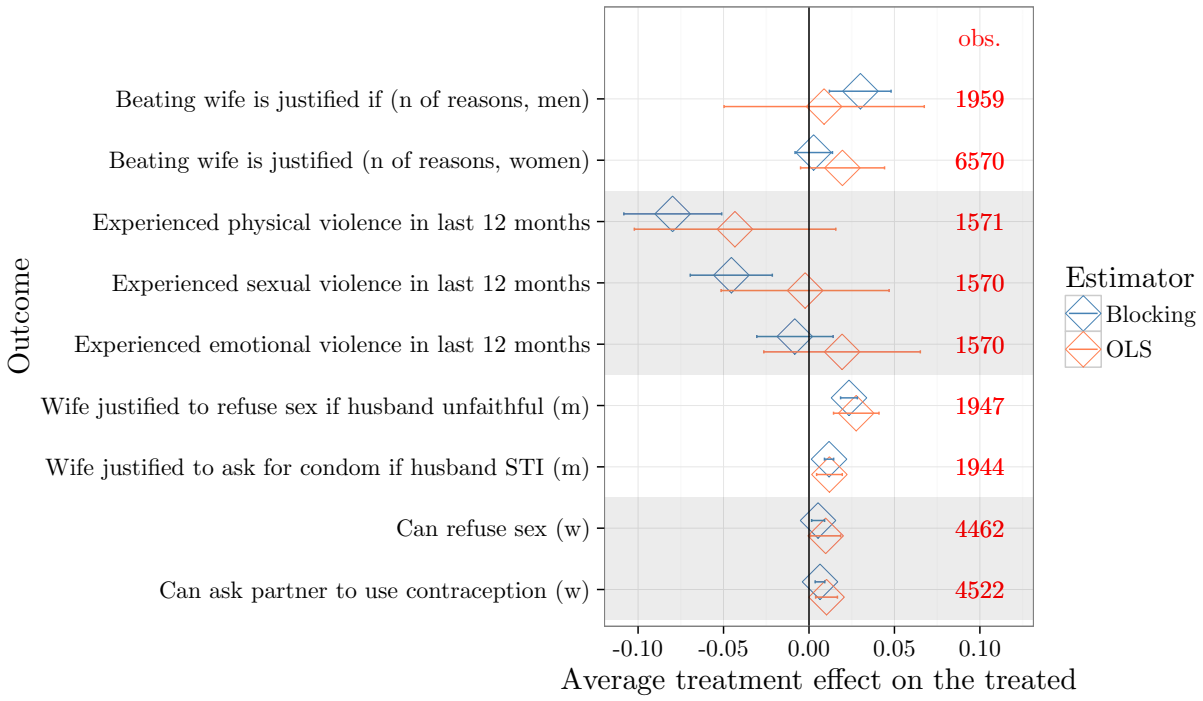


Figure 9: Domestic Violence and Sexual Freedoms: Marginal Effect of Gender Project-Years

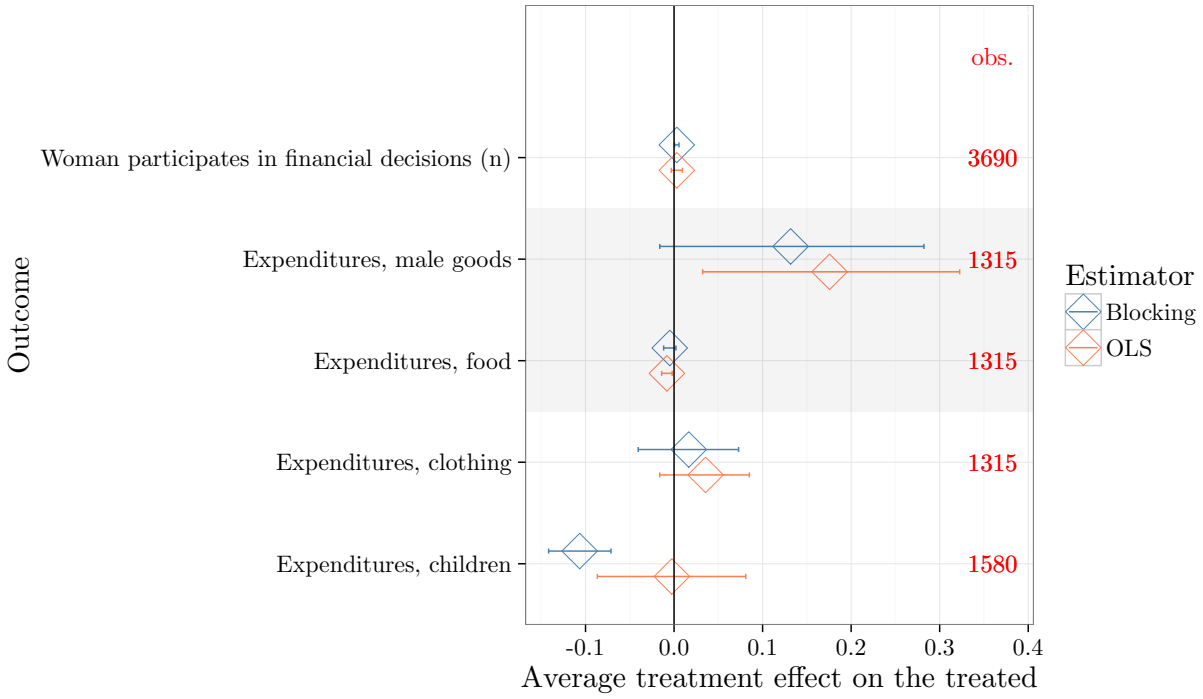


Figure 10: Household Expenditure Decisions: Marginal Effect of Aid Project-Years

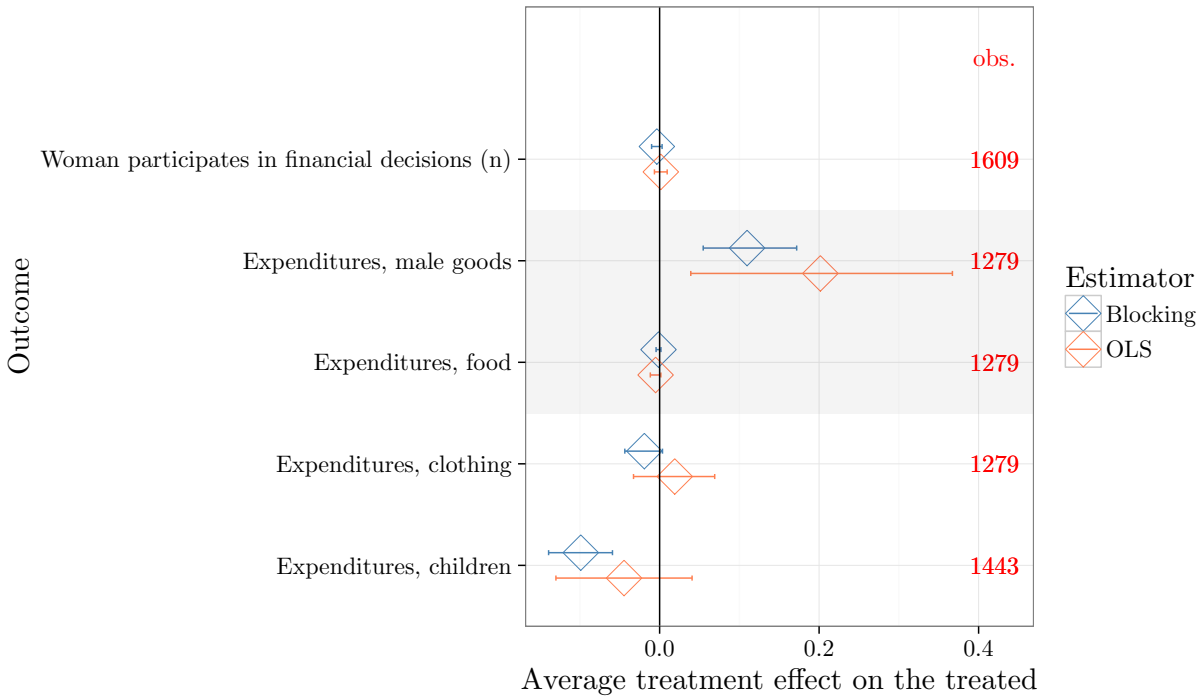


Figure 11: Household Expenditure Decisions: Marginal Effect of Gender Project-Years

