Paid Work for Women and Domestic Violence: Evidence from the Rwandan Coffee Mills

Deniz Sanin[§]

October 19, 2021

Abstract

This paper studies whether providing paid employment opportunities to women decreases the violence they face from their partners. Using the government-induced rapid expansion of the coffee mills in Rwanda in the 2000s, I first provide causal evidence that a mill opening increases women's paid employment, women's and their husbands' earnings and decreases domestic violence. Then I show that the decline in violence is plausibly driven by women's paid employment, not an increase in husbands' earnings. For identification, I uniquely perform two strategies with two sources of domestic violence data, self-reports and novel data on the universe of monthly hospitalizations. The opening of a mill affects coffee farmers who reside in its catchment area, a buffer zone around the mill, during the harvest months. By providing paid job opportunities for female-dominated tasks, a mill opening enables women to transition from being unpaid family workers in their family plots to wage workers in the mills. First, I perform a staggered difference-in-differences (DID) using differential timing of and spatial variation in exposure to a mill opening during the 2000s. I show that upon a mill opening, women in the catchment areas are 18% more likely to work for cash and 26% less likely to self-report domestic violence in the past 12 months. Second, I focus at the end of the expansion, where the number of mills is fixed. I perform a DID event-study, using the spatial variation in exposure to a mill and the timing of the harvest months within a year. I show that it is 23% less likely for hospitals in the catchment areas to have a domestic violence patient in a harvest month compared to one month before the beginning of the harvest season. After a mill opening, women in the catchment areas are also more likely to participate in household decisions. The decline in violence is present even among couples where there is no change in husbands' earnings.

Keywords: Domestic violence, female labor supply, agriculture, Rwanda

JEL Codes: J12, J16, J21, J31, O12

[§]Economics Ph.D. Candidate, Georgetown University, Department of Economics & Pre-Doctoral Research Fellow, Women and Public Policy Program, Harvard Kennedy School, ds1521@georgetown.edu. I am extremely grateful to my advisor Garance Genicot for her invaluable guidance and support on this project. This paper greatly benefited from detailed discussions with Laurent Bouton, Mary Ann Bronson, Roger Lagunoff, Alexandre Poirier, Martin Ravallion and Andrew Zeitlin and I owe special thanks to them. I also owe special thanks to Leodimir Mfura, who helped me throughout my data application process. This project has human subjects approval from Rwanda National Ethics Committee. I acknowledge the financial support of Georgetown Graduate School's Dissertation Travel Grant and GradGov Research Project Award.

1 Introduction

Domestic violence is an extreme form of gender inequality, violation of human rights and a global health problem of epidemic proportions (WHO, 2013). About 1 in every 3 women worldwide have experienced either physical and/or sexual violence from their partners in their lifetime (World Bank, 2015). To address domestic violence, the economics literature focuses on laws, enforcement, shelters, education, cash transfers, unemployment benefits and job opportunities for women. Although there is evidence on the effects of the former policies (Stevenson and Wolfers 2006, Aizer and Dal Bó 2009, Anderson and Genicot 2015, Brassiolo 2016, Chin and Cunningham 2019, García-Ramos 2021, Sanin 2021, Miller and Segal 2019, Sviatschi and Trako 2021, Farmer and Tiefenthaler 1997, Erten and Keskin 2018, Green et al. 2020, Angelucci 2008, Bobonis et al. 2013, Hidrobo et al. 2016, Haushofer et al. 2019, Bhalotra et al. 2021), there is limited causal evidence on the effects of increased job availability for women on domestic violence.

Understanding the relationship between increased job availability for women and domestic violence is particularly crucial for Sub-Saharan Africa. Since the 1990s, the region has the highest female labor force participation rates globally, where more than 60% of the employed women work in agriculture and are often unpaid family workers in their family plots (FAO, 2010). Recently, the region is under a large-scale economic transition. Between 2000 and 2018, 9 million paid jobs are created per year, mostly in agriculture where women work dominantly (IMF, 2018).

This paper investigates whether providing paid employment opportunities to women decrease the violence they face from their partners. I answer my research question in two steps. First, I provide causal evidence that the government-induced rapid expansion of the coffee mills in Rwanda in the 2000s, increased women's paid employment, increased earnings of women and their husbands and decreased DV. To establish the results, I uniquely use two sources of domestic violence data, self-reports and novel data on the universe of monthly hospitalizations, data on individual employment and earnings simultaneously. Then, I show that the decline in violence is plausibly driven by women's paid employment, not necessarily by an increase in husbands' earnings via two pieces of evidence. First, after a mill opening, women in the catchment areas are more likely to participate in household decisions. Second, I show that there is a decline in violence even among couples where husbands work in occupations with no change in earnings with mill exposure, non-agricultural manual jobs.

In 2002, the Rwandan government adopted the National Coffee Strategy that aimed to shift to mill-processed coffee production to participate in the international specialty coffee market (Boudreaux, 2011). In the beginning of the 2000s, the governmental institutions collaborated with USAID, universities in Rwanda and U.S. and private sector partners under the Partnership to

Enhance Agriculture in Rwanda through Linkages (PEARL) project. The project helped farmers to establish cooperatives and build mills in their communities. After the project, farmers continue to build mills across the country. From 2002 to 2012, the number of mills increased from 5 to 213. In 2012, coffee accounted for nearly 20% of Rwanda's exports and 15% of its GDP (Macchiavello and Morjaria, 2020).

A coffee mill is where one of the early stages of the coffee value chain, processing, takes place. Coffee cherries (harvest) are processed into coffee beans so that they can be prepared for export. The mill expansion provides two unique features for identification. First, a mill only serves coffee farmers that reside within its catchment area, a buffer zone around the mill. It has a specific catchment area since coffee cherries will rot if not transported to a mill within few hours of harvest. Second, a mill opening enables women's transition to paid employment. Before a mill opening, farmers process coffee at home; they clean cherries, dry them and sort out the defective ones all by hand. Then they sell the product in the local market. Drying and sorting are female-dominated tasks and selling is male-dominated. Due to this gendered division of labor, husbands receives the income from coffee as their personal earnings and control it. After a mill opening, within the catchment area, husbands sell coffee cherries to the mill. The cherries are cleaned by the machinery in mills. For drying and sorting tasks, mills demand paid labor. Since the tasks are female-dominated, a mill provides the opportunity for women in the catchment areas to transition from unpaid family workers to daily wage workers in the mills for the same tasks as before. Given that the husbands were already generating income, there is no change in their type of earnings. Yet, the husband earns a higher income from selling cherries to mills compared to selling home-processed coffee in the local market (Macchiavello and Morjaria, 2020). Women outside of the catchment area continue to be unpaid family workers.

To understand how a mill opening affects domestic violence, I first present a conceptual framework and present four effects on domestic violence. First, as the wife's earnings increases, her outside option improves. This increases the probability of her initiating divorce and makes it less likely for the husband to inflict violence. As an opposing effect, the husband is incentivized to choose violence to extract the newly acquired resources from the wife. Third, as the household income increases either via an increase in his or his wife's earnings, his financial stress decreases. This makes him less likely to inflict violence. However, the wife's benefit from the marriage increases as the husband's income increases. This makes her less likely to leave and thus, the husband is more likely to choose violence.

To causally identify the impact of a mill opening, I uniquely employ two empirical strategies using two sources of domestic violence data. First, I perform a staggered DID design using two variations based on the context: within-outside of the catchment area (spatial variation) and before-after a mill opening (time variation). Due to rapid expansion during the 2000s, different areas of the country experienced a mill opening in different points in time. For this identification strategy, I use self-reported domestic violence and labor market outcomes data. For my second empirical strategy, I focus at the end of the expansion where the number of mills are fixed. I perform a DID event-study design. I use the same spatial variation as before and monthly time variation within a year. Mills open/operate only during the harvest months, March-July. I use the beginning of the harvest season, March, as the event. For this specification, I use novel data, the universe of monthly hospitalizations for domestic violence in Rwanda.

Multiple cycles of Rwandan Demographic Health Surveys (DHS) provides information on couples' self-reported labor market outcomes and women's domestic violence. Rwanda Ministry of Health's 2012-2014 Hospital Management Information System (HMIS) data provides the universe of monthly hospitalizations due to domestic violence. I combine data on the universe of mills using Rwanda GeoPortal and Macchiavello and Morjaria (2020). All aforementioned datasets are geocoded that enables me to link the individuals (woman/husband) and hospitals with the mills based on the GPS coordinates. I also use Rwandan Living Standards Surveys (EICV) which provides information on couples' earnings to uncover the mechanisms behind the results. Although the survey is not geocoded, since I have information on the district of residence of the couples, I am able to analyze the impact of a mill opening by using mill per capita at the district level.

When I perform staggered DID design with the DHS data, the treatment group consists of couples who are within the catchment areas where the control group consists of couples who are outside. I define the catchment area as residing within 4 km of a mill based on previous reports and research on Rwandan coffee industry (AgriLogic 2018, Macchiavello and Morjaria 2020). My identification strategy compares the changes in self-reported outcomes of couples within the catchment area of a mill with the ones who are outside of the catchment area before and after a mill opening. I employ two approaches for the control group in my analysis. My main control group consists of DHS clusters outside of the catchment area but are located within the same district with the mills (within district approach). As a second approach, I use an additional control group that consists of entities who are within the donut area between 4 and 8 km from the mill (donut approach).

When within district approach is used, being exposed to a mill increases the probability of working for cash in the past 12 months by 7 percentage points (p-value= 0.001) where the probability of working in the past 12 months remains unchanged. The estimated impact represents an increase of 18% with respect to the sample mean (0.40). Importantly, being exposed to a mill decreases the probability of self-reporting a domestic violence experience in the past 12 months by 9 percentage points (p-value= 0.01). The estimated impact represents a decrease of 26% with respect to the sample mean (0.35). There is no statistically significant change in husbands' probability of working and type of earnings being cash in the past 12 months. The occupation for each spouse remains unchanged. My results are very similar when I use the donut approach for the control group.

When I perform DID event study design with the hospitalizations data, the treatment group consists of hospitals that are within the catchment areas where the control group consists of the hospitals that are outside but located within the same province with the mills.[1] Mills operate only during the harvest months, March-July. Specifically, May-July is the peak of the harvest where majority of the neighboring community around the mills work in the mills. The identification strategy compares the changes in hospitalizations within the catchment areas during harvest months within a year, relative to one month before the harvest season, February.

I show that it is 18 and 14 percentage points (p-value= 0.01, p-value= 0.06) less likely for a hospital in the catchment area of a mill to have a domestic violence patient in June and July respectively compared to one month before the mills' month of operation, February. The estimated impact represents a reduction of 23% and 17% respectively with respect to the sample mean (0.79). There are no statistically significant changes both for January, two months before mills' month of operation, as well as for the post-harvest months when mills do not operate. As a placebo test, I present results for hospitalizations other than domestic violence. I find no changes.

In order the uncover the mechanisms behind the decline in violence, I estimate the impact of mill exposure on bargaining power related outcomes and perform a subsample analysis. First, I show that upon a mill opening, women in the catchment areas are more likely to make household decisions jointly with their partners including large household purchases and contraception usage. Second, I show that the decline in domestic violence is observed even among couples who work in different occupations. The sample primarily consists of farmer women with husbands who work in non-agricultural manual jobs (truck drivers, plumbers etc.). For this subsample, increase in women's earnings is accompanied with no change in husbands' earnings. This shows that the decline in financial stress due to the increase in husbands' earnings is not necessarily the dominant mechanism behind my results. Moreover, unlike farmer couples, these couples do not work together. Thus, a mill opening is not a shock to the time they are exposed to each other during work hours. Thus, exposure reduction is also ruled out as the dominant mechanism. Given that women's earnings increases, a decline in financial stress due to the increase in women's earnings is also a plausible mechanism. Since all plausible channels are due to women's paid employment, the paper suggests that women's paid employment is the driver behind the decline in DV. Moreover,

[1] Province is one geographical unit above the district. There are not enough hospitals in each district for me to use the within district approach. There are in total 42 hospitals in Rwanda.

increase in women's outside options and decrease in financial stress, not exposure reduction are two plausible channels on how women's paid employment can affect domestic violence.

I also investigate the dynamic impact of a mill opening using an DID event study specification exploiting the number of years women/husbands are exposed to a mill opening. Dynamic estimates show that the effects on domestic violence persists for 4 years. After 4 years, there is no decline in domestic violence, which suggests that the effects of women's paid employment in low-paid jobs may not persist for long years. According to dynamic estimates, for all variables, the coefficients are close to zero and statistically insignificant for the years before a mill opening. Thus, the estimates also constitute evidence in favor of the parallel trends assumption.

Recent econometric literature on DID estimators that use variation in treatment timing raises concerns about the validity of estimation results in the presence of treatment effect heterogeneity. My results are robust to using estimators proposed in de Chaisemartin and D'Haultfœuille (2020) and Sun and Abraham (2020) that gives valid results even if the treatment effect is heterogeneous over time and across groups. I perform the decomposition outlined in Goodman-Bacon (2021) and show that the overwhelming majority of my results are based on treated and never treated units. I also perform the test outlined in de Chaisemartin and D'Haultfœuille (2020) and document that treatment effect heterogeneity is not a concern for the validity of my estimates.

Throughout the paper, my main identifying assumption is that conditional on mill suitability, exposure to a mill is random. To better control for mill suitability, as a robustness check, I predict a mill suitability index using a machine learning algorithm, Least Absolute Shrinkage and Selection Operator (LASSO), and use it as a control variable in my main empirical specification. Results are robust to controlling for the LASSO predicted mill suitability index. I also perform robustness checks related to catchment area measurement . First, I show that there is no statistically significant differences in outcome variables across women within the 4 km catchment area. Second, I experiment with different catchment area sizes, 5 and 10 km. I show that at the buffer radius increases, more untreated households are counted as treated, the effects of mill exposure fades out. Moreover,

Contributions and Related Literature. This paper contributes to several strands of the literature. First, it contributes to the economics literature on domestic violence. Combining data on self-reported domestic violence, the universe of hospitalizations for domestic violence and individual employment and earnings, I provide casual evidence that the opening of a coffee mill in Rwanda increases women's paid employment, earnings of women and their husbands and decreases domestic violence. I show that the decline in domestic violence is plausibly driven by women's increased job availability, not an increase in husbands' earnings. My results suggest that women's paid employment decrease domestic violence via an increase in women's outside option, decline in financial stress due to a rise in women's earnings, not exposure reduction.

The domestic violence literature provides evidence for both a positive and negative relationship between women's income and domestic violence. The negative relationship between the two is explained with the theory on the increase in the outside option, the decline in financial stress and exposure reduction. The positive relationship is explained with instrumental/extractive theory of domestic violence and male backlash. In the theory on the increase in the outside option, decline in financial stress and the instrumental theory of domestic violence, violence is incorporated into household bargaining models via men's motives.[2] If a man is violent because it contributes to his utility directly, via a release of stress or frustration, then the violence is *expressive* (Gelles 1974, Tauchen et al. 1991). If a man is violent to extract resources from his wife to increase his consumption of goods or to control the wife's behavior, then the violence is *instrumental* (Gelles 1974, Tauchen et al. 1991). I briefly review the theories and the literature on the relationship between women's income and domestic violence below.

I. Increase in the Outside Option. An economic theory of household bargaining that incorporates only expressive violence (Farmer and Tiefenthaler 1997, Aizer 2010, Anderberg et al. 2016, Hidrobo et al. 2016) or expressive and instrumental violence combined (Tauchen et al. 1991, Haushofer et al. 2019), suggests that an increase in women's income decreases domestic violence by improving her outside option and thus her bargaining power. For the choice of outside options, the models use either value of being divorced as in the divorce threat model (Manser and Brown 1980. McElroy and Horney 1981) or noncooperative equilibrium within the marriage as in the separate spheres model (Lundberg and Pollak, 1993)

II. Decline in Financial Stress. Tauchen et al. 1991 highlights that expressive violence does not necessarily need to be sadistic but may allow the husband to release economic stress. Similarly, Angelucci (2008) argues that an increase in spousal income may reduce the husband's economic stress and decrease domestic violence. Haushofer et al. (2019) incorporates expressive violence into their model to analyze the relationship between economic stress and domestic violence. The paper highlights that if violence and the husband's wealth are substitutes in his utility's expressive component, then they are also substitutes in his total utility. Thus, an increase in the husband's income decreases the violence he inflicts on his wife. The theory is in line with evidence from behavioral neuroscience that finds a link between stress and aggressive behavior (Kruk et al., 2004).

III. Exposure Reduction. Criminologists develop the theory of exposure reduction to argue that an increase in female employment may reduce domestic violence due to the decline in the time couples spent together (Dugan et al., 1999).

[2]See Chiappori and Mazzocco (2017) for a detailed review of household decisions.

IV. Instrumental Violence. There also exists theories which suggest that there is a positive relationship between women's income and domestic violence. Bloch and Rao (2002), Eswaran and Malhotra (2011), Bobonis et al. (2013) and Calvi and Keskar (2021) present a bargaining model that incorporates instrumental violence where the husband uses violence or threats of violence as a bargaining tool to extract resources from his wife or to enhance his bargaining power. Both models argue that violence is an instrument in household bargaining for men and an increase in women's income also increases the domestic violence they face.

V. Male Backlash. Theory of male backlash proposed by sociologists also argues that an increase in women's income may indeed increase domestic violence. According to the theory, female employment challenges the husband being the "breadwinner" and thus, he inflicts violence on his wife to reinstate his traditional gender role (Macmillan and Gartner, 1999). Although both instrumental violence and male backlash theory suggest a positive relationship between female employment and domestic violence, male backlash theory do not necessarily argue that the husband's motive to inflict violence is to extract resources from the wife. The theory focuses on the husband perceiving the wife's employment as a status threat in the household due to entrenched social norms.

There is growing empirical evidence that support the theory on the increase in women's outside options (Aizer 2010, Anderberg et al. 2016, Hidrobo et al. 2016, Haushofer et al. 2019), reduction in financial stress (Angelucci 2008, Bhalotra et al. 2019, Heath et al. 2020, Arenas-Arroyo et al. 2021), exposure reduction (Chin, 2011), instrumental violence (Bloch and Rao 2002, Eswaran and Malhotra 2011, Bobonis et al. 2013, Heath 2014, Erten and Keskin 2018, Bhalotra et al. 2019, Erten and Keskin 2018, Erten and Keskin 2020, Calvi and Keskar 2021 Erten and Keskin 2021b) and male backlash (Angelucci 2008, Luke and Munshi 2011, Tur-Prats 2021, Alesina et al. 2020, Guarnieri and Rainer 2021). The evidence is coming from using dowry payments, randomized cash transfers, compulsory schooling laws, labor demand shocks, gender norms and colonial borders. My empirical results are closely related to Aizer (2010) that provides causal evidence that an increase in women's relative wages decreases domestic violence due to the improvement in women's outside options in California.

Second, this paper contributes to the literature on the consequences of female employment in developing countries. Heath and Jayachandran (2017) reviews the literature and highlights that increased job availability for women is found to have both positive consequences including increases in women's bargaining power (Heath and M. Mobarak, 2015) and children's health (Qian, 2008) and negative consequences including domestic violence. The review highlights the aforementioned papers that shows a positive relationship between female employment and domestic violence and confirms that the causal evidence on the relationship is limited and mixed. My results fill this gap in the literature. Female employment is strikingly low in developing countries and recent literature finds that correcting husbands' beliefs and opening bank accounts increase female labor supply (Bursztyn et al. 2020, Field et al. 2021). My results suggest that interventions that increase female employment have the potential to benefit women beyond their labor supply. More broadly, my research relates to the large literature on female empowerment in developing countries.[3]

Lastly, this paper contributes to the literature on domestic violence and female empowerment in Rwanda (La Mattina 2017, Rogall 2021, Sanin 2021). Sanin (2021) investigates the adoption of domestic violence laws on domestic violence in Rwanda and provides causal evidence that the laws protect women in violent marriages. This paper provide evidence that the rapid expansion of coffee mills in the country also protects women from domestic violence.

The organization of the paper is as follows. Section 2 provides background information on the Rwandan coffee industry, rapid expansion of coffee mills and women's paid employment in the coffee value chain. Section 3 introduces the theoretical framework that guides the empirical analysis. Section 4 introduces all data sources. Section 5 outlines the identifying assumptions and threats to identification and propose the identification strategy. Section 6 present the empirical results and investigate the mechanisms. Section 7 provides robustness checks. Section 8 concludes.

2 Institutional Context

In this section, I first provide background information on the coffee industry in Rwanda. Second, I present the agricultural reforms in the early 2000s and the rapid expansion of mills following the reforms. Third, I provide information on why the expansion of mills is likely to have pronounced effects on women's employment.

2.1 The Coffee Industry in Rwanda

Value chain and coffee quality. There are four main steps of the coffee value chain: Cultivation, processing, roasting, and consumption. The first two stages, cultivation and processing are the stages of interest throughout the paper. Both stages includes labor intensive processes which affect the quality of the end product.

A coffee tree produces coffee cherries that contains coffee beans. It takes at least three to five years for a coffee tree to produce cherries after it is planted. When the cherries ripen, they should be harvested (picked) by hand. The cherries do not ripen all at once which makes harvesting a

[3] See Duflo (2012) for a detailed review of female empowerment in developing countries.

labor-intensive process. The harvest season lasts for approximately four to five months. It lasts from March to July in Rwanda and the peak of the harvest is from May to July.

After harvest, coffee needs to be processed. Processing is made up of three main tasks. The first task is cleaning, which is removing the outer layer (pulp) of the coffee beans. In the second task, the coffee beans are dried under the sun on flat surfaces.[4] In the third task, the defective dried coffee beans are sorted out by hand based on their color and size, which is a very labor-intensive process like harvesting.

Coffee processing methods. There are two possible methods to process the coffee cherries: dry processing and wet processing. In the dry processing, which is also known as the traditional method due to being the oldest coffee processing method, all tasks are done by farmers at their homes without any machinery.[5] The outcome of this process yields a low quality product which is sold in the local market for a very low price.

In the wet processing method, farmers sell their coffee cherries to coffee mills (called washing stations in Rwanda). A mill can be thought as a large firm for developing country standards where coffee is wet-processed. In the mills, cleaning is done with specific machinery which uses plenty of water. That is why the method is called wet processing. Drying and sorting tasks are done by hand as in dry processing. Mills demand seasonal wage labor from the neighboring community for these tasks. The method yields a high quality product which is sold in the international market for a high price to international buyers.[6] Multinational companies like Starbucks is an example of an international buyer of Rwandan wet-processed coffee.

2.2 Rapid expansion of the mills

Agricultural Reforms. In 2000, Paul Kagame came into power and prioritized economic growth to rebuild the country in the aftermath of the Rwandan Genocide (1994). He launched the Vision 2020 program in 2000 (Boudreaux, 2011). The program outlined a list of goals which the government aimed to achieve by 2020. One of the main goals was to transform agriculture into a high value sector. In light of this goal, the government adopted the National Coffee Strategy in 2002 which aimed to shift to high quality, wet-processed coffee production to participate in the international specialty coffee market (Boudreaux, 2011). At the time, 90% of the Rwandan coffee was dry-processed and thus classified as low-quality (MINAGRI and MINICOM, 2008). The government also liberalized the coffee industry and removed barriers to trade (Boudreaux, 2011).

[4] The dried beans are called parchment coffee.

[5] Farmers clean cherries using rocks and then dry the beans on mats (Macchiavello and Morjaria, 2020).

[6] In 2012, wet-processed (also called as fully washed) coffee export gate prices are roughly 40% higher than for dry-processed coffee in Rwanda (Macchiavello and Morjaria, 2020).

The reforms in the coffee industry changed the incentives of individuals in the coffee value chain. First and foremost, farmers started to have the incentive to use wet-processing rather than dry-processing since it yields a higher income (Boudreaux, 2011). Exporters compete to sell the wet-processed coffee to foreign buyers. Foreign buyers are interested in buying the Rwandan coffee since the quality increased as the coffee is now wet-processed.

Expansion of the Mills. Between 2000-2006, governmental institutions collaborated with USAID, universities in Rwanda and U.S., and private sector partners under the Partnership to Enhance Agriculture in Rwanda through Linkages (PEARL) project. The project helped farmers to establish cooperatives, find loans and build mills in their communities in few locations. After 2006, farmers continue to establish cooperatives and build mills in their communities across the country. The number of mills expanded rapidly between 2005-2011. Figure 1 shows the expansion of the mills in Rwanda over the last decades. In 2005, the total number of mills was 49. In 2011, the number quadrupled to 197.

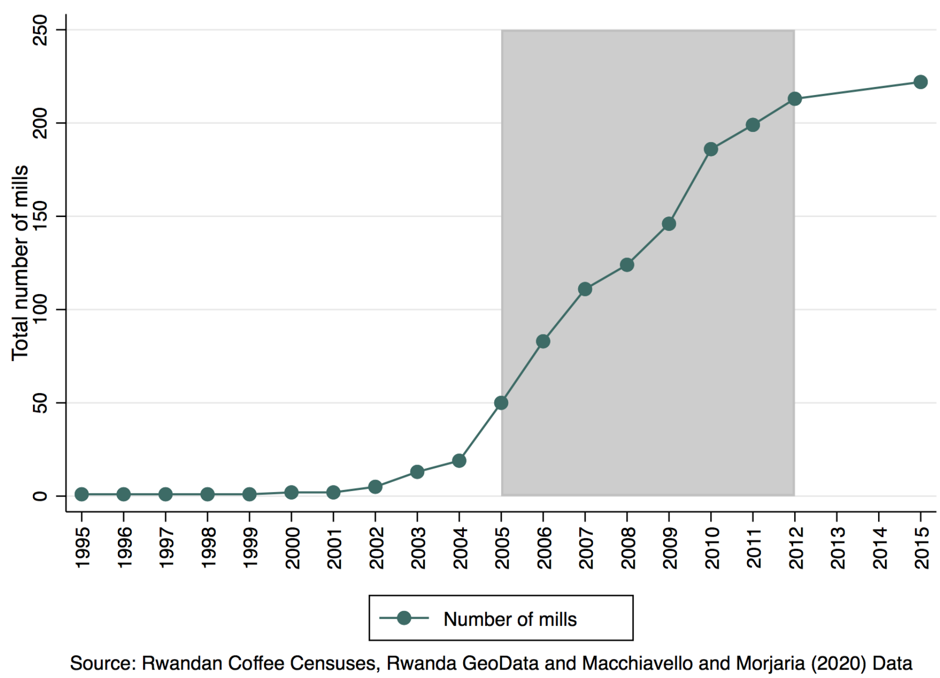


Figure 1: Mill Expansion in Rwanda

As it is shown in Figure 1, there is differential timing of mill openings in Rwanda. Moreover, there is geographical variation in mill locations. Figure 2 visualizes the spatial variation. The expansion of the mills are mostly concentrated in the areas that had a high number of coffee trees in 1999, before the adoption of National Coffee Strategy in 2002.

This paper exploits the differential timing of and spatial variation in exposure to a mill opening.

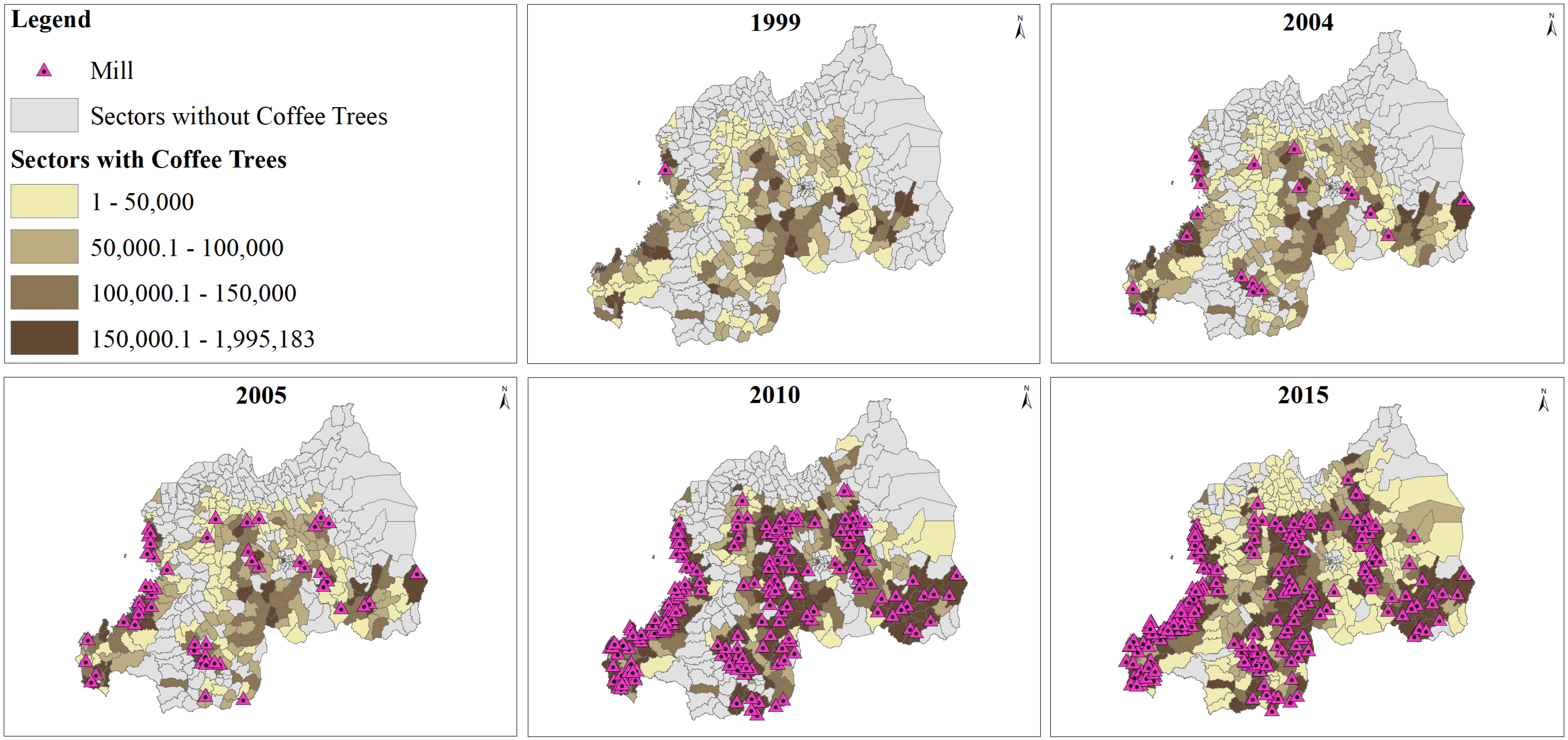


Figure 2: Expansion of Mills in Rwanda

Note: The maps are constructed by combining data on mills, The Rwandan Coffee Census 1999, 2003, 2009 and 2015, The Rwandan GeoPortal spatial data on sector boundries.