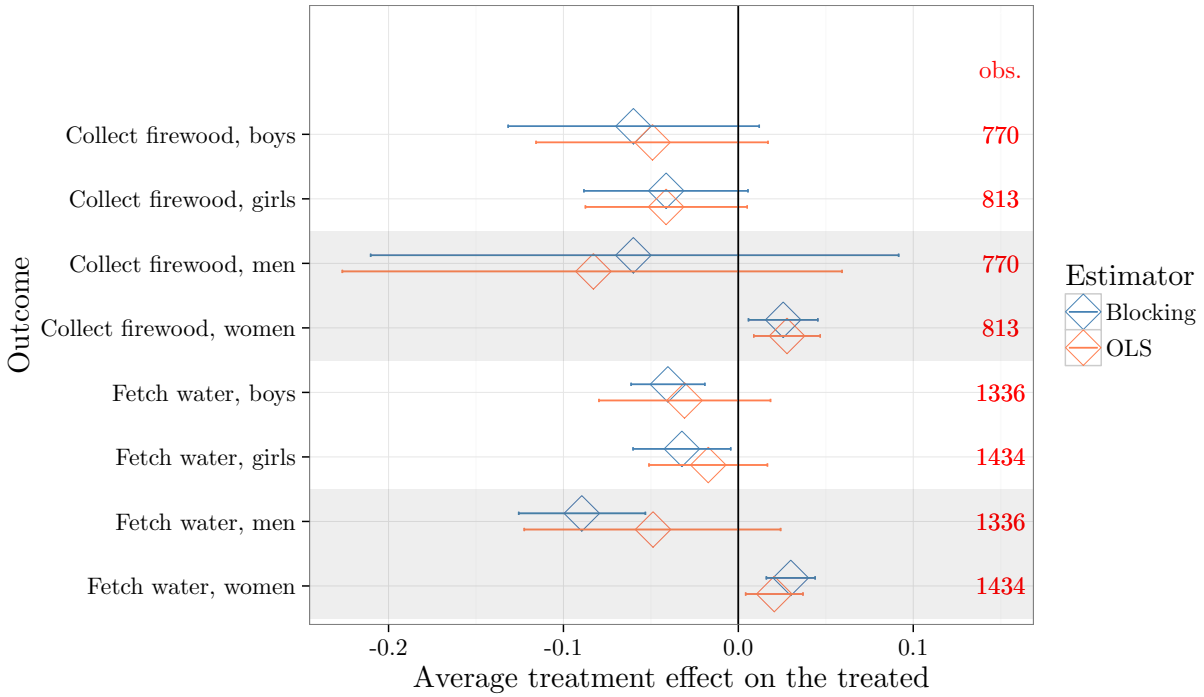


Figure 12: Domestic Chores: Marginal Effect of Aid Project-Years

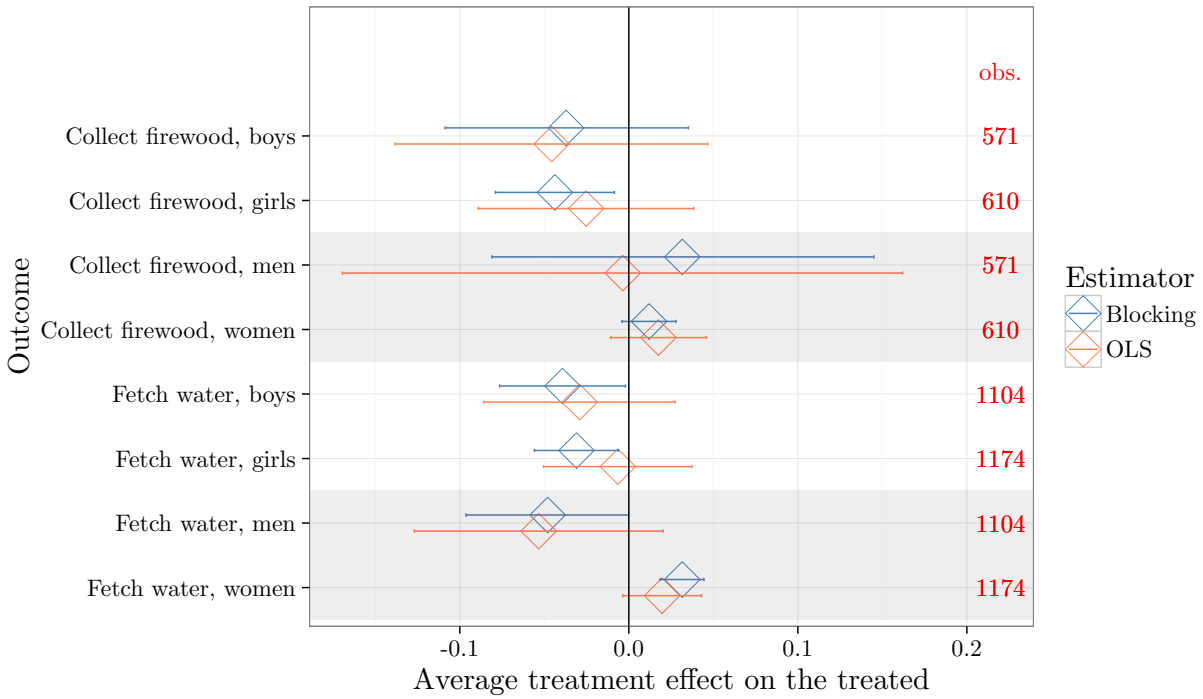


Figure 13: Domestic Chores: Marginal Effect of Gender Project-Years

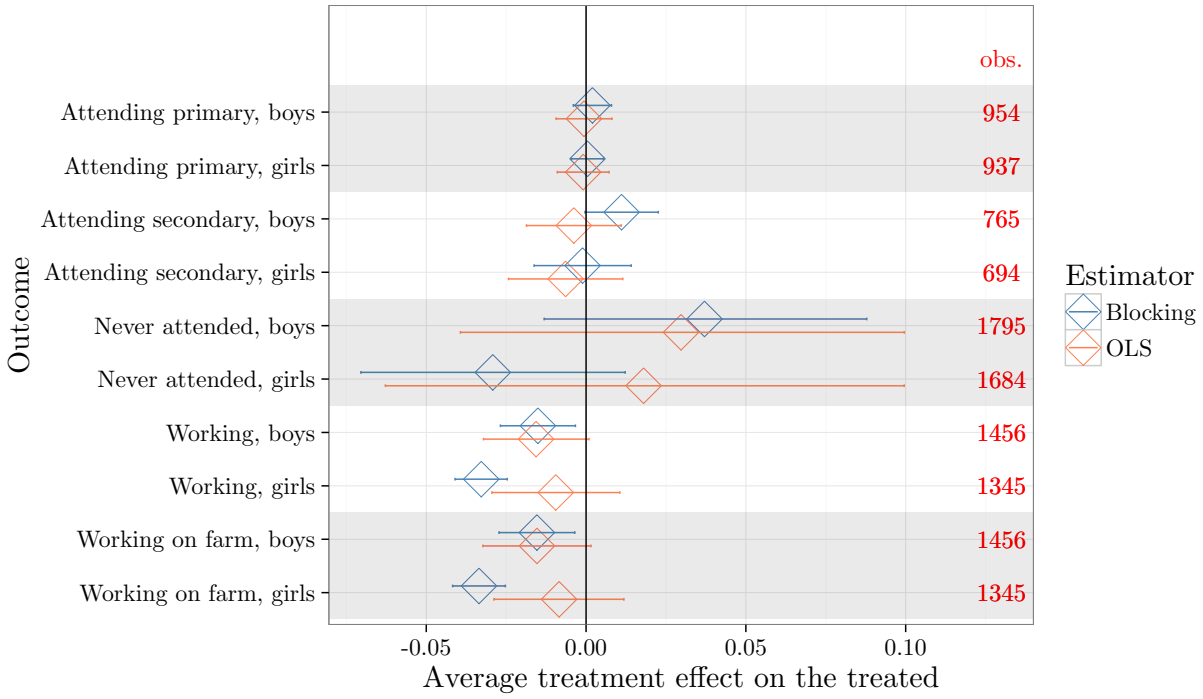


Figure 14: Children's Outcomes: Marginal Effect of Aid Project-Years

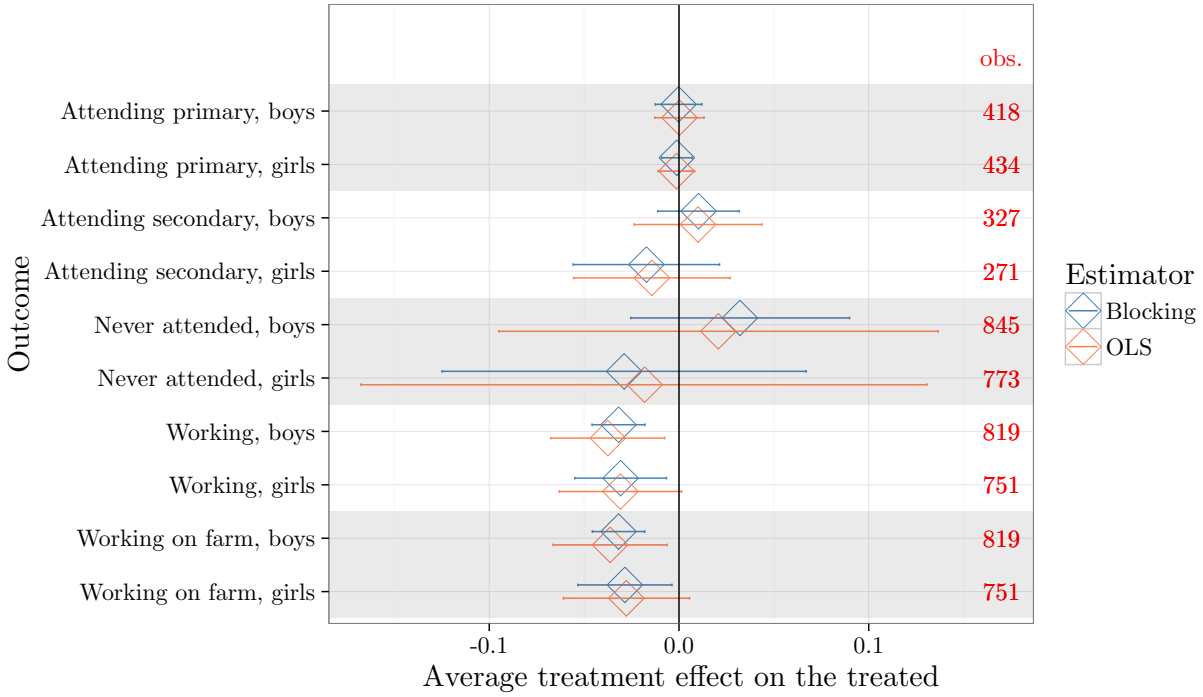


Figure 15: Children's Outcomes: Marginal Effect of Gender Project-Years

the share of the budget spent on male goods happens partly at the expense of spending on children (health and education). The other expenditure categories considered are hardly affected.

The presence of aid seems to affect also the division of domestic labor between children and adults, and men and women. Though more likely to work for pay, women still undertake a larger share of (typically female) household chores such as collecting firewood and fetching water. The small impact shown in Figures 12 and 13 corresponds to an increase of slightly more than half an hour (during the previous week) spent fetching water and around 20 minutes collecting firewood. Results in Figures 14 and 15 show that aid projects do not significantly affect schooling decisions. However, children, and in particular girls, are less likely to work, both outside the household and on the family farm. Again the impacts are larger, and even larger for girls, in the presence of aid in general, when women are more likely to work.

All in all, the picture painted by the results is consistent with the framework outlined in Section 2. The impact of aid seems to be clearest on outcomes related to women's access to economic opportunities. However, the spillover from this main impact to other important outcomes for women and girls is not overwhelmingly positive across the board, although we can't exclude small signs of improvements in some areas.

6 Conclusions

Until recently, studies of the impact of foreign aid have typically taken one of two forms: on the one hand, cross-country comparisons of the impact of aggregate aid inflows on countrywide indicators such as growth, savings, or level of democracy; on the other, micro-level impact evaluations of the effect on immediate beneficiaries from a particular aid-financed intervention. The results from these studies are often contradictory,