I first show that a mill opening at a specific location in a given year is uncorrelated with other determinants of changes in women's paid employment, domestic violence and female empowerment over time. In addition, I show that conditional on agricultural conditions related to mill suitability, the timing and location of a mill opening is random.

Mill Suitability. The number one condition for mill suitability is the area having a sufficient amount of coffee trees (Schilling and McConnell, 2004). This is mainly because the harvested coffee cherries should be transported to the mills within two hours of harvest. Otherwise they will rot (Schilling and McConnell 2004, Macchiavello and Morjaria 2020).[7] An overwhelming majority of the country is suitable for coffee cultivation.[8] Figure A.1 visualizes FAO-GAEZ coffee suitability index and confirms that the country is suitable for coffee production. Arabica coffee, the dominant plant type in Rwanda, grows best at an altitude over 1,000 meters (3,000 and 6,000 feet) and at an average temperature between 15 and 24 °C (59 and 75 °F). Nicknamed as "The Land of the Thousand Hills", Rwanda's hillsides provide the ideal conditions to cultivate coffee.

Figure 2 and Figure A.1 show that over time, the number of coffee trees increased also in the coffee suitable areas that did not have coffee trees in 1999. However, specifically the rapid expansion between 2005 and 2010 is concentrated in the areas where there were already coffee trees back in 1999. This is mainly because it takes at least 3-5 years for a newly planted coffee tree to produce coffee beans. Thus, a newly planted coffee tree that is planted the day the National Coffee Strategy is adopted will start to produce coffee cherries between 2005-2007.

In Section 5.1, I statistically investigate the determinants of a mill opening and present supporting evidence that conditional on agricultural conditions related to mill suitability, the timing and location of a mill opening is random.

2.3 Women's paid employment in the mills

In this subsection, I provide evidence that the expansion had a more pronounced effect on women compared to men.

The division of labor in the coffee value chain is gendered. This holds true not just for Rwanda but in coffee producing countries in the world in general. According to International Trade Forum's (ITC) survey that was conducted in 15 coffee producing ITC member countries -most of them are developing countries and Rwanda is one of them- women mostly take place in the earlier

[7] Most farmers carry their product to mills by walking.

[8] Coffee plants (*coffea*) is found in the tropical areas of Africa, South America and Asia, which is termed as "the coffee/bean belt". The belt provides the necessary conditions for the coffee plants to grow. Such conditions include specific ranges of temperature, rainfall, altitude and soil characteristics.

labor intensive steps in the value chain compared to the latter steps. Survey results is reported in Table 1. Women constitute 70% of the workforce in harvesting and processing. Only 10% of them are doing sale activities like in-country trading or exporting (Scholer, 2008). Drying and sorting are female-dominated tasks and marketing and selling the product are male-dominated. Tasks become male-dominated as coffee transforms from a commodity into a value-added product (SCAA Sustainability Council, 2015).

Tasks in the value chain	Participation Variation (Min-Max)	Typical Participation
Fieldwork	10-90%	70%
Harvest	20-80%	70%
Sorting	20-95%	75%
In-country trading	5-50%	10%
Export	0-40%	10%
Other (certification, laboratories)	5-35%	20%

Source: International Trade Forum, 2008

Table 1: Women's Participation in the Coffee Value Chain in Coffee Producing Countries

Before the expansion, the gendered division of labor generated a disparity in generating labor income in Rwanda. Although women worked in labor intensive stages, men receive and control the income. Women do not have their personal income and they depend on their husbands financially. With the expansion, mills demand seasonal wage labor for drying and sorting tasks. Women transition from being unpaid family workers to wage workers in the mills. What is unique about the context is that women are doing the exact same tasks as before the expansion, drying and sorting, but now they are paid for their labor by the mills and generate personal income. Since men were already generating income from their labor (selling the product), there is no change in their type of earnings. However, I will show that they generate a higher income from selling their product. Before the expansion, men sell dry-processed coffee beans to middlemen in the domestic market for a very low price. With the expansion, they sell their coffee cherries to mills for a higher price.[9] Thus, mill exposure is a plausibly exogenous increase in the ratio of women to their husbands' earnings. In the next section, I conceptualize the impact of an increase in women's and their husband's income on domestic violence in a simple theoretical framework to guide my empirical analysis.

[9]Macchiavello and Morjaria (2020) surveys the coffee farmers in Rwanda in 2019. The survey asks the relative profitability of the two processing methods. 98% of the farmers reported that selling cherries to mills are more profitable than home processing.

3 Theoretical Framework

To guide my empirical analysis, I introduce a simple conceptual framework below that shows how changes in women and their husband's income affect domestic violence.

3.1 Setup

Preferences. The household consists of a wife and husband, $j \in \{w,h\}$. The preferences of the man and the woman depend on their marital status. If they are partnered, I assume that the preferences can be represented by the utility functions

$$U_h = yI_h + (1-x+T)I_w + T\alpha_h(I_h,I_w) + \theta_h \quad and \quad U_w = (1-y)I_h + (x-T)I_w + \theta_w - v_w \qquad (1)$$

where I_j for $j \in \{w,h\}$ indicates personal earnings for the wife and husband. $I_h > I_w$, the husband earns more than the wife as in my data. According to Rwandan social norms, the husband controls household income and household decision-making (ICRW 2011, Bayisenge 2010, Ya-Bititi et al. 2019). The husband keeps yI_h of his personal income, $y \in (0,1)$, and gives $(1-y)I_h$ of it to the wife for her expenses. The wife keeps xI_w of her personal income, $x \in (0,1)$, and turns $(1-x)I_h$ of it to the husband. In my data, 65% of women decide how to spend their earnings jointly with their husband, not by themselves, which provides support for the model. x and y are predetermined in the beginning of the marriage. The main difference between the husband and the wife is that the husband can extract more money than his spouse turns over to him (based on the predetermined x value) by inflicting violence. Mainly, the husband can use violence to get a bigger share of the pie and maximize his utility from marriage. Thus, the violence is *instrumental* and used as a tool by the husband to extract resources from his wife. The extraction is parametrized by $T \in \{0,1\}$. If the husband chooses to inflict violence on his wife, $T = 1$. If he chooses not to be violent, $T = 0$.

$\alpha_h(I_h,I_w)$ denotes the husband's non-monetary utility from inflicting violence and it is decreasing in the husband's and wife's income, $\frac{\partial \alpha_h}{\partial I_h} < 0$ and $\frac{\partial \alpha_h}{\partial I_w} < 0$. It captures that an increase in household income provides reduction in his financial stress and thus, he receives less utility from inflicting violence to extract resources from the wife. θ_j for $j \in \{w,h\}$ indicates the private level of satisfaction with the marriage for the wife and husband. As a key assumption of the model, the level of satisfaction with the marriage remains private information for each spouse. θ_j follows distribution F_j with support $[\underline{\theta_j}, \overline{\theta_j}]$. $\underline{\theta_j}$ is high enough for the marriage to be intact. v_w is the disutility from violence for the wife. If the husband chooses to inflict violence, $v_w > 0$. If he chooses not to be violent, $v_w = 0$. U_w is decreasing in domestic violence and increasing in her and the husband's earnings and her private level of satisfaction with the marriage. U_h is increasing in his and his wife's earnings and his private level of satisfaction with the marriage.

If they are single, I assume that their preferences can be represented by the utility functions

$$S_h = I_h \quad and \quad S_w = I_w. \tag{2}$$

These represent the outside options. Each spouse only enjoys their personal incomes when they are single. The outside option is assumed to be the utility upon separation/divorce following Manser and Brown (1980) and McElroy and Horney (1981) rather than the noncooperative equlibrium as in Lundberg and Pollak (1993). This is because 15 percent of the sample is either divorced or not living together in my data. Moreover, in Sanin (2021), I study the adoption of domestic violence laws in Rwanda that enables women to divorce their husbands unilaterally if their husbands are violent to them. I show that the adoption increased the divorce rates. Specifically, the rates increased more in geographical units where women are more likely to be in violent marriages. This suggests that divorce is a credible threat in the context, specifically in case of domestic violence.

Timing. The husband keeps a portion of his personal income to himself and gives the remaining portion to the wife for her expenses. The wife keeps a portion of her personal income and turns the rest to her husband. The husband observes the resources he receives from his wife, $(1-x)I_w$, and either choose to inflict violence to extract more resources from the wife, TI_w, or he chooses not to be violent and remains with $(1-x)I_w$. If he chooses to inflict violence, his utility becomes $U_h = yI_h + (1-x+T)I_w + T\alpha_h(I_h, I_w) + \theta_h$ where $T = 1$. If he chooses not to, then $U_h = yI_h + (1-x)I_w + \theta_h$. Then, the wife either chooses to stay in the marriage or divorce the husband. If the husband chooses to inflict violence and wife remains married, her utility becomes $U_w = (1-y)I_h + (x-T)I_w + \theta_w - v_w$ where $T = 1$ and $v_w > 0$. In the absence of violence, her utility becomes $U_w = (1-y)I_h + xI_w + \theta_w$ and she stays in the marriage assuming $(1-y)I_h + xI_w + \theta_w > S_w$. The husband's utility is high enough to stay in the marriage in the absence of violence as well. When they are separated/divorced (single), the wife and the husband receive S_w and S_h respectively, their individual labor market potential. To highlight again, the level of satisfaction with the marriage remains private information for each spouse throughout the marriage.

Decisions. The solution of the game between the husband and wife can be found via backward induction. The wife chooses between staying in the marriage and divorcing the husband given the husband's decision to choose violence. She will divorce her husband if her utility of being married (where she experienced violence) is smaller than the utility of being divorced as in

$$\underbrace{(1-y)I_h + (x-T)I_w + \theta_w - v_w}_{\text{Utility of being in a violent marriage}} \leq \underbrace{I_w}_{\text{Utility of being single}}. \tag{3}$$

Based on Equation 3, there is a threshold value of the wife's private level of satisfaction with the marriage, $\overline{\theta_w}$, that makes her indifferent between remaining married and divorcing her husband. The value of $\overline{\theta_w}$ is determined by

$$\overline{\theta_w} = (1-x+T)I_w - (1-y)I_h + v_w. \tag{4}$$

Recall that the cumulative distribution function of θ_w is given by F_w. Therefore, the probability of the wife divorcing her husband is $P(\theta_w \leq \overline{\theta_w}) = F_w(\overline{\theta_w})$. Observe that as I_w increases, the probability of divorce increases. This is due to the fact that an increase in wife's personal income is improving her outside option. Also, when I_h increases, the probability of divorce decreases. This is because the wife enjoys a portion of the husband's personal income if she remains married.

Since the level of satisfaction with the marriage is private information, the husband chooses to inflict violence only knowing the probability that she will divorce him if her utility of being married is less than her utility of being divorced. He compares the expected utility of choosing violence, EU_h, with utility of not choosing to be violent. His expected utility of inflicting violence on his wife is

$$EU_h = F_w(\overline{\theta_w})I_h + [1-F_w(\overline{\theta_w})][yI_h + (1-x+T)I_w + T\alpha_h(I_h, I_w) + \theta_h]. \tag{5}$$

The husband chooses to be violent if

$$\underbrace{EU_h}_{\text{Expected utility of inflicting violence}} \geq \underbrace{yI_h + (1-x)I_w + \theta_h}_{\text{Utility from not inflicting violence}}. \tag{6}$$

$yI_h + (1-x)I_w + \theta_h$ captures the utility he receives when he chooses not to inflict violence and remains in the marriage with the money he receives from the wife. Based on Equation 6, there is a threshold value of the husband's private level of satisfaction with the marriage, $\overline{\theta_h}$, that makes the husband indifferent between choosing violence (to extract a portion of the wife's resources) and not being violent. The value of $\overline{\theta_h}$ is determined by

$$\overline{\theta_h} = (1-y)I_h - (1-x+T)I_w - T\alpha_h(I_h, I_w) + \frac{T[I_w + \alpha_h(I_h, I_w)]}{F_w(\overline{\theta_w})}. \tag{7}$$

Therefore, the probability of the husband choosing violence is $P(\theta_h < \overline{\theta_h}) = F_h(\overline{\theta_h})$. Differentiating it with respect to the wife's personal income, $\frac{\partial F_h(\overline{\theta_h})}{\partial I_w}$, highlights the impact of providing paid employment opportunities to women on domestic violence. Differentiating it with respect to the husband's personal income, $\frac{\partial F_h(\overline{\theta_h})}{\partial I_h}$, highlights the impact of an increase in husband's earnings on him inflicting violence. Both impact of an increase in I_w and I_h on $F_h(\overline{\theta_h})$ are both ambiguous.The effects still exist if decreasing marginal utility of consumption is assumed such as $U_h = \log(.) + \theta_h$ and $U_w = \log(.) + \theta_w - v_w$.[10]

Proof. See Appendix B.

3.2 Channels

Although the derivatives cannot be signed, I present the main channels that affect domestic violence below.

Increase in Outside Option. As the wife's earnings increases, her outside option improves. This increases the probability of her initiating divorce and makes it less likely for the husband to inflict violence on her.

Increased Incentives for Extraction. As the wife's earnings increases, the husband is incentivized to choose violence to extract the newly acquired resources from the wife.

Reduction in Financial Stress. As the husband's and/or women's earnings increases, his financial stress decreases. The husband would receive less non-monetary benefit from using violence to extract resources from the wife. This makes him less likely to inflict violence.

Benefit of Marriage. As the husband's earnings increases, the wife's benefit from the marriage increases. This makes her less likely to react domestic violence by leaving. This makes the husband more likely to choose violence to increase his share of resources.

For the couples who are exposed to a mill, the mill provides paid employment opportunities to the wife who was working as unpaid family workers. Thus, I_w increases. Since the husband receives a higher income from selling to the mill compared to selling to a middleman in the domestic market, there is also an increase in I_h. In the rest of the paper, I first estimate the causal impact of a mill on women's probability of experiencing domestic violence. Then, I investigate which of the aforementioned effect dominates the result empirically.

4 Data

4.1 Panel of Mills

I combine multiple data sources to create a panel of mills to exploit the differential timing of and spatial variation in exposure to a mill opening.

[10] Results available upon request.

Mills Data. The Rwanda GeoPortal provides data on the universe of mills as of 2014 in Rwanda. The data is geocoded and includes information on the characteristics of the mills like owner, number of drying tables etc. I combine the data with Macchiavello and Morjaria (2020)'s data that includes the year of operation for each mill between 1995-2012.[11] By combining the datasets, I create a panel of mills that has information on GPS coordinates, year of operation and characteristics of every mill in Rwanda.

Spatial Data, Coffee Census and FAO-GAEZ Suitability Index. First, I complement the mills data with spatial data by The Rwanda GeoPortal. The portal provides maps of different geographical boundries like sector, district and province. There are 416 sectors, 30 districts and 5 provinces in Rwanda. The mean area of each is approximately 58.4 km^2, 810 km^2 and 5552 km^2. I spatially match the mills data with the maps of different geographical boundries to find the sector, district and province a mill is located in. Then I match the data with several rounds (1999, 2003, 2009 and 2015) of Rwandan Coffee Census. The coffee censuses provide information on the universe of coffee trees in Rwanda at the sector level. I also match the mills panel with FAO's Global Agro-Ecological Zones (FAO-GAEZ) coffee suitability index for 1980-2010. The index provides a coffee suitability score for Rwanda at 9 km^2 resolution. I aggregated the index at the sector level.

The final panel of mills have information on a mill's location at different levels (including the exact GPS coordinates), number of coffee trees in the sector the mill is located for different years and FAO-GAEZ coffee suitability score. The data also has information on mill characteristics and spatial features of the location of the mill including presence of water bodies, road network, area etc. Figure 2 visualizes the finalized data. The years are selected based on the availability of the coffee census and data on women's outcome variables (2004, 2005, 2010 and 2014). The rapid expansion of the mills between 2005 and 2010 are observed in the areas that had a high number of coffee trees in 1999.