as suggested by the so-called micro–macro paradox. The newly emerging geocoded data on the history of aid project locations have given rise to the opportunity to find a middle ground: looking at community-level effects in the vicinity of aid projects, including both direct beneficiaries and others who may, or may not, be affected indirectly. In this paper, we have applied this methodology to better understand the impact of foreign-aid-financed activities on gender equality, an important objective of most Western donors that, beyond impact evaluations of individual aid-financed interventions, has gotten surprisingly little attention in the economics literature.

Based on an intra-household bargaining framework we look at outcomes across several dimensions. We find that the presence of aid projects generates paid job opportunities outside the household that strengthen women's outside options, thereby potentially also strengthening their control over other areas influenced by relative bargaining power. However, in this aspect the results are more mixed. Attitudes towards female sexual empowerment seem more positive and exposure to domestic violence lower, but we find no impact on the perceived influence over household revenues, and adult women (if anything) take responsibility for a greater share of household time spent on key domestic chores such as fetching water and collecting firewood. We also find no effect on girls' absolute or relative schooling, but both boys and girls are less likely to work both outside the household and on the farm. We believe the diversity in outcomes reflects in a reasonable way the variety in challenges faced across different areas depending on differences between attitudes and actual behavior, the strength of existing gender norms, and what is required of men to give up. Overall the outcomes lend some ground for optimism concerning the potential for aid to facilitate a process towards greater gender equality, but also clearly point to the challenges and the need for realistic expectations. It should also be emphasized that, as with all methods, our approach has strengths and weaknesses, and it is important to understand that the impact of specific interventions on direct beneficiaries may indeed be stronger. The focus here is on the extent to which such interventions can spread beyond that and have a more lasting impact on the surrounding community.