Table A.1 provides summary statistics. 78% of the mills started to operate between 2005-2010, the rapid expansion period. 50% of them are owned by cooperatives where only 25% of them are owned by NGOs. The remaning are owned by entepreneurs or private companies. Log number of coffee trees increased over time in the sectors where mills are built. 65% of the mills are built in areas where FAO-GAEZ coffee suitability index is either moderate, medium or good.

4.2 Individual and Household Level Data

Rwandan Demographic Health Surveys. As a first step of my analysis, I show the impact of the expansion of the mills on paid employment using the Rwandan Demographic Health Surveys

[11] The number of mills opened between 2012-2014 is very small. I coded those mills' year of operation as 2013.

(DHS). DHS are nationally representative, cross-section individual and household level surveys that are conducted in developing countries every 5 years. I use 2005, 2010/11 and 2014/15 cycles for my analysis. The surveys collect demographic and health information from women aged 15-49 and men aged 18-59. Such information includes marital status, employment, occupation, type of earnings, education, fertility and household decision-making. The data also provide household member information, which enable me to link couples to each other for my analysis.

The information on individuals' employment and type of earnings is collected retrospectively during the individuals' interview: individuals are asked to answer whether they are employed for the last 12 months and if so, whether they worked for cash. I create a binary variable which takes the value one if an individual worked for cash in the last 12 months. I also use other labor market outcomes including occupations to analyze whether exposure to mills induce a change in sectors.

Tables A.2 and A.3 provide descriptive statistics for women and their husbands. 88% of women worked in the past 12 months, where only 39% of them worked for cash. The share of working women is the same across areas that are exposed and not exposed to a mill. However, 47% of women are working for cash in mill exposed areas where this number is 37% for women in unexposed areas. The dominant occupation for women is agricultural where women are self-employed (farmers). 79% and 76% of women are working in agriculture in exposed and unexposed areas respectively. Agriculture is the dominant occupation for women's husbands as well. 73% and 71% of the husbands are working in agriculture in exposed and unexposed areas respectively. In contrast to women's outcomes, for husbands, the share of working and working for cash in the past 12 months are both similar across both areas.

The women's questionnaire also has a domestic violence module which collects information on family violence.[12] DHS randomly selects one women per household for the module. Thus, the number of women who have information for domestic violence will always be less than the number of women who answered the women's questionnaire. The module asks partnered women whether they experience physical, sexual or emotional violence in the last 12 months by their partners. If women's current marital status is widowed, divorced or separated, the module asks whether they experienced domestic violence in the last 12 months by their most recent partner.

Domestic violence categories (physical, sexual and emotional) are classified by DHS with respect to World Health Organization (WHO) guidelines. Since emotional violence questions are not asked in the 2010/11 cycle, I use physical and sexual domestic violence in my analysis. Physical domestic violence consists of being pushed, shaken, thrown something at, slapped, kicked, dragged, strangled, burnt and sexual domestic violence consists of physically forced into unwanted

[12] The module is prepared with respect to WHO guidelines "Putting Women First: Ethical and Safety Recommendations for Research on Domestic Violence against Women" World Health Organization, 2001.

sex and perform sexual acts. I create a binary variable which takes the value one if a partnered woman experienced physical or sexual domestic violence in the last 12 months. 34% of women self reported experiencing domestic violence in the past 12 months.

I also exploit variables related to household-decision making to analyze the relationship between exposure to mills and women's bargaining power within the household. DHS asks women whether they make certain decisions alone, jointly with their husband or the decisions are made for her by their husbands or someone else. Such decisions include decisions regarding large household purchases, women's own health and women visiting their own family.

The surveys also collect GPS coordinates for every cluster of households. Using the GPS coordinates, I spatially merge the DHS data with the mills data. One potential concern is that DHS randomly displace the GPS coordinates to maintain confidentiality. Due to the random displacement, urban clusters contain a minimum of 0 and a maximum of 2 kilometers of error and rural clusters contain a minimum of 0 and a maximum of 5 kilometers of positional error. Thus, GPS displacement may leads to measurement error and can bias the results (Perez-Heydrich et al., 2013). In order to reduce distance measurement error, I follow Perez-Heydrich et al. (2013), which suggests using a buffer distance rather than a closest distance when using a distance measure. All the distances calculated in the paper are based on a buffer distance measure. Thus, exposure to mill are constructed via creating buffers. Details on measuring exposure to mills are outlined at the end of the section.

DHS data allows me to observe the outcome variables for 2004, 2005, 2010 and 2014. The 2005 cycle was collected in February-July 2005, the 2010/11 cycle was collected in September 2010-March 2011 and the 2014/15 cycle was collected in November 2014-April 2015. Given that the harvest season runs from April to July and key variables captures individuals' experience in the past 12 months, 2005 data cycle contains information on both 2004 and 2005 harvest season. The 2010/11 cycle contains information on 2010 harvest season and the 2014/15 cycle provides information on 2014 harvest season. Thus, DHS data enables me to observe 4 harvest years in total.

I restrict the sample to women who are partnered (civil marriage and those who are living together) before the expansion of the mills to avoid changes in the marriage market matching.

Integrated Household Living Conditions Surveys. I complement my analysis using Integrated Household Living Conditions Survey (Enquête Intégrale sur les Conditions de Vie des Ménages, EICV). EICV is nationally representative, cross-section individual and household level surveys that are conducted in Rwanda. I use 2005, 2011 and 2013/14 cycles for my analysis.

The surveys collect demographic and socieconomic information from households. Such information includes consumption, income, agricultural activity employment and education. I use last daily labor income amount (cash) to further investigate the impact of the mill expansion on the labor market and uncover the mechanisms behind the effects on domestic violence. Table A.5 provides descriptive statistics for women and their husbands. On average women earn less daily labor income compared to their husbands. Women-to-husband daily labor income ratio is 0.65 in mill exposed areas where it us 0.55 in unexposed areas.

The surveys collect district of residence as the smallest geographical unit for each individual/household. There are in total 30 districts in Rwanda. I aggregate the panel of mills at the district level and merge it with individual and household level variables for my analysis. Similarly to DHS, EICV data allows me to observe the outcome variables for 2005, 2010 and 2014 harvest years.

I again restrict the sample to women who are partnered (civil marriage and those who are living together). Since EICV does not provide information on the year of marriage, unfortunately, I cannot restrict the sample to couples who married before the expansion. However, I include cohort fixed effects in my specifications that can provide a control for the year of marriage. Moreover, I restrict the sample to individuals who are in the same age groups with the DHS data.

4.3 Administrative Hospital Level Domestic Violence Data

In order to investigate the impact of providing employment opportunities to women on domestic violence, I also use confidential, administrative, geocoded data on the universe of public/district hospitals from the Rwandan Ministry of Health (MOH). Rwandan Health Management Information System (HMIS) data is a monthly district hospital level data on hospitalizations between January 2012 to December 2020. I focus on the years between 2012-2014 for my analysis.

There are in total 42 district hospitals in Rwanda. District hospitals constitute the overwhelming majority of the hospitals (47 in total) in the country.[13] The remaining 5 hospitals are either referral and teaching hospitals. Referral hospitals are located either in Kigali (capital) or urban areas, have the latest technology and provide care under the public national health insurance system if a district hospital or health clinic refers a current patient to visit a referral hospital. Thus, they do not constitute a first stop for a patient. Teaching hospitals focus on medical research.

Due to the institutional structure of the hospitals, the data is on the universe of official domestic violence reports in Rwanda. Figure A.2 visualizes the hospitals on a map. The district hospitals have gender based violence centers in them which aim to combat gender based violence in Rwanda.

[13]Given that there are in total 30 districts, some districts have more than one district hospital.

These centers, also known as the Isange One Stop Centers (IOSCs), are state institutions which provide medical, police and legal help for gender based violence victims under one roof and are free of charge. By combining all services under one roof, MOH, Rwanda National Police, the Ministry of Justice (MOJ) and the Ministry of Gender and Family Promotion (MIGEPROF) aim to ease the reporting process of gender based violence victims. Sviatschi and Trako (2021) finds that such centers in Peru reduced gender-based violence and increased human capital investments in children, raised children's enrollment, attendance, and test scores.

The data collect information on the number of individuals (both women and men) who show symptoms of physical and sexual violence for different age groups (10-18, older than 18). Unfortunately the data does not provide information on the marital status of the patient. In order to create a measure of domestic violence, I focus on the gender based violence reports of individuals who are older than 18 years old. This is because 70% of women over 18 are married in Rwanda.

I construct a binary variable coded as 1 if a hospital had hospitalizations due to physical or sexual violence for women older than 18 in a month and 0 otherwise. This creates a non self-reported measure of domestic violence. Geocoded nature of the data enable me to match this measure of domestic violence with the panel of mills. Since the data is at the monthly level, I analyze the impact of the mills on domestic violence by exploiting the timing of the harvest season within a year. Months between March-July are defined as the harvest season for my analysis, where May-July is defined as the peak of the harvest.

Table A.4 provides descriptive statistics for the hospitals. 79% of the hospitals have hosted at least a domestic violence patient (GBV patient older than 18) in a given month. The share is 70% for the hospitals who are exposed to the mills where it is 85% for the ones who are not exposed. Mean number of patients with physical domestic violence and sexual domestic violence symptoms in a month in a hospital is 2.52 and 1.71 respectively. The numbers are 1.56 and 1.16 for the hospitals who are exposed to the mills. For the hospitals who are not exposed to the mills, the mean number of patients with physical domestic violence and sexual domestic violence symptoms in a month are 3.07 and 2.02 respectively.

4.4 Measuring Exposure to Mills: Treatment and Control Groups

Throughout the paper, the treatment group consists of entities (woman/husband/hospital) who are exposed to the mills where the control group consists of the entities who are not. I construct two measures of exposure to the mills, being in the catchment area of a mill and mills per capita in a district. I also construct multiple control groups with respect to the measures I create.

The first mill exposure measure, being in the catchment area of a mill, is used when the outcome variables are created with geo-coded datasets (DHS and HMIS). The measure uses the GPS coordinates of the mills and DHS clusters/hospitals and calculates whether the clusters/hospitals are within the catchment area of a mill. In most cases, mills are located within a radius of 3–5 km away from the farmers (AgriLogic, 2018). On average, managers of the mills in Rwanda report that the catchment areas have a radius of approximately 4.5 km (Macchiavello and Morjaria, 2020). I create a buffer with 4 km radius centered around the mills to construct the catchment area (treatment group). I define a DHS cluster/hospital within the catchment area of a mill, if the GPS coordinate of the DHS cluster/hospital falls inside the buffer around a mill.

When using DHS data, I consider 2 potential control groups in my analysis. The first control group consists of the DHS clusters that are outside of the catchment area of the mills (4 km buffer) but are located within the same district with the mills. I define using this control group as the "within district approach". An alternative control group consists of only the DHS clusters that are within the donut area between 4 and 8 km from the mills. Since I restrict the control group only to DHS clusters that are within the surface area of a donut, I define using this control group as the "donut approach". Figure 3 shows a visual representation of the treatment and the 2 control groups using a mill in Nyarugenge district and DHS clusters. In both maps, the orange area is the treatment group. In the top map, which visualizes the within district approach, everywhere outside the orange buffer characterizes the control group. In the bottom map, which visualizes the donut approach, only the green area constitutes the control group.

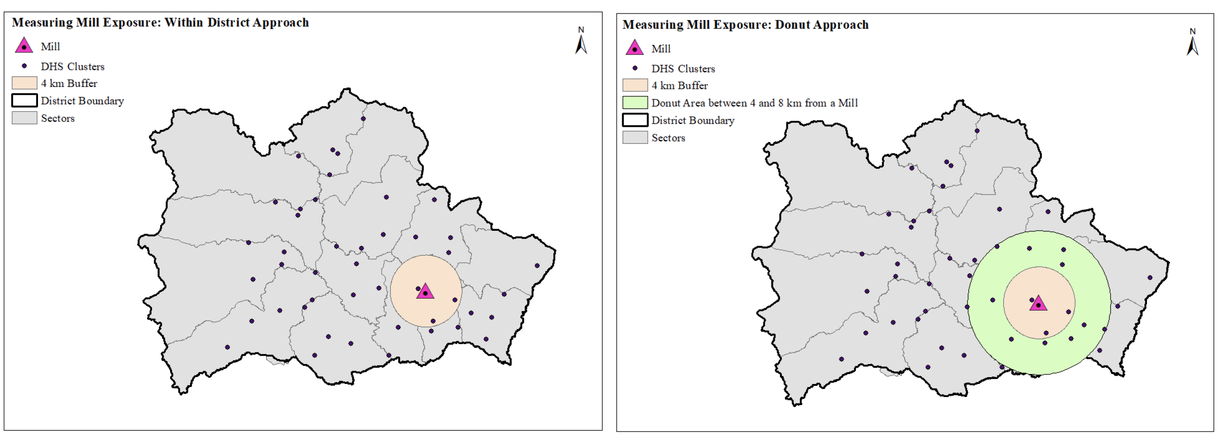


Figure 3: Visualization of the Treatment and Control Groups using Nyarugenge District: Within District and Donut Approach

For HMIS (hospital) data, the control group consists of the hospitals that are outside of the catchment area of the mills but are located within the same province with the mills.This choice is due to the sample size and locations of the hospitals. There are only 42 hospitals in the country and majority of the hospitals who are not in the catchment area of a mill are located outside of the same district with the mill as well as the donut area between 4 and 8 km from the mill.

The second mill exposure measure, mills per capita, is used when the outcome variables are created with a dataset that is not geo-coded (earnings data, EICV). The measure is the number of mills in a district in a given year divided by the number of working age individuals living in the district in that year.[14]

5 Empirical Specification

In this section, I propose my identification strategy to estimate the causal impact of the expansion of the coffee mills on women's labor market outcomes and domestic violence. I use a difference-in-differences strategy that exploits the differential timing of and spatial variation in mill openings.

5.1 Identifying Assumptions and Threats to Identification

The key identifying assumption of my empirical strategy is that the average outcomes for the treatment and control groups would have parallel trends in the absence of treatment, a mill placement. A mill opening at a specific location in a given year is assumed to be uncorrelated with other determinants of changes in the outcome variables (women's labor market outcomes and domestic violence) over time (the treatment is not endogenous). If a mill placement is endogenous, then the parallel trends assumption is violated. This is because in counterfactual, the areas who are exposed to the mill would have diverged anyway, regardless of the mill.

To provide suggestive evidence in favor of my identifying assumption, I estimate, at the sector level, the determinants of having the first mill during the period when mills expanded rapidly, between 2005-2011, and having a mill by the end of my sample, 2014. Firstly, I focus on the variables that are related to coffee cultivation including the historical number of coffee trees and FAO-GAEZ coffee suitability index. Then, in order to confirm that mill placement is uncorrelated with factors that can affect the evolution of female paid employment and domestic violence rates in Rwanda, I included the share of women who are unpaid family workers and completed primary school at the sector level before the expansion of the mills in the regression. I use the Rwanda Population and Housing Census 2002 to create the variables. I also included total population and female population at the sector level. Unfortunately, I am unable to include domestic violence before the expansion in the regression, since domestic violence data started to be collected in

[14] Individuals who are aged 19-59 are selected in EICV to match the age profile in the DHS sample.

2005, after the expansion started. However, since education are correlated to the probability of experiencing domestic violence, including it in the regression constitutes a test on whether the placement of the mills are correlated with the evolution of domestic violence rates in Rwanda.

I also include other variables related to promoting female empowerment including share of women in a consensual union, polygamous marriage, without assets and number of daughters per woman. If these variables are uncorrelated with mill placement, then, it provides supporting evidence that the mills are not placed to promote female empowerment and hence mill placement is not endogenous. I also include genocide intensity since it is also a variable that is found to impact the probability of experiencing domestic violence in Rwanda (La Mattina, 2017). It is also possible that the government or NGOs are more likely to financially support the opening of a mill in genocide intense areas to promote female empowerment that can make the opening of a mill endogenous. District fixed effects are included in all specifications.

I report the results in Table 2. The results provide support for the evidence that variables related to coffee, the number of coffee trees in 1999 and FAO-GAEZ coffee suitability predicts mills placement. Specifically, sectors that have a higher number of coffee trees in 1999 are more likely to have their first mill between 2005-2010. Those areas are also more likely to have a mill by 2014. Similarly, sectors with a higher FAO-GAEZ coffee suitability index are more likely to have their first mill between 2005-2010. Based on results, historical number of coffee trees are more likely to be correlated with the rapid expansion of the mill compared to FAO-GAEZ coffee suitability index. This is not surprising. Before the expansion, not every coffee suitable area in Rwanda had coffee trees and it takes minimum 3-5 years for a coffee tree to produce coffee cherries. Thus, it is likely to observe the rapid expansion of the mills between 2005-2010 in the areas that already have coffee trees before the expansion, like 1999.

All variables other than the number of coffee trees in 1999 and FAO-GAEZ coffee suitability index are statistically insignificant. Specifically, pre-expansion female unpaid employment and female primary education rates do not predict mill placement. None of the variables related to female empowerment is statistically significant. Genocide-intensity does not have a statistically significant impact on mill expansion. These results suggest that mills are not placed to promote female empowerment. They are uncorrelated to the factors that can have an impact on the evolution of female paid employment and domestic violence over time. Thus, conditional on mill suitability based on coffee related variables, opening of a mill is assumed to be random. I control for historical number of coffee trees and FAO-GAEZ coffee suitability index in all of my specifications in the paper.

In Section 7, I also perform an event-study analysis for all of my main outcome variables that exploits the variation in the number of years couples are exposed to the mill. In the event-study analysis, the patterns of outcome variables before a mill opening is investigated. I provide evidence that there is no pre-trends in outcomes.

5.2 Paid Work and Self-Reported Domestic Violence

In order to investigate the impact of a mill opening on women's and their husbands' outcomes, I estimate the specification below:

$$Y_{ist} = \beta_0 + \beta_1 Mill_{ist} + \mathbf{X_{ist}}\phi + \lambda_c + \omega_m + \alpha_s + \gamma_{dt} + (\mathbf{X_s} \times t)\theta + \varepsilon_{ist}. \tag{8}$$

The dependent variable Y_{ist} is the outcome of interest of woman i (or the husband of woman i), in sector s and at year t. $Mills_{ist}$ is a binary variable coded as 1 if woman/husband i at year t resides within the catchment area of a mill and zero otherwise. I have a rich set of individual controls, $\mathbf{X_{ist}}$, that includes information on women's/husbands' occupation, education, religion, number of children, marital status, age and duration of marriage, residence (rural/urban) and household wealth. I also control for partner characteristics like partner's age, occupation and education. λ_c is the cohort fixed effects and controls for factors that vary across cohorts. As an example, compared to younger cohorts, older cohorts grew up during a period when Rwanda did not have pro-women laws which may affect their labor market participation and acceptance of domestic violence. ω_m is the year of marriage fixed effects. It controls for time-variant shocks to the marriage market such as the Rwandan Genocide that is documented to affect marriage quality in Rwanda (La Mattina 2017, Sanin 2021).[15] α_s is the sector fixed effects and controls for time-invariant local observable and unobservable characteristics, such as social norms related to women's employment, domestic violence and gender roles.

γ_{dt} is the district-by-year fixed effects. I allow time fixed effects to vary by district. This is to control for factors that change over time and across districts and may determine both a mill opening and female empowerment such as female political participation. As of 2008, Rwanda is the first country in the world with a female majority in parliament. The share of women in local government varies across districts and increases over time. Districts (akarere) are the geographical units with the highest tier in local government. District councils decide on local development programs within districts. As a possible concern, after an increase in the female political representation in a district

[15] La Mattina (2017) shows that the timing of marriage (before or after the genocide) has an impact on probability of experiencing domestic violence in Rwanda. Sanin (2021) investigates the impact of the adoption of the domestic violence legislation in Rwanda in 2008, which allows women to unilaterally divorce their husbands if their husbands are violent towards them. After the law, among the women who married after the genocide, the divorce rates increase more and sexual domestic violence rates increase less in the formerly genocide-intense areas, where women are more likely to be in violent marriages. No effect is observed for women who married right before the genocide.

Table 2: Sector Level Baseline Characteristics that Predict Mill Opening

	(1) First Mill in 2005-2010	(2) Mill by 2014
Log Coffee Trees in 1999	0.03***	0.04***
	(0.01)	(0.01)
FAO-GAEZ Coffee Suitability Index	0.07*	0.04
	(0.04)	(0.04)
Log Population in 2002	-1.01	-0.74
	(0.73)	(0.74)
Log Female Population in 2002	0.97	0.79
	(0.73)	(0.74)
Share of Self-Employed Women in 2002	0.44	0.67
	(0.55)	(0.55)
Share of Unpaid Worker Women in 2002	0.39	0.90
	(0.58)	(0.59)
Share of Primary-Educated Women in 2002	1.34	0.87
	(1.01)	(1.02)
Number of daughters per Woman in 2002	-0.20	-0.15
	(0.22)	(0.22)
Share of Women in a Consensual Union in 2002	-1.08	-0.59
	(0.77)	(0.78)
Share of Women in a Polygamous Marriage in 2002	-0.26	-1.34
	(1.92)	(1.94)
Share of Women without Assets in 2002	0.47	-0.25
	(0.59)	(0.60)
Age at Genocide among Women in 2002	-0.08	-0.05
	(0.05)	(0.05)
Genocide Intensity Index at the Commune Level	0.00	-0.03
	(0.04)	(0.04)
District FE	✓	✓
Number of Observations	348	348
Dependent variable mean	0.28	0.33
Adjusted R^2	0.27	0.33

Note: FAO-GAEZ coffee suitability and genocide intensity index are both standardized. The data is at the sector level. FAO-GAEZ index is originally defined at the 9 km^2 resolution. I aggregated the index at the sector level (on average 50 km^2 area) for analysis. *** p<.01, ** p<.05, * p<.1

council, the local government may financially support cooperatives who want to open mills to promote female paid employment. District-by-year fixed effects control for such a scenario and avoids ommitted variable bias. $\mathbf{X_s}$ is a vector of baseline geographical variables at the sector level such as the historical number of coffee trees in 1991 and FAO-GAEZ coffee suitability index. I interact these initial conditions with linear time trends to allow their impact vary over time. The interaction mitigates potential omitted variable bias. This is because coffee tree presence in 1999 is correlated with mill openings after 2002 and it may also effect the evolution of female paid employment over time. ε_{ist} is the error term. I cluster standard errors at the sector level.

The main dependent variables for this specification are being employed, working for cash and experiencing domestic violence in the past 12 months. All outcomes are indicator variables. As an example, working for cash in the past 12 months variable takes the value one if the respondent worked for cash in the past 12 months and 0 otherwise. Domestic violence in the past 12 months variable is a combination of physical and sexual domestic violence. It takes the value one if a partnered woman experienced physical or sexual domestic violence in the past 12 months and 0 otherwise. I also use dependent variables like occupations (manager, sales, agricultural self-employed, agricultural employee, skilled manual, unskilled manual) and household decision-making (whether large household purchases are decided by women alone or jointly with husband compared to husband alone).

The coefficient of interest is β_1, which identifies the impact of a mill opening on the outcome variables. The treated group consists of woman/husband who reside in the catchment area of a mill. The control group either consists of woman/husband who do not reside in the catchment area of a mill within the same district (within district approach) or woman/husband who reside within the donut area between 4 and 8 km from a mill (donut approach). Tables 3 and 4 examine whether predetermined covariates including women's agricultural occupation, education, religion, marriage type, age at first marriage, partner characteristics and living conditions (electricity, cement floor in the household) are balanced across treatment and control groups. I perform the test by estimating the specification given in Equation 8 using each of the predetermined covariates as the dependent variable. None of the estimates are statistically significant which suggests that the baseline characteristics are balanced across the treatment and each control group.

I also perform a placebo test to show that outcome variables are balanced across treatment and control groups before a mill opening. For this test, I drop the women who reside within the catchment area of a mill in a given year (2004, 2005, 2010 and 2014) from the sample. I construct the treatment group as women who live in the areas that do not have a mill yet. The areas will receive a mill and become catchment areas in the upcoming years. Thus, I am *falsely assuming* that the women who reside in those areas are exposed to a mill and in the treatment group. The

Table 3: Balance Check: Within District Approach

	Husband		Women		
	(1) Occupation: Agricultural	(2) Education in Years	(3) Occupation: Agricultural	(4) Education in Years	(5) Civil Marriage
Mill	-0.00	0.10	-0.01	-0.08	-0.00
	(0.02)	(0.15)	(0.02)	(0.14)	(0.02)
Observations	9823	9823	9823	9823	9823
Dependent variable mean	0.71	4.34	0.76	4.02	0.73
	Women		Household		
	(1) Age at: First Marriage	(2) Religion: Christian	(3) Residence: Rural	(4) Cement Floor	(5) Electricity
Mill	0.01	-0.00	0.02	-0.00	0.01
	(0.06)	(0.01)	(0.03)	(0.01)	(0.01)
Observations	9823	9823	9823	9823	9823
Dependent variable mean	19.90	0.96	0.83	0.17	0.12

Note: Robust standard errors clustered at the sector level are in parentheses. 4 km catchment area is used for the treatment group. Within district approach is used for the control group. The estimates are based on DHS data and estimated with the main specification presented in Section 5.2.1
*** p<.01, ** p<.05, * p<.1

Table 4: Balance Check: Donut Approach

	Husband		Women		
	(1) Occupation: Agricultural	(2) Education in Years	(3) Occupation: Agricultural	(4) Education in Years	(5) Civil Marriage
Mill	-0.01	-0.02	-0.01	-0.18	-0.01
	(0.02)	(0.17)	(0.02)	(0.17)	(0.02)
Observations	4900	4900	4900	4900	4900
Dependent variable mean	0.70	4.46	0.75	4.41	0.79
	Women		Household		
	(1) Age at: First Marriage	(2) Religion: Christian	(3) Residence: Rural	(4) Cement Floor	(5) Electricity
Mill	-0.01	0.00	0.04	-0.00	0.02
	(0.07)	(0.01)	(0.03)	(0.01)	(0.01)
Observations	4900	4900	4900	4900	4900
Dependent variable mean	20.15	0.96	0.81	0.20	0.16

Note: Robust standard errors clustered at the sector level are in parentheses. 4 km catchment area is used for the treatment group. Donut approach is used for the control group. The estimates are based on DHS data and estimated with the main specification presented in Section 5.2.1 *** $p<.01$, ** $p<.05$, * $p<.1$

control group consists of women who live in the areas outside of the future catchment areas. I again use both the within district and donut approaches. Women in both groups self-report their labor market outcomes and domestic violence experience for a given year before a mill opening. Since both groups are not exposed to a mill, I expect to see no statistical differences between them. I perform the placebo test by estimating the specification given in Equation 8 using the aforementioned treatment and control groups. None of the estimates are statistically significant which suggests that the outcome variables are balanced across treatment and control groups before a mill opening.

Table 5: Placebo Test: Outcome Variables Before a Mill Opening

	Within District			Donut		
	(1) Work	(2) Cash	(3) Violence	(4) Work	(5) Cash	(6) Violence
Mill	0.01	-0.02	-0.00	0.00	-0.03	0.01
	(0.01)	(0.04)	(0.07)	(0.01)	(0.04)	(0.07)
Observations	6472	6475	2717	3105	3106	1348
Dependent variable mean	0.99	0.36	0.35	0.98	0.36	0.34

Note: Robust standard errors clustered at the sector level are in parentheses. 4 km catchment area is used for the treatment group. The estimates are based on DHS data and estimated with the main specification presented in Section 5.2.1 *** p<.01, ** p<.05, * p<.1

5.3 Earnings

I also test the impact of a mill on labor income (women and their partners' each). Unfortunately the data on these variables is not geo-coded and the smallest geographical unit in the data is the district. Thus, in order to estimate the impact of the mills on those variables, I employ the empirical specification below:

$$Y_{idt} = \beta_0 + \beta_1 Mill_{idt} + \mathbf{X_{it}}\phi + \lambda_c + \alpha_d + \gamma_{dt} + (\mathbf{X_d} \times t)\theta + \varepsilon_{idt}. \tag{9}$$

The dependent variable Y_{idt}, is the logarithm of daily labor income of partnered woman/husband i, in district d and at year t for estimating the impact on labor income. I use logarithm of women and their husband's earnings as well as the ratio of their earnings as dependent variables. $Mills_{idt}$ is the logarithm of the total number of mills per capita in the district of residence of a woman/husband i

at year t. $\mathbf{X_{idt}}$ is the same set of controls in the main specification. λ_c is the cohort fixed effects.[16] α_d is the district fixed effects. γ_{dt} is the district-by-year fixed effects. $\mathbf{X_d}$ is a vector of baseline geographical variables at the district level including the historical number of coffee trees in 1999 and FAO-GAEZ coffee suitability index. I interact these initial conditions with linear time trends, to allow their impact vary over time. ε_{idt} is the error term. I cluster standard errors at the district level.

The coefficient of interest is β_1, which identifies the impact of an increase in log mill per capita on the outcome variables.

5.4 Monthly Hospitalizations due to Domestic Violence

A unique feature of this paper is that I investigate the relationship between paid work opportunities and domestic violence using both annual self-reported and monthly administrative data. Using the universe of monthly hospitalizations due to domestic violence in Rwanda, I test whether a mill affects domestic violence when it is in operation. This variation is due to the fact that a mill is operating only during the harvest period: March-July. I use a specification at the hospital level. To estimate the impact of a mill on hospitalizations due to domestic violence, I estimate the empirical specification below:

$$Y_{hdtm} = \beta_0 + \sum_{m=1}^{12} Mill_{hd} \times \beta_m \mathbb{1}[\tau = m] + \mathbf{X_{ht}}\phi + \lambda_h + \alpha_d + \sigma_m + \gamma_{pt} + (\mathbf{X_d} \times t)\theta + \varepsilon_{hdtm}. \quad (10)$$

The dependent variable Y_{hdtm} is the monthly hospitalization outcome due to domestic violence in hospital h, in district d, in year t and at event-time m. $Mills_{hd}$ is a binary variable coded as 1, if hospital h is within the catchment area of a mill and zero otherwise during the sample period, 2012-2014. Being in the catchment area of a mill did not change between 2012-2014 for the hospitals. Thus, $Mill_{hd}$ is a time-invariant characteristic. It is interacted with event-month dummies, $\mathbb{1}[\tau = m]$, to investigate the dynamic impact of a mill during the harvest period (March-July), the months mills operate. τ denotes the event-month. $\tau = 3$, March, represents the month harvest period begins and mills start to operate. For $3 \leq m \leq 7$, March-July, $\tau = m$ represents the months mills are operating. For $m < 3$, $\tau = m$ represents the months before a mill's month of operation. The omitted category is $\tau = -1$, February, which means that the dynamic impact of being exposed to a mill is estimated with respect to one month prior to a mill's month of operation.