We see several avenues for further elaboration of this research. The AidData Research Consortium is continuously updating its data, completing the donor footprint for more and more countries. It should thus hopefully be possible to substantially expand the current research data-wise in the not too far future, covering a much larger share of the African continent. The method can also be applied to other outcomes, for instance household-level consumption or earnings, or attitudes towards traditional authority and democracy.

References

Agénor, P.-R., O. Canuto, and L. P. da Silva (2014). On gender and growth: The role of intergenerational health externalities and women's occupational constraints. *Structural Change and Economic Dynamics 30*, 132–147.

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

Akresh, R., D. de Walque, and H. Kazianga (2016). Evidence from a randomized evaluation of the household welfare impacts of conditional and unconditional cash transfers given to mothers or fathers. Oklahoma State University, Department of Economics and Legal Studies in Business.

Allcott, H. (2015). Site selection bias in program evaluation. *The Quarterly Journal of Economics 130*(3), 1117–1165.

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

Andrabi, T. and J. Das (2017). In aid we trust: Hearts and minds and the Pakistan earthquake of 2005. *The Review of Economics and Statistics*. Forthcoming.

Attanasio, O. and V. Lechene (2002). Tests of income pooling in household decisions. *Review of Economic Dynamics 5*(4), 720–748.

Beaman, L., R. Chattopadhyay, E. Duflo, R. Pande, and P. Topalova (2009). Powerful women: Female leadership and gender bias. *Quarterly Journal of Economics 124*(4), 1497–1540.

Beaman, L., E. Duflo, R. Pande, and P. Topalova (2012). Female leadership raises aspirations and educational attainment for girls: A policy experiment in india. *Science 335*(6068), 582–586.

Belloni, A., D. Chen, V. Chernozhukov, and C. Hansen (2012). Sparse models and methods for optimal instruments with an application to eminent domain. *Econometrica 80*(6), 2369–2429.

Benhassine, N., F. Devoto, E. Duflo, P. Dupas, and V. Pouliquen (2015). Turning a shove into a nudge? A "labeled cash transfer" for education. *American Economic Journal: Economic Policy 7*(3), 86–125.

BenYishay, A. and A. M. Mobarak (2015). Social learning and incentives for experimentation and communication. *Review of Economic Studies. Forthcoming*.

Blackden, C. M. and Q. Wodon (2006). *Gender, Time Use, and Poverty in Sub-Saharan Africa*, Volume 73. Washington, DC: World Bank Publications.

Bobonis, G. J., M. González-Brenes, R. Castro, et al. (2013). Public transfers and domestic violence: The roles of private information and spousal control. *American Economic Journal: Economic Policy 5*(1), 179–205.

Cai, J., A. De Janvry, and E. Sadoulet (2015). Social networks and the decision to insure. *American Economic Journal: Applied Economics 7*(2), 81–108.

Case, A. and A. Deaton (1998). Large cash transfers to the elderly in South Africa. *The Economic Journal 108*(450), 1330–1361.

Chaaban, J. and W. Cunningham (2011). Measuring the economic gain of investing in girls: The girl effect dividend. World Bank Policy Research Working Paper Series 5753. Washington, DC: World Bank.

Chakravarty, S., Z. Iqbal, and A. Z. Shahriar (2013). Are women naturally better credit risks in microcredit? evidence from field experiments in patriarchal and matrilineal societies in Bangladesh. Technical report. West Lafayette, IN: Purdue University, Department of Consumer Sciences.

Cochrane, W. (1968). The effectiveness of adjustment by subclassification in removing bias in observational studies. *Biometrics 24*(2), 295–313.

Conley, T. G. and C. R. Udry (2010). Learning about a new technology: Pineapple in Ghana. *The American Economic Review 100*(1), 35–69.

Crump, R. K., V. J. Hotz, G. W. Imbens, and O. A. Mitnik (2008). Nonparametric tests for treatment effect heterogeneity. *The Review of Economics and Statistics 90*(3), 389–405.

De, R. and C. Becker (2015). The foreign aid effectiveness debate: Evidence from malawi. AidData Working Paper 6. Williamsburg, VA: College of William and Mary.

Dreher, A. and S. Lohmann (2015). Aid and growth at the regional level. *Oxford Review of Economic Policy 31*(3-4), 420–446.

Duflo, E. (2012). Women empowerment and economic development. *Journal of Economic Literature 50*(4), 1051–79.

Findley, M. G., J. Powell, D. Strandow, and J. Tanner (2011). The localized geography of foreign aid: A new dataset and application to violent armed conflict. *World Development 39*(11), 1995–2009.

Friedberg, L. and A. Webb (2005). The chore wars: Household bargaining and leisure time. Unpublished manuscript. Charlottesville, VA: University of Virginia.

Gakidou, E., K. Cowling, R. Lozano, and C. J. Murray (2010). Increased educational attainment and its effect on child mortality in 175 countries between 1970 and 2009: A systematic analysis. *The Lancet 376*(9745), 959–974.

Galiani, S., S. Knack, L. C. Xu, and B. Zou (2017). The effect of aid on growth: Evidence from a quasi-experiment. *Journal of Economic Growth 22*(1), 1–33.

Hasan, A. et al. (2010). Time allocation in rural households: the indirect effects of conditional cash transfer programs. Policy Research Working Paper 5256. Washington, DC: World Bank.

Imbens, G. W. (2015). Matching methods in practice: Three examples. *Journal of Human Resources 50*(2), 373–419.

Imbens, G. W. and D. B. Rubin (2015, May). *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction* (1st ed.). New York: Cambridge University Press.

Isaksson, A.-S. and A. Kotsadam (2016). Chinese aid and local corruption. AidData Working Paper 33. Williamsburg, VA: College of William and Mary.

Iyer, L., A. Mani, P. Mishra, and P. Topalova (2012). The power of political voice: Women's political representation and crime in india. *American Economic Journal: Applied Economics 4*(4), 165–193.

Kabunga, N. S., T. Mogues, E. Bizimungu, A. Erman, and B. Van Campenhout (2016). *The State of Public Service Delivery in Uganda: Report of a baseline survey.* Washington, DC: International Food Policy Research Institute.

Kim, J. C., C. H. Watts, J. R. Hargreaves, L. X. Ndhlovu, G. Phetla, L. A. Morison, J. Busza, J. D. Porter, and P. Pronyk (2007). Understanding the impact of a microfinance-based intervention on women's empowerment and the reduction of intimate partner violence in south africa. *American Journal of Public Health 97*(10), 1794–1802.

Koenig, M. A., S. Ahmed, M. B. Hossain, and A. K. A. Mozumder (2003). Women's status and domestic violence in rural Bangladesh: Individual-and community-level effects. *Demography 40*(2), 269–288.

Lundberg, S. J., R. A. Pollak, and T. J. Wales (1997). Do husbands and wives pool their resources? Evidence from the United Kingdom Child Benefit. *Journal of Human Resources 32*(3), 463–481.

Maxwell, N. L. and N. Wozny (2017). Gender gaps in time use and earnings: What's norms got to do with it? Working Paper 54. Oakland, CA: Mathematica Policy Research.

McCrary, J. and H. Royer (2011). The effect of female education on fertility and infant health: Evidence from school entry policies using exact date of birth. *The American Economic Review 101*(1), 158–195.

Miguel, E. and M. Kremer (2004). Worms: Identifying impacts on education and health in the presence of treatment externalities. *Econometrica 72*(1), 159–217.