$\mathbf{X_{ist}}$ is the set of hospital level time-varying controls related to gender based violence specific

[16] I am unable to control for year of marriage fixed effects since the data does not have such information. However, I am controlling for cohort fixed effects.

hospital quality. It includes information on whether gender based violence patients are referred for care to a health facility with higher level of resources. It is a proxy for poor hospital quality. $\mathbf{X_{ist}}$ also includes information on whether gender based violence victims are referred to the hospital by the police and community health workers. They are proxies for high hospital quality.[17] λ_h is the hospital fixed effects which controls for any hospital specific characteristic that is fixed over time like its location. α_d is the district fixed effects. District is chosen for the level of geographical unit since the unit of observation is a district hospital. σ_m is the month fixed effects and controls for month specific trends. γ_{dt} is the province-by-year fixed effects. I allow year fixed effects to differ by province, one unit higher than the district.[18] This way, I am comparing the hospitals who are in the catchment area of a mill to the ones that are not, within the same province. Hospitals who are not within the catchment area of a mill within the same province constitutes a more accurate control group. $\mathbf{X_d}$ is the vector of baseline geographical variables at the district level interacted with linear time trends. ε_{hdtm} is the error term. I cluster standard errors at the district level.

The main dependent variables for this specification are whether the hospital had hospitalizations and deaths due to gender based violence for women/men older than 18 and aged 10-18. Variables that focus on individuals older than 18 capture domestic violence. This is because according to DHS and census data, majority of the individuals who are older than 18 are married in Rwanda. Variables that focus on individuals aged 10-18 capture gender based violence not due to a partner. Moreover, that age group do not work in the mills. Such variables are used as placebo outcomes as a robustness check.

All outcomes are indicator variables. As an example, monthly hospitalizations due to gender based violence for women older than 18 variable takes the value 1 if a hospital had a hospitalization due to gender based violence for a female victim older than 18 in a given month and 0 otherwise. Dummy variables rather than logarithm of hospitalizations are used since the number of hospitalizations for gender based violence is low in a month.

The coefficients of interest are β_m's when $3 \leq m \leq 7$ (March-July). For $3 \leq m \leq 7$ and aged older than 18 category, each β_m provides the change in hospitalizations due to domestic violence within the cathment area of a mill relative to the hospitalizations that are not, within the same province, during a mill's month of operation, relative to one month before the operation. The harvest season is at peak during $5 \leq m \leq 7$, May-July, where majority of the farmers work in the

[17] Being a hospital that is referred to a gender based violence victim by the police and community health workers may show that the hospital is known for the resources and knowledge related to gender based violence cases. Moreover, referrals may also show that the hospital is in collaboration with the police and community health workers.

[18] Unfortunately there is not enough observations/hospitals to have district-by-year fixed effects. Adding district-by-year fixed effects mean comparing the hospitals who are in the catchment area of a mill to the ones that are not within the same district. Most of the districts have only one district hospital.

mills. Thus, I specifically expect to see statistically significant estimates during those months. Since $m = 1$ (January) and $8 \leq m \leq 12$ (August-December) respectively represent the months before and after the harvest season when the mills do not operate, I expect to see small and statistically insignificant estimates for those months.

6 Results

6.1 Employment and Self-Reported Domestic Violence

I first estimate the impact of mill exposure on women's employment, type of earnings and self-reported domestic violence using Equation 8. Table 6 presents the results of estimating the impact of exposure to a mill on women's probability of working in the past 12 months, her type of earnings being cash in the past 12 months and self-reporting domestic violence in the past 12 months using different control groups. To measure mill exposure 4 km catchment area is used. The first three columns represents estimating Equation 8 when the control group is defined as DHS clusters that are outside of a mill's catchment area but within the same district the mill is located (within district approach). Columns 4-6 represents estimating Equation 8 when the control group consists of DHS clusters that are within the donut area between 4 and 8 km from a mill (donut approach). All estimations include individual controls, cohort fixed effects, year of marriage fixed effects, sector fixed effects, district-by-year fixed effects, linear time trends interacted with baseline sector level characteristics. Sample consists of partnered women who married before the expansion of the mills.

When within district approach is used for the control group, being exposed to a mill increases the probability of working for cash in the past 12 months by 7 percentage points (p-value= 0.001) where the probability of working in the past 12 months remain unchanged. The estimated impact represents an increase of 18% with respect to the sample mean (0.40). Moreover, being exposed to a mill decreases the probability of experiencing domestic violence in the past 12 months by 9 percentage points (p-value= 0.01). The estimated impact represents a decrease of 26% with respect to the sample mean (0.35).

When using the donut approach, being exposed to a mill increases the probability of working for cash in the past 12 months again by 7 percentage points (p-value= 0.003) where the probability of working in the past 12 months remain unchanged. The estimated impact represents an increase of 16% with respect to the sample mean (0.45). Moreover, being exposed to a mill decraeses the probability of experiencing domestic violence in the past 12 months by 8 percentage points (p-value= 0.05). The estimated impact represents a decrease of 22% with respect to the sample mean (0.37). In summary, results continue to be statistically significant and very similar even when the control group is restricted only to the clusters within the donut area between 4 and 8 km from a mill. There is a statistically significant increase in the probability of working for cash and a decline in the probability of experiencing domestic violence for women.

Table 6: Effect of Mill Exposure on Women's Employment, Type of Earnings and Self-Reported Domestic Violence in the Past 12 Months

	Within District			Donut		
	(1) Work	(2) Cash	(3) Violence	(4) Work	(5) Cash	(6) Violence
Mill	-0.00	0.07***	-0.09***	-0.00	0.07***	-0.08**
	(0.01)	(0.02)	(0.03)	(0.01)	(0.02)	(0.04)
Observations	9823	8766	3583	4900	4415	1658
Dependent variable mean	0.88	0.40	0.35	0.88	0.45	0.37

Note: Robust standard errors clustered at the sector level are in parentheses. All estimations include individual controls, cohort fixed effects, year of marriage fixed effects, sector fixed effects, district-by-year fixed effects, linear time trends interacted with baseline sector level characteristics. Sample consists of partnered women who married before the expansion of the mills. All dependent variables are measured for the past 12 months. The catchment areas is constructed by buffers around the mills with a 4 km radius. *** p<.01, ** p<.05, * p<.1

Table 7: Effect of Mill Exposure on Husband's Employment and Type of Earnings in the Past 12 Months

	Within District		Donut	
	(1) Work	(2) Cash	(3) Work	(4) Cash
Mill	-0.00	0.03	-0.01	0.02
	(0.02)	(0.02)	(0.02)	(0.02)
Observations	4099	3548	2168	1961
Dependent variable mean	0.87	0.81	0.90	0.82

Note: Robust standard errors clustered at the sector level are in parentheses. All estimations include individual controls, cohort fixed effects, year of marriage fixed effects, sector fixed effects, district-by-year fixed effects, linear time trends interacted with baseline sector level characteristics. Sample consists of the husbands of the women in the sample. All dependent variables are measured for the past 12 months. The catchment areas is constructed by buffers around the mills with a 4 km radius. *** p<.01, ** p<.05, * p<.1

Table 7 presents the results of estimating the impact of exposure to a mill on women's husbands' probability of working in the past 12 months and their type of earnings being cash in the past 12 months for both control group approaches. There is no statistically significant change in those variables. I also estimated the impact of mill exposure on women's and their husbands' occupation. It is possible that a mill opening can induce individuals to sort into the agricultural sector. Results are reported in Tables A.6, A.7, A.8 and A.9. The first two tables use the within district approach and the last two use the donut approach. Each column represents one occupation where the occupations are managers, sales, agricultural self-employed, agricultural employee and skilled and unskilled manual. For women, exposure to a mill do not induce a change in occupations. The dominant occupation is the agricultural self-employed category where 80% of women have such occupation. Being a farmer constitutes an example of working as self-employed in agriculture. For husbands, there is a statistically significant increase in the sales category. However, there is not a statistically significant impact of mill exposure on husbands being self-employed in the agricultural sector. Being self-employed in the agricultural sector is also the dominant occupation for husbands where 66% of the husbands have such occupation. Overall, individuals who are exposed to a mill do not sort into being self-employed in agriculture (become farmers).

According to the results, women and their husbands were already self-employed in the agricultural sector (farmers). They were already working in the past 12 months. Being exposed to a mill, did not induce a shift on these variables. However, exposure to a mill changed the type of earnings of women. Women start to be paid cash for their labor in agriculture. Their husbands were already working for cash and being exposed to a mill did not change their type of earnings.

6.2 Earnings

I also estimate the impact of mill exposure on women's and their husbands' earnings using Equation 9. Table 8 presents the results of estimating the impact of mill exposure on women's and their husbands' log labor income. Since EICV data is not geocoded, mill exposure is defined as the logarithm of mills per capita in a district. Columns 1-3 present the results for the whole sample. Columns 4-6 present the results among couples working in agriculture.

Among all sample, 1% increase in mills per capita increases women's and their husbands' daily labor income by 7% each. Among couples working in agriculture, an increase in mills per capita also increases each spouse's labor income. These results show that mill exposure increases women's and their husbands' earnings. As an additional information, in columns 3 and 6, I show that among all women and couples working in agriculture, 1% increase in mills per capita increases women's relative earnings in the household (the ratio of women and their husbands' log labor income) by 0.4% and 0.3% respectively.

Table 8: Effect of Mill Exposure on Log Daily Labor Income

	All Sample			Occupation: Agriculture		
	(1) W	(2) H	(3) W/H	(4) W	(5) H	(6) W/H
Mill	7.37***	7.09***	0.24***	8.24***	6.91***	0.32***
	(0.48)	(0.36)	(0.02)	(0.07)	(0.09)	(0.02)
Observations	5254	8343	4275	4192	3519	2331

Note: Robust standard errors clustered at the district level are in parentheses. All estimations include individual controls, cohort fixed effects, district fixed effects, district-by-year fixed effects, linear time trends interacted with baseline district level characteristics. Sample consists of partnered women (and their husbands) who married before the expansion of the mills. Since EICV is not geocoded, the mill variable is mill per district area in km^2 in a given year. In Columns 3-5, the sample consists of women and their husbands who reported their occupation as agricultural. *** p<.01, ** p<.05, * p<.1

It should be noted that the results are based on data where I observe the labor income of women who work for pay (positive, not missing labor income). The missing labor income observations belong to women who do not work or do not work for pay. Their labor income is zero either way. As a first robustness check for my results on labor income, I code the missing values as zero and transform the dependent variables by using an inverse hyberbolic sine (IHS) function. Unlike in a log transformation, zeroes are defined in an IHS transformation. Table A.11 presents the results. Estimates are in line with the main labor income results.

Since the data on labor income are available only for a self-selected group of women (women who work for pay), sample selection bias is a potential concern. In order to address this concern, as a second robustness check, I use the two-step Heckman estimator. An exclusion restriction, a variable that affects women's decision to work for pay (her reservation labor income), but not the labor income (labor income offer), identifies the two-step model. I use the number of young children (age 0-5) in the household as such variable in my model. Having young children at home makes it more costly for women to work due to taking care of children. However, it does not affect the amount of income received for per unit labor. The variable is commonly used in the literature that uses the correction in models on female wages and labor force participation (Mulligan and Rubinstein 2008, Blau and Kahn 2017). Table A.12 presents the results. The estimates are still positive and statistically significant.

6.3 Monthly Hospitalizations for Domestic Violence

One of the contributions of this paper is to use monthly administrative hospital records for domestic violence in addition to using annual, self-reported domestic violence data. I estimate the dynamic impact of mill exposure on monthly hospitalizations for gender based violence using Equation 10. Figure 4 plots the coefficient of the interaction terms for every month in a calendar year. Mills operate only during the harvest season, March-July. In a given year, January and February are the pre-harvest period, May-July is the peak of the harvest where majority of the neighboiring community around the mills work in the mills. August-December is the post-harvest period. In the specification, the omitted category is February, which means that the dynamic impact of being exposed to a mill is estimated with respect to one month prior to a mill's month of operation. The top subfigure is for women aged older than 18, where the bottom subfigure is for women younger than 19. Since overwhelming majority of women aged older than 18 is married in Rwanda, I define hospitalizations for gender based violence for women age older than 18 as domestic violence. I find a statistically significant decline in the probability of a hospital having a hospitalization for domestic violence for the months during the peak of the harvest. It is 18 and 14 percentage points less likely for a hospital in the catchment area of a mill to have a domestic violence patient in June and July respectively compared to one month before the mills' month of operation (p-value= 0.01 and p-value= 0.06). The estimated impact represents a reduction of 23% and 17% respectively with respect to the sample mean (0.79). There is not a statistically significant change in hospitalizations when mills do no operate (January, August-December).

I do the same analysis for men and do not find any statistically significant results. I also do not find any statistically significant change in deaths due to gender based violence for both age groups and genders. I plot the results in Figures 5 and A.4 respectively. All results are also shown in Tables A.14.

No change in hospitalizations during the post-harvest months is not surprising. First, the share of women who do not a have savings account is high. 65% of the women in my sample do not have a savings account and this number increases to 73% in rural areas. Second, even if more women have access to savings accounts, the amount of money they earn from the mills is small. It is plausible that their earnings are not enough for them to save for the remaining 7 months of the year which they are not working for pay. In 2010, women's daily earnings is only 17% of their household's daily food consumption expenditures. Moreover, the schools reopen right after the harvest season, August. Although education is free in Rwanda upto nine years, parents have school expenses, primarily a school uniform fee. Anecdotal evidence suggests that women buy school uniforms with the money they earn from the mills, so that their children can go to elementary school. In Table A.13, I report the results of estimating the impact of mill exposure on household's education expenditures for elementary school-age children. 1% increase in mills per capita increases the education expenditures by 4%. If women are more likely to invest in their children's education compared to their husbands, they have less resources left to save.[19]

Placebo Tests. One potential concern with the main results is that women may go to hospitals less during the harvest season due to the increased opportunity cost of a hospital visit. For women who are exposed to a mill, a day in the hospital is costly during the harvest season since it means loss of daily labor income. This can make women less likely to go to a hospital and drive the decline in hospitalizations for domestic violence during the harvest months. To rule out this channel, as a placebo test, I perform the same analysis for universe of monthly hospitalizations for other diseases, bone and joint disorders other than fractures. Examples of such disorders include osteoarthritis, gout, rheumatoid arthritis, lupus and bursitis. I report the results in Figure 6 and Table A.15. I find no changes in hospitalizations for these disorders within a year, specifically during the harvest months relative to one month before the harvest season. This suggests that the decline in hospitalizations due to domestic violence is not driven by a decreased likelihood of going to a hospital during the harvest season.