Munshi, K. and J. Myaux (2006). Social norms and the fertility transition. *Journal of Development Economics 80*(1), 1–38.

Nanivazo, M. and L. Scott (2012). Gender mainstreaming in Nordic development agencies: Seventeen years after the beijing conference. WIDER Working Paper 2012/91. Helsinki: UNU-WIDER.

Peratsakis, C., J. Powell, M. Findley, J. Baker, and C. Weaver (2012). Geocoded activity-level data from the government of malawis aid management platform. Washington, DC: AidData and the Robert S. Strauss Center for International Security and Law.

Pickbourn, L. and L. Ndikumana (2016). The impact of the sectoral allocation of foreign aid on gender inequality. *Journal of International Development 28*(3), 396–411.

Pulerwitz, J., L. Hughes, M. Mehta, A. Kidanu, F. Verani, and S. Tewolde (2015). Changing gender norms and reducing intimate partner violence: Results from a quasi-experimental intervention study with young men in Ethiopia. *American Journal of Public Health 105*(1), 132–137.

Rajan, R. G. and A. Subramanian (2008). Aid and growth: What does the cross-country evidence really show? *The Review of Economics and Statistics 90*(4), 643–665.

Ravishankar, V., S. E.-T. El-Kogali, D. Sankar, N. Tanaka, and N. Rakoto-Tiana (2016). *Primary Education in Malawi: Expenditures, service delivery, and outcomes*. Washington, DC: World Bank Publications.

Richey, L. A. (2003). Gender Equality and Foreign Aid. In F. Tarp (ed.), *Foreign Aid and Development. Lessons Learnt and Directions for the Future*, pp. 247–270. London and New York: Routledge.

Ringdal, C. and I. Hoem Sjursen (2017). Household bargaining and spending on children: Experimental evidence from Tanzania. WIDER Working Paper 2017/128. Helsinki: UNU-WIDER.

Sankoh, A. J., M. F. Huque, and S. D. Dubey (1997). Some comments on frequently used multiple endpoint adjustment methods in clinical trials. *Statistics in Medicine 16*(22), 2529–2542.

Seguino, S. (2008). Micro-macro linkages between gender, development, and growth: Implications for the Caribbean region. *Journal of Eastern Caribbean Studies 33*(4), 8–42.

Ssewamala, F. M., L. Ismayilova, M. McKay, E. Sperber, W. Bannon, and S. Alicea (2010). Gender and the effects of an economic empowerment program on attitudes toward sexual risk-taking among aids-orphaned adolescent youth in Uganda. *Journal of Adolescent Health 46*(4), 372–378.

Strandow, D., M. G. Findley, and J. Young (2014). Foreign aid and the intensity of violent armed conflict. AidData Working Paper 24. Williamsburg, VA: College of William and Mary.

Thomas, D. (1990). Intra-household resource allocation an inferential approach. *The Journal of Human Resources 25*(4), 635–664.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological) 58*(1), 267–288.

Wood, R. M. and C. Sullivan (2015). Doing harm by doing good? the negative externalities of humanitarian aid provision during civil conflict. *The Journal of Politics 77*(3), 736–748.

OECD (2016). Aid in support of gender equality and women's empowerment: Statistical overview. Paris: OECD-DAC.

Data References

AidData. 2016. UgandaAIMS_GeocodedResearchRelease_Level1_v1.4.1 geocoded dataset [Dataset]. Williamsburg, VA and Washington, DC: AidData. Accessed from http://aiddata.org/research-datasets on August 15, 2016.

Bureau of Statistics [Uganda]. 2002 Uganda Population and Housing Census [Dataset]. Minnesota Population Center. Integrated Public Use Microdata Series, International: Version 6.3 [Machine-readable database]. Minneapolis: University of Minnesota, 2014. Accessed from http://idhsdata.org on December 4, 2015.

National Statistical Office [Malawi]. Population and Housing Census 1998 [Dataset]. Minnesota Population Center. Integrated Public Use Microdata Series, International: Version 6.4 [Machine-readable database]. Minneapolis: University of Minnesota, 2015. Accessed from http://idhsdata.org on September 2, 2016.

National Statistical Office [Malawi] and ICF Macro. Malawi Demographic and Health Survey 2010 [Dataset]. Data Extract from mwir61fl.dta, mwmr61fl.dta, and mwge61fl.shp. IPUMS-Demographic and Health Surveys (IPUMS-DHS), version 3.0, Minnesota Population Center and ICF International [Distributors]. Accessed from http://idhsdata.org on September 26, 2016.

National Statistical Office [Malawi] and ORC Macro. Malawi Demographic and Health Survey 2000 [Dataset]. Data Extract from mwir41fl.dta, mwmr41fl.dta, and mwge43fl.shp. IPUMS-Demographic and Health Surveys (IPUMS-DHS), version 3.0, Minnesota Population Center and ICF International [Distributors]. Accessed from http://idhsdata.org on January 10, 2017.

Peratsakis, Christian, Joshua Powell, Michael Findley, Justin Baker and Catherine Weaver. 2012. Geocoded Activity-Level Data from the Government of Malawi's Aid Management Platform. [Dataset] Washington D.C. AidData and the Robert S. Strauss Center for International Security and Law. Accessed from http://aiddata.org/research-datasets on September 26, 2016.

Uganda Bureau of Statistics [Uganda] and ICF International. Uganda Demographic and Health Survey 2011 [Dataset]. Data Extract from ugir60fl.dta, ugmr60fl.dta, and ugge61fl.shp. IPUMS-Demographic and Health Surveys (IPUMS-DHS), version 3.0, Minnesota Population Center and ICF International [Distributors]. Accessed from http://idhsdata.org on July 20, 2015.

Uganda Bureau of Statistics [Uganda] and ORC Macro. Uganda Demographic and Health Survey 2000-01 [Dataset]. Data Extract from ugir41fl.dta, ugmr41fl.dta, and ugge43fl.shp. IPUMS-Demographic and Health Surveys (IPUMS-DHS), version 3.0, Minnesota Population Center and ICF International [Distributors]. Accessed from http://idhsdata.org on January 10, 2017.

Appendix

	Treat, N=4402		Control, N=458		
	Mean	(s.d.)	Mean	(s.d.)	Nor Diff
Distance to major road	7.426	8.739	12.953	11.388	-0.565
Distance to pop center	27.323	20.763	38.043	18.461	-0.561
Distance to borderpost	66.597	38.491	69.052	43.063	-0.062
Annual temperature	213.754	18.627	220.280	18.586	-0.361
Temperature wettest q	226.022	18.496	222.234	17.388	0.217
Annual rainfall	1098.488	240.702	1170.171	187.176	-0.340
Rainfall wettest m	231.696	49.191	206.024	50.625	0.530
Rainfall wettest q	625.683	127.001	545.939	149.258	0.595
Share under agriculture within 1km	31.461	20.363	38.590	21.570	-0.351
Landcover within 1km	5.754	4.209	4.305	3.409	0.388
Agro-ecological zones	15.792	2.029	16.927	2.021	-0.577
Rainfall July-June	900.086	157.618	1038.776	174.710	-0.861
Rainfall wettest q, July-June	592.773	95.036	532.663	136.241	0.532
Start of wettest q	14.815	3.967	11.249	5.642	0.760
Change in greenness, pre-tr. district avg.	62.906	12.785	66.026	21.183	-0.186
Onset of greenness, pre-tr. district avg.	141.583	30.422	108.649	35.335	1.033
Onset of greenness decrease, pre-tr. district avg.	227.515	40.400	198.218	57.907	0.610
Peak of greenness, pre-tr. district avg.	0.504	0.040	0.511	0.046	-0.181
Household size	5.210	2.454	5.312	2.676	-0.041
Father education	5.332	15.407	13.700	26.150	-0.503
Mother's education	7.493	20.033	23.604	34.253	-0.712
Number of daughters	0.979	1.070	1.049	1.131	-0.068
Number of sons	0.978	1.076	1.272	1.251	-0.275
Share of women employed, pre-tr. district	0.545	0.105	0.526	0.120	0.272
Share of women never in school, pre-tr. district	0.401	0.180	0.394	0.096	0.089
Share of women illiterate, pre-tr. district	0.474	0.189	0.468	0.124	0.060
Share of women w pri edu, pre-tr. district	0.341	0.177	0.310	0.086	0.425
Share of girls never in school, pre-tr. district	0.209	0.178	0.183	0.044	0.436
Share of girls attending primary, pre-tr. district	0.791	0.178	0.817	0.044	-0.436
Years of school, pre-tr. district avg.	20.944	1.968	21.465	1.048	-0.629
Educational attainment, pre-tr. district avg.	1.117	0.163	1.074	0.053	0.733
Never been in school, pre-tr. district share	0.290	0.163	0.273	0.045	0.308
Access to grid, pre-tr. district share	0.080	0.147	0.034	0.028	0.961
Access to piped water, pre-tr. district share	0.114	0.197	0.045	0.035	1.064
Distance to market	22.252	16.479	37.929	13.846	-1.057
Distance to adm center	22.780	21.519	37.984	22.818	-0.707
Housing with no floor, pre-tr. district share	0.774	0.225	0.825	0.103	-0.572
Mobile phone, pre-tr. district share	0.060	0.092	0.032	0.015	0.938
Landline, pre-tr. district share	0.008	0.016	0.002	0.001	1.073
Elevation	6.157	3.173	6.680	3.443	-0.163
Slope	272.271	516.464	726.646	591.650	-0.846
	Mean	(s.d.)	Mean	(s.d.)	p-val
Urban dummy	0.598	0.491	0.363	0.481	0.000
Poor status	0.329	0.471	0.287	0.452	0.055
Malawi dummy	0.383	0.487	0.793	0.405	0.000

Table 1: Summary Statistics with Normalized Differences, LSMS

		Aid projects		Gender projects	
	Variable	Estimate	p.value	Estimate	p.value
1	Intercept	-4.346	0.000	0.206	0.803
2	Malawi dummy	1.959	0.000	3.498	0.000
3	Annual temperature			0.005	0.136
4	Annual rainfall	-0.000	0.248		
5	Rainfall wettest m	0.001	0.359		
6	Distance to adm center	-0.014	0.000	-0.023	0.000
7	Distance to borderpost	0.003	0.001		
8	Distance to market	-0.016	0.000	-0.045	0.000
9	Distance to pop center	-0.007	0.000		
10	Distance to major road	-0.004	0.186	-0.019	0.002
11	Change in greenness, pre-tr. district avg.	0.009	0.000		
12	Peak of greenness, pre-tr. district avg.	3.409	0.000		
28	Onset of greenness, pre-tr. district avg.	0.016	0.000		
29	Household size	0.012	0.292	-0.001	0.956
30	Poor status	-0.068	0.297	-0.434	0.003
31	Slope	0.002	0.000		
32	Urban dummy	-0.031	0.758	-0.405	0.026

Notes: Column 1 includes dummies for the land coverage within 1km.

Table 2: Estimated Parameters of Propensity Score Equation, LSMS

	Treat, N=1959		Control, N=516			
	Mean	(s.d.)	Mean	(s.d.)	Nor Diff (Trim)	Nor Diff (Untrim)
Distance to major road	11.207	10.254	12.399	11.110	-0.112	-0.565
Distance to pop center	36.215	20.186	36.486	18.631	-0.014	-0.561
Distance to borderpost	63.444	41.916	72.747	48.652	-0.205	-0.062
Annual temperature	217.700	18.722	216.574	17.161	0.063	-0.361
Temperature wettest q	227.143	19.819	220.506	15.455	0.373	0.217
Annual rainfall	1140.835	259.981	1136.886	188.105	0.017	-0.340
Rainfall wettest m	224.118	55.367	206.130	46.524	0.352	0.530
Rainfall wettest q	607.481	143.918	549.116	141.124	0.409	0.595
Share under agriculture within 1km	37.866	18.881	41.512	20.328	-0.186	-0.351
Landcover within 1km	4.142	2.829	4.047	3.085	0.032	0.388
Agro-ecological zones	16.153	2.093	17.200	1.988	-0.513	-0.577
Rainfall July-June	948.182	179.744	1023.153	168.166	-0.431	-0.861
Rainfall wettest q, July-June	571.688	103.138	545.281	132.495	0.222	0.532
Start of wettest q	13.813	4.731	11.649	5.670	0.414	0.760
Change in greenness, pre-tr. district avg.	62.786	12.759	63.096	18.266	-0.020	-0.186
Onset of greenness, pre-tr. district avg.	130.972	34.913	115.411	38.163	0.425	1.033
Onset of greenness decrease, pre-tr. district avg.	217.412	48.905	200.256	55.517	0.328	0.610
Peak of greenness, pre-tr. district avg.	0.505	0.039	0.508	0.050	-0.064	-0.181
Household size	5.340	2.542	5.390	2.765	-0.018	-0.041
Father education	7.477	19.342	12.527	25.903	-0.268	-0.503
Mother's education	10.947	24.542	21.298	33.824	-0.410	-0.712
Number of daughters	1.046	1.094	1.070	1.162	-0.022	-0.068
Number of sons	1.075	1.112	1.239	1.256	-0.145	-0.275
Share of women employed, pre-tr. district	0.553	0.100	0.536	0.118	0.235	0.272
Share of women never in school, pre-tr. district	0.382	0.128	0.379	0.086	0.037	0.089
Share of women illiterate, pre-tr. district	0.467	0.147	0.453	0.117	0.167	0.060
Share of women w pri edu, pre-tr. district	0.341	0.145	0.322	0.082	0.260	0.425
Share of girls never in school, pre-tr. district	0.178	0.106	0.173	0.043	0.102	0.436
Share of girls attending primary, pre-tr. district	0.822	0.106	0.827	0.043	-0.102	-0.436
Years of school, pre-tr. district avg.	21.478	1.470	21.479	0.995	-0.001	-0.629
Educational attainment, pre-tr. district avg.	1.106	0.125	1.077	0.050	0.508	0.733
Never been in school, pre-tr. district share	0.265	0.097	0.265	0.041	-0.000	0.308
Access to grid, pre-tr. district share	0.062	0.109	0.037	0.026	0.547	0.961
Access to piped water, pre-tr. district share	0.090	0.147	0.043	0.034	0.749	1.064
Distance to market	31.516	16.093	36.354	13.899	-0.322	-1.057
Distance to adm center	33.258	23.925	39.994	25.882	-0.270	-0.707
Housing with no floor, pre-tr. district share	0.795	0.187	0.809	0.098	-0.154	-0.572
Mobile phone, pre-tr. district share	0.049	0.069	0.032	0.014	0.584	0.938
Landline, pre-tr. district share	0.005	0.011	0.002	0.001	0.697	1.073
Elevation	5.975	3.284	6.610	3.265	-0.194	-0.163
Slope	427.053	593.831	651.893	604.516	-0.375	-0.846
	Mean	(s.d.)	Mean	(s.d.)	p-val	p-val
Urban dummy	0.517	0.500	0.308	0.462	0.000	0.000
Poor status	0.328	0.470	0.359	0.480	0.185	0.055
Malawi dummy	0.455	0.498	0.664	0.473	0.000	0.000