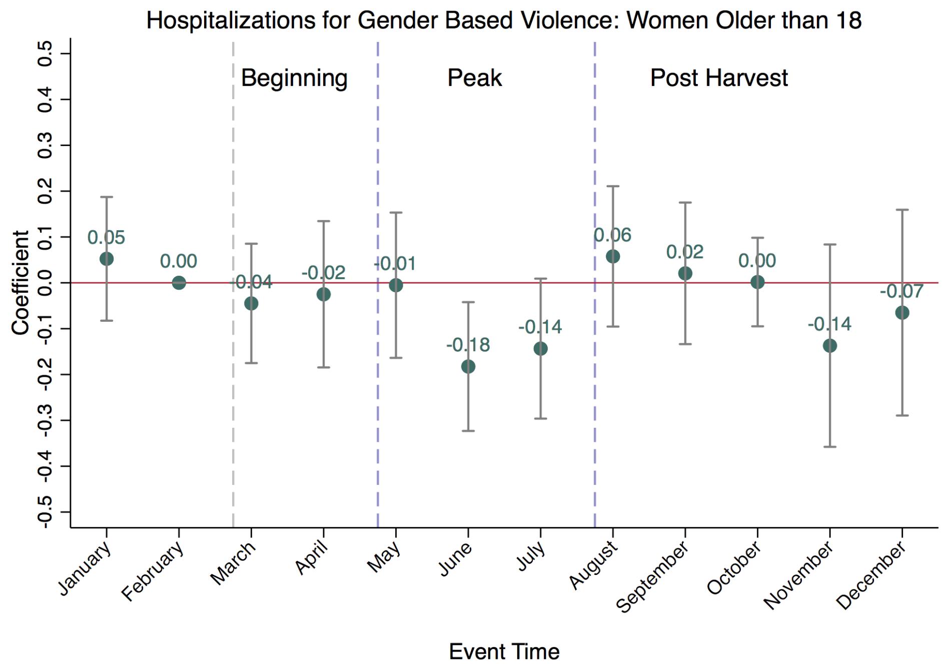
My results for hospitalizations for gender based violence among girls aged 10-18 constitute as another placebo test. Girls aged 10-18 are not the age group who work in the mills and are also most likely to be not partnered. Thus, they are not treated by the mills. As expected, I find no statistically significant changes when the dependent variable is hospitalizations for gender based violence among girls aged 10-18. The placebo tests strengthen the argument that the reduction in hospitalizations for domestic violence is driven by exposure to a mill when mills operate.

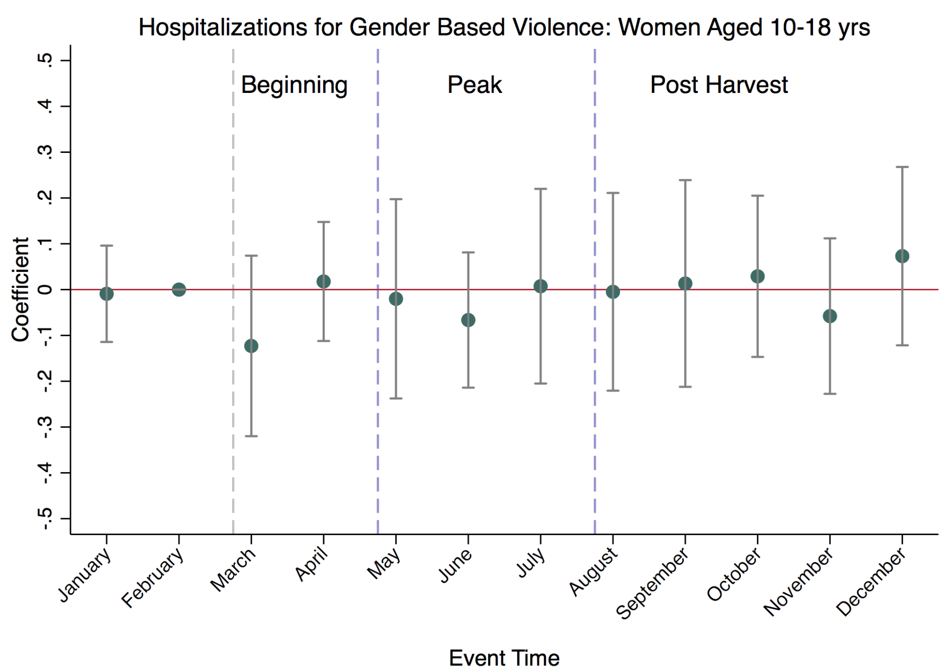
6.4 Mechanisms behind the Decline in Domestic Violence

Staggered DID and DID event-study estimates based on monthly hospitalizations both show a decline in domestic violence. In this subsection, I investigate the potential mechanisms behind the decline in domestic violence via analyzing the impact of a mill opening on bargaining power related outcomes and doing a subsample analysis. As outlined in the model section, there can be multiple reasons behind a decline in domestic violence. These include the increase in women's outside option, exposure reduction and decline in financial stress. While I cannot disentangle the exact mechanism, my analysis suggests that increase in women's outside options due to paid employment is a potential channel behind the decline in domestic violence. I also provide evidence that the decline in domestic violence is observed even among couples where husbands' earnings

[19] Doepke and Tertilt (2019) finds that in Mexico, cash transfers to women (PROGRESA) lead to an increase in spending on children and a decline in the savings rate.

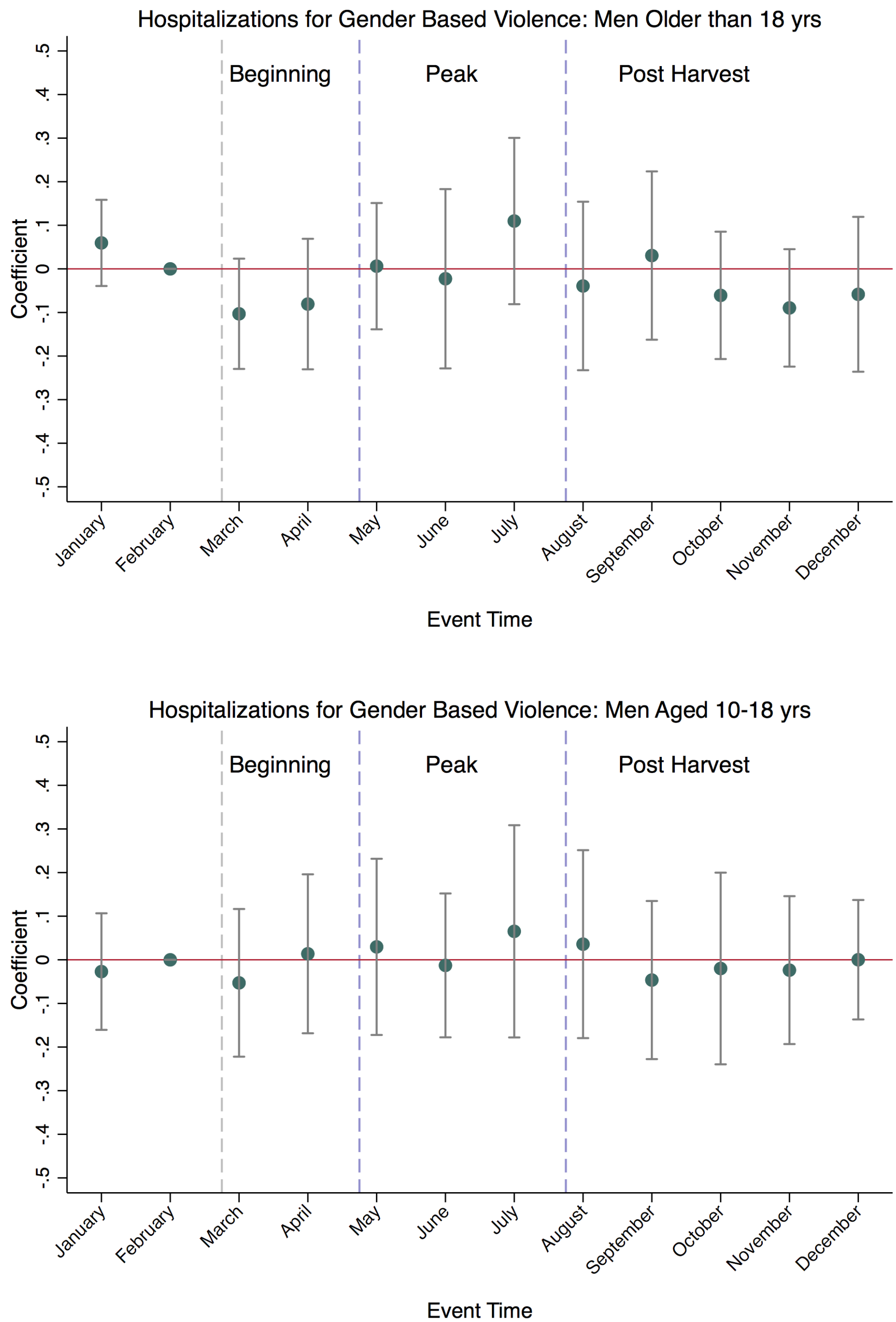
Figure 4: Dynamic Impact of a Mill Opening on Hospitalizations for Gender Based Violence (Women)





Note: Robust standard errors clustered at the sector level are in parentheses. All estimations include hospital controls, hospital fixed effects, district fixed effects, province-by-year fixed effects, linear time trends interacted with baseline district level characteristics. 4 km catchment area is used. *** p<.01, ** p<.05, * p<.1

Figure 5: Dynamic Impact of a Mill Opening on Hospitalizations for Gender Based Violence (Men)



Note: Robust standard errors clustered at the sector level are in parentheses. All estimations include hospital controls, hospital fixed effects, district fixed effects, province-by-year fixed effects, linear time trends interacted with baseline district level characteristics. 4 km catchment area is used. *** p<.01, ** p<.05, * p<.1

Figure 6: Dynamic Impact of a Mill Opening on Log Hospitalizations for Bone and Joint Disorders other than Fractures (Women)

Note: Dependent variable is in logs. Fractures (broken bones) are not included. Examples of bone and joint disorders include osteoarthritis, gout, rheumatoid arthritis, lupus, bursitis. Robust standard errors clustered at the sector level are in parentheses. All estimations include hospital controls, hospital fixed effects, district fixed effects, province-by-year fixed effects, linear time trends interacted with baseline district level characteristics. 4 km catchment area is used. *** p<.01, ** p<.05, * p<.1

do not increase in with mill exposure.

Increase in the Outside Option. According to the family economics literature, an increase in women's outside option increases her bargaining power within the household. Using data on household decision-making, I investigate whether women's employment opportunities translate into an improvement in women's bargaining power. Table 9 presents the results of estimating the impact of mill exposure on household decision-decision making. Three different decisions are investigated. These are making large household purchases, making decisions on own health and visiting family. The estimation captures whether women are more likely to make these decisions by themselves or jointly with their husbands compared to their husbands or someone else in the family is making the decision for them. When a within district approach is used, women who are exposed to a mill are 5 percentage points (p-value= 0.02) more likely to make decisions on large household purchases by themselves or jointly with their husbands. The estimated impact represents an increase of 7% with respect to the sample mean (0.69). There is no change in decision making related to women's own health and visiting family. These results suggests that mill exposure improves women's say on financial decisions in the household.

I also investigate the impact of mill exposure on couple's decision-making on contraception. Table 10 presents the results. The estimation captures whether women's decision to use contraception is taken jointly with their husbands compared to alone or husbands are making the decision for them. For this variable, making the decision alone is not a proxy for bargaining power. Research finds that women in Sub-Saharan Africa may use contraception covertly as a response to their husbands' desire to have more children (Ashraf et al., 2014). When a within district approach is used, among women under the age of 40 who are exposed to a mill are 5 percentage points (p-value= 0.07) more likely to decide to use contraception jointly with their husbands. The estimated impact represents an increase of 6% with respect to the sample mean (0.87). Results on household decisions suggest that women's improvement in their outside options due to mill exposure also improved their bargaining power, say in some household decisions. These results support the argument that an increase in women's outside options due to paid employment is a potential mechanism behind the decline in domestic violence.

Table 9: Effect of Mill Exposure on Household Decision Making: Woman Alone or Jointly with Husband

	Within District			Donut		
	(1) Large HH Purchases	(2) Own Health	(3) Family Visit	(4) Large HH Purchases	(5) Own Health	(6) Family Visit
Mill	0.05**	0.02	0.02	0.03	-0.01	-0.00
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Observations	9823	9823	9823	5165	5165	5165
Dependent variable mean	0.69	0.73	0.81	0.71	0.75	0.82

Note: Robust standard errors clustered at the sector level are in parentheses. All estimations include individual controls, cohort fixed effects, year of marriage fixed effects, sector fixed effects, district-by-year fixed effects, linear time trends interacted with baseline sector level characteristics. Sample consists of partnered women who married before the expansion of the mills. All dependent variables are coded as 1 if the decision is taken by women alone or jointly with their husbands and zero otherwise. Otherwise category includes husband only or someone else. The survey question on household decisions changed significantly over the data cycles. in 2005, the question asks "Who has the final say on decision x?" After 2005, the question is "Who usually decides on x". The answer option "Respondent and other" is also discontinued in 2010 and 2014 rounds. To make a comparable variable, I create the variable based on whether the woman had a say in the decision, either alone or jointly. Results should be interpreted with caution. 4 km catchment area is used. *** p<.01, ** p<.05, * p<.1

Exposure Reduction. Another possible mechanism behind the decline in domestic violence can

Table 10: Effect of Mill Exposure on (Modern) Contraception Decision-Making: Woman Jointly with Husband

	Within District	Donut
	(1) Contraception Decision Maker: Joint	(2) Contraception Decision Maker: Joint
Mill	0.05*	0.05
	(0.03)	(0.03)
Observations	2820	1620
Dependent variable mean	0.87	0.88

Note: Robust standard errors clustered at the sector level are in parentheses. All estimations include individual controls, cohort fixed effects, year of marriage fixed effects, sector fixed effects, district-by-year fixed effects, linear time trends interacted with baseline sector level characteristics. Sample consists of partnered women younger than 40 who married before the expansion of the mills. 4 km catchment area is used.. *** p<.01, ** p<.05, * p<.1

be the reduction in the time couples spend together. Working hours in the mill are long. If women were farming in the same plot with their husbands before the expansion, starting to work in the mills for pay is a shock to the time they are exposed to their husbands as well as their earnings. To provide evidence that exposure reduction is unlikely to drive my results, I take advantage of the information on women and their husbands' occupation. I perform a subsample analysis. I run Equation 8 using the sample of women who do not share the same occupation with their husbands. Self-employed women working in agriculture (smallholder farmers) whose husbands have skilled or unskilled manual jobs (truck drivers, plumbers, refuse workers, labourers in mining, construction and manufacturing etc.) constitute the majority of the subsample. Given that those couples have different occupations, they do not work together during working hours. Thus, women starting working in the mills will not be a shock to the time couples spend together. Their exposure to each other will continue to be the same (no exposure during work hours). However, the workplace of women will be shifted to mills where they work for pay. If there is a statistically significant increase in the probability of women working for cash and a decline in the probability of experiencing domestic violence in the past 12 months for this subsample, it suggests that my main result, reduction in domestic violence, is plausibly not driven by exposure reduction.

Results for women's employment and domestic violence are reported in Table 11. Only within the district approach is used due to sample size. Being exposed to a mill increases the probability of working for cash in the past 12 months by 10 percentage points where the probability of working in the past 12 months remain unchanged. The estimated impact represents an increase of 32% with respect to the sample mean (0.31). Moreover, being exposed to a mill still decreases the probability of experiencing domestic violence in the past 12 months by 11 percentage points. The estimated impact represents a decrease of 32% with respect to the sample mean (0.34). The results show that there is a statistically significant decline in the probability of experiencing domestic violence even among women whose exposure to their husbands are presumably not affected by the expansion. One potential concern is whether the subsample is among couples with very educated husbands compared to the main sample. This is because in the current subsample, husband occupations are non-agricultural jobs. However, according to the data, the mean number of husbands' education is 5.7 for this subsample and it is 4.4 for the whole sample (husbands working in agriculture is not dropped). Taken together, the exercise provides suggestive evidence that exposure reduction is not the mechanism behind my main results. I also further restricted the subsample to women who are farmers and married to men with non-agricultural jobs. There is a statistically significant increase in working for cash for this subsample. I still find a negative estimate for the domestic violence result. However, due to low sample size, I do not have enough power to defend my (statistically insignificant) negative estimate.