Table 3: Summary Statistics with Normalized Differences in the Trimmed Sample, LSMS

	Low	Middle	High	Sum
Control	9	516	13	538
Treated	3	1959	2264	4226
Sum	12	2475	2277	4764

Table 4: LSMS Sample Sizes for Subsamples with a Propensity Score between α and $1-\alpha$ ($\alpha = 0.1$) by Treatment Status

	Treat, N=2653		Control, N=2602		
	Mean	(s.d.)	Mean	(s.d.)	Nor Diff
Distance to adm center	0.209	0.130	0.226	0.136	-0.124
Distance to market	0.270	0.185	0.277	0.147	-0.043
Distance to pop center	0.493	0.378	0.587	0.400	-0.242
Distance to major road	0.038	0.039	0.036	0.043	0.047
Household size	6.097	2.798	6.198	2.711	-0.037
Uganda dummy	0.517	0.500	0.632	0.482	-0.233
Annual temperature	219.064	18.588	223.423	20.162	-0.225
Age	28.463	9.813	28.792	10.065	-0.033
Education level	2.055	0.584	2.018	0.588	0.063
Married	0.654	0.476	0.660	0.474	-0.013
Muslim	0.069	0.253	0.076	0.265	-0.026
Info from radio	0.823	0.381	0.818	0.386	0.014
Info from tv	0.280	0.449	0.228	0.420	0.118
Urban	0.105	0.307	0.054	0.226	0.190
Wealth Index	2.849	1.397	2.621	1.380	0.165

Table 5: Summary Statistics with Normalized Differences, DHS

	Middle	High	Sum
Control	3573	1195	4768
Treated	11707	23193	34900
Sum	15280	24388	39668

Table 6: DHS Sample Sizes for Subsamples with a Propensity Score between α and $1-\alpha$ ($\alpha = 0.1$) by Treatment Status

		Aid projects		Gender projects	
	Variable	Estimate	p.value	Estimate	p.value
1	Intercept	4.266	0.000	5.402	0.000
2	Distance to adm center	-2.772	0.000	-2.995	0.000
3	Distance to market	-1.352	0.000	-3.221	0.000
4	Distance to pop center	-1.853	0.000		
5	Distance to major road	-4.189	0.000	-1.368	0.026
6	Household size	0.002	0.749	-0.049	0.000
7	Uganda dummy	-2.622	0.000	-4.555	0.000
8	Annual temperature			-0.011	0.000
9	Age	-0.001	0.564	-0.004	0.070
21	Muslim	0.108	0.139		
24	Info from radio	-0.141	0.002		
25	Info from tv	0.313	0.000		
26	Urban	0.787	0.000	2.901	0.000
27	Wealth Index	0.112	0.000	0.161	0.000

Notes: All models include dummies for education level, occupation sector, ethnicity, and language. Column 1 also includes marital status.

Table 7: Estimated Parameters of Propensity Score Equation, DHS

Uganda	Below median	Above median	Diff	p-value
Women/men attended primary school	0.717	0.718	-0.000	0.494
Women/men attended secondary school	0.599	0.584	0.015	0.368
Girls/boys attended primary school	0.983	0.987	-0.004	0.349
Girls/boys attended secondary school	0.995	0.970	0.025	0.299
Num of reasons men agree beating wife is justified (0–5)	1.621	1.791	-0.170	0.275
Num of reasons women agree beating wife is justified (0–5)	1.849	2.252	-0.403	0.011
Malawi				
Women/men attended primary school	0.702	0.677	0.025	0.313
Women/men attended secondary school	0.431	0.430	0.001	0.491
Girls/boys attended primary school	0.994	0.992	0.002	0.459
Girls/boys attended secondary school	0.944	0.902	0.042	0.205
Num of reasons men agree beating wife is justified (0–5)	0.489	0.563	-0.074	0.255
Num of reasons women agree beating wife is justified (0–5)	1.098	0.899	0.198	0.094

Table 8: Pre-Existing Gender Gaps

Outcome	beta	Confidence	interval	p-value	Obs.	Blocks	Mean	Impact at mean
Working on farm, women	0.003	[-0.0021	0.0073]	0.863	1856	5	0.871	0.079
Working on farm, men	-0.021	[-0.0254	-0.0161]	0.000	1771	6	0.802	-0.630
Working, women	0.006	[0.0017	0.0104]	0.038	1856	5	0.915	0.184
Working, men	-0.012	[-0.0154	-0.0096]	0.000	1771	6	0.924	-0.380

Notes: p-values are modified with a Bonferroni correction adjusted for correlation.

Table 9: Adults' Outcomes, Aid vs. No Aid

Outcome	beta	Confidence	interval	p-value	Obs.	Blocks	Mean	Impact at mean
Working on farm, women	0.002	[-0.0075	0.0115]	0.683	872	3	0.899	0.035
Working on farm, men	-0.005	[-0.0122	0.0022]	0.172	834	4	0.912	-0.088
Working, women	0.000	[-0.0078	0.0087]	0.910	872	3	0.931	0.008
Working, men	-0.005	[-0.0112	6e-04]	0.081	834	4	0.961	-0.093

Notes: p-values are modified with a Bonferroni correction adjusted for correlation.

Table 10: Adults' Outcomes, Gender Projects vs. No Aid

Outcome	beta	Conf.	int.	p	Obs.	Bl.	Mean	Impact at mean
Beating wife is justified if (n of reasons, men)	-0.002	[-0.0086	0.0048]	0.577	3191	14	0.712	-0.052
Beating wife is justified (n of reasons, women)	-0.002	[-0.0069	0.0028]	0.404	12002	20	1.192	-0.057
Experienced physical violence in last 12 months	-0.008	[-0.0132	-0.0021]	0.018	2689	8	0.240	-0.211
Experienced sexual violence in last 12 months	-0.003	[-0.0083	0.0027]	0.625	2686	8	0.203	-0.077
Experienced emotional violence in last 12 months	-0.000	[-0.0064	0.0062]	1.000	2687	8	0.280	-0.003
Wife justified to refuse sex if husband unfaithful (men)	0.005	[0.0024	0.0074]	0.000	3159	14	0.732	0.134
Wife justified to ask for condom if husband has STI (men)	0.002	[6e-04	0.0039]	0.007	3164	14	0.882	0.062
Can refuse sex (w)	0.000	[-0.0017	0.0018]	0.956	7985	15	0.786	0.001
Can ask partner to use contraception (w)	0.006	[0.004	0.0072]	0.000	8085	16	0.736	0.155

Notes: p-values are modified with a Bonferroni correction adjusted for correlation.

Table 11: Domestic Violence and Sexual Freedoms, Aid vs. No Aid

Outcome	beta	Conf.	int.	p	Obs.	Bl.	Mean	Impact at mean
Beating wife is justified if (n of reasons, men)	0.011	[0.0043	0.0173]	0.001	1959	12	0.360	0.213
Beating wife is justified (n of reasons, women)	0.001	[-0.0038	0.0064]	0.628	6570	10	0.466	0.025
Experienced physical violence in last 12 months	-0.012	[-0.0159	-0.0075]	0.000	1571	8	0.147	-0.231
Experienced sexual violence in last 12 months	-0.008	[-0.0123	-0.0038]	0.001	1570	8	0.177	-0.158
Experienced emotional violence in last 12 months	-0.002	[-0.0067	0.0031]	0.814	1570	8	0.220	-0.036
Wife justified to refuse sex if husband unfaithful (men)	0.016	[0.013	0.0199]	0.000	1947	11	0.703	0.324
Wife justified to ask for condom if husband has STI (men)	0.010	[0.0079	0.0126]	0.000	1944	12	0.873	0.202
Can refuse sex (w)	0.004	[0.0011	0.0065]	0.006	4462	10	0.712	0.074
Can ask partner to use contraception (w)	0.005	[0.003	0.0076]	0.000	4522	9	0.824	0.105

Notes: p-values are modified with a Bonferroni correction adjusted for correlation.

Table 12: Domestic Violence and Sexual Freedoms, Gender Projects vs. No Aid

Outcome	beta	Conf.	int.	p	Obs.	Bl.	Mean	Impact at mean
Woman participates in financial decisions (n)	0.010	[0.002	0.0185]	0.015	3690	14	3.319	0.281
Expenditures, food	-0.003	[-0.0078	0.0014]	0.527	1315	2	0.662	-0.098
Expenditures, clothing	0.000	[-0.001	0.0018]	0.972	1315	2	0.025	0.012
Expenditures, male goods	0.002	[-2e-04	0.0035]	0.305	1315	2	0.012	0.050
Expenditures, children	-0.013	[-0.0175	-0.0088]	0.000	1580	5	0.123	-0.400

Notes: p-values are modified with a Bonferroni correction adjusted for correlation.

Table 13: Household Expenditure Decisions, Aid vs. No Aid

Outcome	beta	Conf.	int.	p	Obs.	Bl.	Mean	Impact at mean
Woman participates in financial decisions (n)	-0.012	[-0.0334	0.0104]	0.302	1609	6	3.325	-0.227
Expenditures, food	-0.001	[-0.0029	0.0011]	0.841	1279	5	0.654	-0.016
Expenditures, clothing	-0.001	[-0.0012	1e-04]	0.440	1279	5	0.028	-0.009
Expenditures, male goods	0.001	[7e-04	0.0022]	0.001	1279	5	0.013	0.025
Expenditures, children	-0.015	[-0.0212	-0.009]	0.000	1443	4	0.152	-0.265

Notes: p-values are modified with a Bonferroni correction adjusted for correlation.

Table 14: Household Expenditure Decisions, Gender Projects vs. No Aid

Outcome	beta	Confidence	interval	p-value	Obs.	Blocks	Mean	Impact at mean
Fetch water, women	0.017	[0.0092	0.0252]	0.000	1434	5	0.574	0.522
Fetch water, men	-0.006	[-0.0078	-0.0033]	0.000	1336	7	0.062	-0.169
Fetch water, girls	-0.008	[-0.0152	-0.0011]	0.188	1434	5	0.253	-0.247
Fetch water, boys	-0.005	[-0.008	-0.0025]	0.001	1336	7	0.131	-0.160
Collect firewood, women	0.016	[0.0037	0.0289]	0.094	813	3	0.636	0.496
Collect firewood, men	-0.002	[-0.0078	0.0034]	0.995	770	3	0.037	-0.068
Collect firewood, girls	-0.010	[-0.0208	0.0013]	0.544	813	3	0.236	-0.295
Collect firewood, boys	-0.006	[-0.0132	0.0012]	0.614	770	3	0.100	-0.183

Notes: p-values are modified with a Bonferroni correction adjusted for correlation.

Table 15: Domestic Chores, Aid vs. No Aid

Outcome	beta	Confidence	interval	p-value	Obs.	Blocks	Mean	Impact at mean
Fetch water, women	0.015	[0.0085	0.0224]	0.000	1174	5	0.538	0.272
Fetch water, men	-0.003	[-0.0075	0.001]	0.131	1104	5	0.088	-0.058
Fetch water, girls	-0.008	[-0.0149	-0.002]	0.011	1174	5	0.260	-0.148
Fetch water, boys	-0.005	[-0.0105	8e-04]	0.090	1104	5	0.150	-0.086
Collect firewood, women	0.007	[-0.0024	0.0172]	0.140	610	5	0.615	0.130
Collect firewood, men	0.001	[-0.0051	0.0069]	0.774	571	4	0.053	0.016
Collect firewood, girls	-0.010	[-0.0188	-0.0015]	0.022	610	5	0.247	-0.179
Collect firewood, boys	-0.004	[-0.0113	0.0038]	0.329	571	4	0.105	-0.066

Notes: p-values are modified with a Bonferroni correction adjusted for correlation.

Table 16: Domestic Chores, Gender Projects vs. No Aid

Outcome	beta	Confidence	interval	p-value	Obs.	Blocks	Mean	Impact at mean
Working on farm, girls	-0.021	[-0.0258	-0.0156]	0.000	1345	7	0.618	-0.629
Working on farm, boys	-0.010	[-0.0177	-0.0023]	0.065	1456	6	0.650	-0.304
Working, girls	-0.020	[-0.0256	-0.0154]	0.000	1345	7	0.625	-0.622
Working, boys	-0.010	[-0.0177	-0.0022]	0.068	1456	6	0.661	-0.303
Never attended, girls	-0.002	[-0.0046	8e-04]	0.164	1684	6	0.065	-0.058
Never attended, boys	0.003	[-0.001	0.0067]	0.152	1795	5	0.076	0.086
Attending secondary, girls	-0.001	[-0.013	0.0113]	0.887	694	4	0.800	-0.027
Attending secondary, boys	0.009	[-3e-04	0.0189]	0.057	765	4	0.837	0.282
Attending primary, girls	0.000	[-0.0043	0.0051]	0.863	937	4	0.935	0.013
Attending primary, boys	0.002	[-0.0037	0.0074]	0.517	954	4	0.930	0.056

Notes: p-values are modified with a Bonferroni correction adjusted for correlation.

Table 17: Children's Outcomes, Aid vs. No Aid

Outcome	beta	Confidence	interval	p-value	Obs.	Blocks	Mean	Impact at mean
Working on farm, girls	-0.018	[-0.0346	-0.0024]	0.024	751	3	0.648	-0.325
Working on farm, boys	-0.022	[-0.0311	-0.0123]	0.000	819	4	0.681	-0.382
Working, girls	-0.021	[-0.0366	-0.0044]	0.012	751	3	0.667	-0.361
Working, boys	-0.022	[-0.0312	-0.0122]	0.000	819	4	0.681	-0.381
Never attended, girls	-0.003	[-0.0108	0.0058]	0.553	773	3	0.086	-0.044
Never attended, boys	0.004	[-0.003	0.0106]	0.277	845	4	0.118	0.067
Attending secondary, girls	-0.014	[-0.0441	0.0169]	0.384	271	2	0.790	-0.238
Attending secondary, boys	0.008	[-0.0092	0.026]	0.347	327	3	0.819	0.148
Attending primary, girls	-0.001	[-0.009	0.0068]	0.781	434	2	0.982	-0.020
Attending primary, boys	-0.000	[-0.0118	0.0114]	0.972	418	2	0.945	-0.004

Notes: p-values are modified with a Bonferroni correction adjusted for correlation.

Table 18: Children's Outcomes, Gender Projects vs. No Aid