Table 11: Effect of Mill Exposure on Women's Employment, Type of Earnings and Self-Reported Domestic Violence in the Past 12 Months: Couples with Different Occupations

	Within District		
	(1) Work	(2) Cash	(3) Violence
Mill	-0.00	0.09**	-0.11*
	(0.03)	(0.04)	(0.06)
Observations	3981	2893	1410
Dependent variable mean	0.71	0.31	0.34

Note: Robust standard errors clustered at the sector level are in parentheses. All estimations include individual controls, cohort fixed effects, year of marriage fixed effects, sector fixed effects, district-by-year fixed effects, linear time trends interacted with baseline sector level characteristics. Sample consists of partnered women who married before the expansion of the mills. For the exposure reduction robustness check, the sample is restricted to women who are not working with their husbands. The restriction is made based on the occupation of women and their husbands. All dependent variables are measured for the past 12 months. 4 km catchment area is used. *** $p<.01$, ** $p<.05$, * $p<.1$

Decline in Financial Stress. Since mill exposure increases husbands' earnings, decline in husbands' financial stress can also be the mechanism behind the decline in domestic violence. In order to investigate this mechanism, I build on the results from the previous exercise. Recall that there is a decline in domestic violence even among couples with different occupations. Women working in agriculture and husbands working in non-agricultural jobs constitute the majority of that subsample. To provide evidence that the decline in financial stress due to an increase husbands' earnings is unlikely to be the main mechanism, I estimate the impact of mill exposure on earnings of women who work in agriculture and husbands who work in non-agricultural jobs. If there is no change in husbands' earnings for this subsample, then there is a decline in violence even among couples where husbands' earnings do not increase with mill exposure. This will suggest that the decline in husbands' financial stress due to an increase in their earnings is not the dominating mechanism behind the decline in domestic violence.

Results are reported in Table 12. I find that mill exposure significantly increases women's daily earnings, however, there is no statistically significant change in their husbands' earnings. The magnitude is also close to 0. This is plausible. Although mills may demand paid labor from men with non-agricultural occupations (construction workers to build a mill, technicians to provide repairs in a mill), the expansion is primarily a (persistent) shock to individuals who are working in agriculture. No effect on husbands' earnings shows that there is decline in violence even among couples where husbands' earnings do not change with mill exposure. This confirms that the decline in husbands' financial stress due to an increase in their earnings is unlikely to be the main mechanism behind my results.

Mechanisms and Women's Paid Employment. Recall from the model that the husband's financial stress can also be reduced due to an increase in women's earnings. Unfortunately I cannot rule out the decline in financial stress due to an increase in women's earnings as a mechanism. However, ruling out the decline in financial stress due to an increase in husbands' earnings as a mechanism suggests that the effects on domestic violence is driven by women's paid employment. This is because the two channels that decreases domestic violence, an increase in women's outside options (a mechanism I provide evidence for), and decline in financial stress due to an increase in women's earnings (a mechanism I cannot rule out) are both due to women's paid employment. Thus, eliminating the decline in financial stress due to an increase in husbands' earnings answers my research question. Providing paid employment opportunities to women plausibly decreases the violence they face from their partners. My empirical analysis further shows that women's paid employment plausibly affects domestic violence via an increase in women's outside options, not exposure reduction.

Table 12: Effect of Mill Exposure on Log Daily Labor Income: Couples with Different Occupations

	Couples with Different Occupations	
	(1)	(2)
	W: Agriculture	H: Non-Agriculture
Mill	7.22***	0.64
	(0.76)	(1.36)
Observations	1828	1134

Note: Robust standard errors clustered at the district level are in parentheses. All estimations include individual controls, cohort fixed effects, district fixed effects, district-by-year fixed effects, linear time trends interacted with baseline district level characteristics. Sample consists of couples who married before the expansion of the mills and have different occupations. Women report their occupation as agricultural where the husbands report their occupation as non-agricultural. Non-agricultural occupations mostly consist of manual (unskilled and skilled) jobs. Since EICV is not geocoded, the mill variable is mill per district area in km^2 in a given year. *** p<.01, ** p<.05, * p<.1

7 Robustness Checks

7.1 Dynamic Impact of Mill Exposure and Pre-Trends

I also estimate the impact of a mill opening using an even-study specification. The specification exploits the variation in years when individuals are exposed to a mill. The results provide insights on the dynamics of the effects, the impact of a mill in the short and long-run. Moreover, estimates during the years before mill exposure constitute a test for the parallel trends assumption. I estimate the dynamic impact of a mill opening with the event study specification below:

$$Y_{istk} = \beta_0 + \sum_{k=-3}^{6} Mill_{is} \times \beta_k \mathbb{1}[\tau = k] + \mathbf{X_{ist}}\phi + \lambda_c + \omega_m + \alpha_s + \gamma_{dt} + (\mathbf{X_s} \times t)\theta + \varepsilon_{istk}. \tag{11}$$

The dependent variable Y_{istk} is the outcome of interest of woman i (or the husband of woman i), in sector s, in year t and at event-time k. $Mills_{is}$ is a binary variable coded as 1 if woman/husband i resides within the catchment area of a mill and zero otherwise by the end of the sample, 2014. It is interacted with event-year dummies, $\mathbb{1}[\tau = k]$, to investigate the dynamic impact of a mill opening. τ denotes the event-year. $\tau = 0$ represents mill's year of operation. For $k > 0$, $\tau = k$ represents k years after a mill's opening. For $k < 0$, $\tau = k$ represents k years before a mill's year of operation. The omitted category is $\tau = -1$, which means that the dynamic impact of being exposed to a mill is estimated with respect to one year prior to a mill opening. $\mathbf{X_{ist}}$ is the same set of controls in the static empirical specification. λ_c is the cohort fixed effects. ω_m is the year of marriage fixed effects. α_s is the sector fixed effects. γ_{dt} is the district-by-year fixed effects. $\mathbf{X_s}$ is a vector of baseline geographical variables interacted with linear time trends. ε_{istk} is the error term. I clustered standard errors at the district level. The main dependent variables for this specification are being employed, working for cash and experiencing domestic violence in the past 12 months.

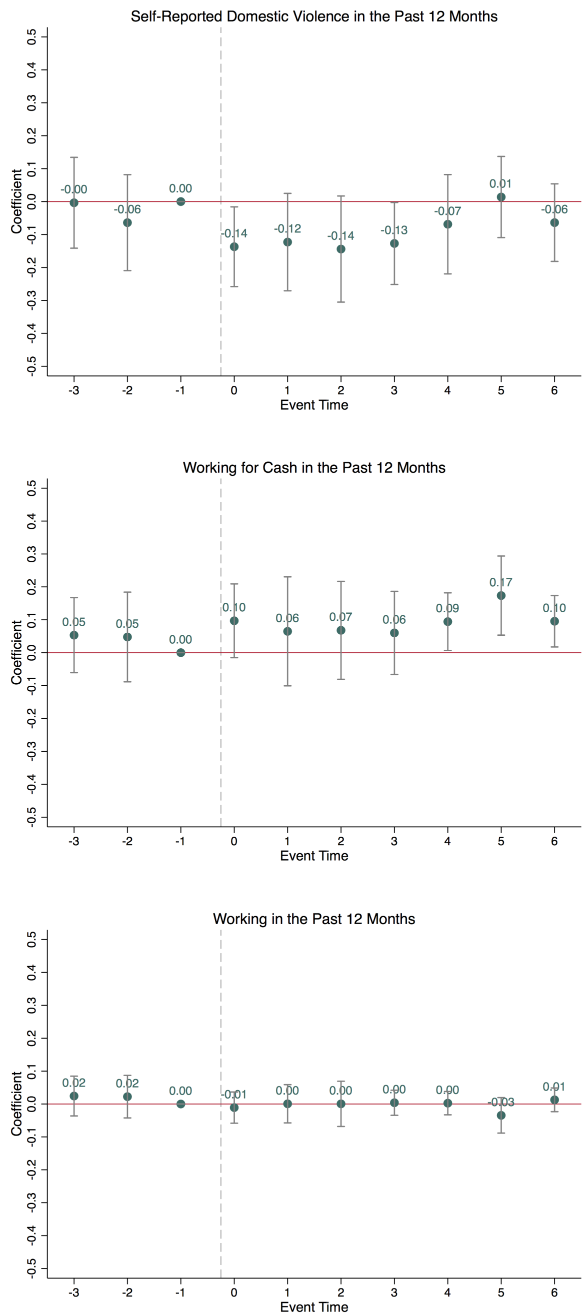
The coefficients of interest are β_ks when $k \geq 0$. For $k \geq 0$, each β_k provides the change in outcomes for the individuals who are exposed to a mill relative to the individuals who are not within the same district, k years after a mill opening, relative to one year before a mill's year of operation. Since $k < 0$ represents the years before a mill opening, I expect to see small and statistically insignificant estimates for such β_ks when $k < 0$. If so, such estimates will constitute supportive evidence in favor of the parallel trends assumption.

Figure 7 presents the results from estimating Equation 11 (event-study specification). It plots the coefficient of the interaction term for 3 years before and 6 years after a mill opening. I use the within district approach to have balanced number of observations every time period. The dynamic results are mostly in line with the static estimates. I find a statistically significant decrease in the probability of experiencing domestic violence in the past 12 months right after a mill opening. Exposure to a mill opening also increases women's probability of working for cash starting from right after a mill opening. There is no change in women's probability of working. I need to note that number of women who are in the treatment group is lower for the event times 1, 2 and 3 compared to the remaining event times. This is because the number of mill openings are not balanced across the 2000s and I have the DHS data cycles for specific years. This can plausibly be the reason why the coefficients at event times 1, 2 and 3 are statistically insignificant for the working for cash variable although they are positive.

I should also note that there is no statistically significant decline in DV at event times 4, 5 and 6, although women are more likely to work for cash in these years relative to one year before a mill opening. The number of observations is not low compared to other years for these event times. This suggests that the effect of women's paid employment in low-paid jobs on domestic violence may not persist for long years.

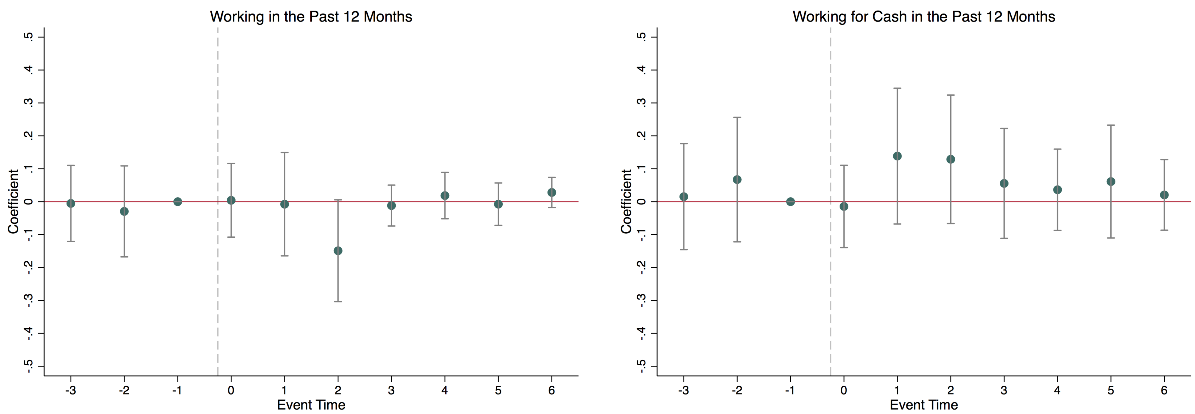
Figure 8 plots the coefficient of the interaction terms for the same specification for husbands. As expected, there is no change in probability of working and working for cash in the past 12 months right after a mill opening. I need to note that there is a statistically significant decline in the husband's probability of working only at event time 2. For working for cash variable, the estimates for event times 1 and 2 are also much bigger than 0 compared to the remaining event times, although insignificant. This is plausibly due to the lower number of observations at event times 1, 2 and 3 based on my data and the imbalance in the mill expansion.

Figure 7: Dynamic Impact of a Mill Opening on Self-Reported Domestic Violence, Women's Type of Earnings and Employment in the Past 12 Months



Note: Robust standard errors clustered at the sector level are in parentheses. All estimations include individual controls, cohort fixed effects, year of marriage fixed effects, sector fixed effects, district-by-year fixed effects, linear time trends interacted with baseline sector level characteristics. Sample consists of partnered women who married before the expansion of the mills. All variables are measured for the past 12 months. 4 km catchment area is used. Within district approach is used. *** p<.01, ** p<.05, * p<.1

Figure 8: Dynamic Impact of a Mill Opening on Husbands' Employment and Type of Earnings in the Past 12 Months



Note: Robust standard errors clustered at the sector level are in parentheses. All estimations include individual controls, cohort fixed effects, year of marriage fixed effects, sector fixed effects, district-by-year fixed effects, linear time trends interacted with baseline sector level characteristics. Sample consists of partners of women in the first specification. All variables are measured for the past 12 months. 4 km catchment area is used. *** p<.01, ** p<.05, * p<.1

For all variables, the coefficients are close to zero and statistically insignificant for the years before a mill opening. Thus, the estimates constitute evidence in favor of the parallel trends assumption. All results are also shown in Tables A.17, A.18, A.19 and A.20.

7.2 Role of Variation in Treatment Timing

Goodman-Bacon (2021) demonstrates that twoway fixed effects difference-in-differences (TWFEDD) estimator is a weighted average of all possible 2x2 DD estimators that compare timing groups to each other. Under time-varying treatment effects, TWFEDD estimates a weighted average treatment effect with *negative weights* (Borusyak and Jaravel 2018, de Chaisemartin and D'Haultfœuille 2020, Sun and Abraham 2020, Callaway and Sant'Anna 2020, Goodman-Bacon 2021) which biases the results. This occurs when the later-treated group uses the earlier-treated group as control. Goodman-Bacon (2021) highlights that specifically the units treated near the beginning or the end of the panel can get more weight as controls than treatments, which is always the case in designs without untreated units. I do not expect this to be a major concern in my context. I have untreated units and out of the 211 mill openings as of 2002, only 21 percent of them opened near the beginning or the end of the panel (2012). 79% of the mills opened between 2005-2010. In order to analyze the 2x2 DD comparisons and weights formally, I perform the Goodman-Bacon (2021) decomposition.[20] The decomposition calculates the weights of each type of 2x2 DD comparison and how it contributes to the TWFEDD estimate. I find that 80-81% of the TWFEDD estimate for labor market outcomes (working for pay, working cash for the past 12 months. and domestic violence is derived from Treatment vs. Never Treated comparisons. Earlier Group Treated vs. Later Group Control and Later Group Treatment Control each has 0.01 weights and do not contribute much to the TWFEDD estimate as expected. The weight for Treated vs. Already Treated comparison is only 0.16.

I also follow the test proposed by de Chaisemartin and D'Haultfœuille (2020) which enables researchers to understand when treatment effect heterogeneity would be a serious concern for the validity of their estimates.[21] de Chaisemartin and D'Haultfœuille (2020) recommends researchers to calculate both the weights and the ratio of the TWFE estimate divided by the standard deviation of the weights. If many weights are negative, and if the ratio is not very large, then treatment effect heterogeneity would be a serious concern for the validity of the estimate.[22] For the working for pay and cash in the past 12 months outcomes, I find that 92 and 93% of my treatment effects receive positive weights respectively where the remaining receive negative weights. I also calculated the ratios based on my estimates and the standard errors of the weights and they are sufficently high. Thus, given that approximately only 7% of the weights are negative and the ratios are sufficiently high, treatment effect heterogeneity is not a concern for the validity of my estimates. For a complete robustness check, I estimate my results using the estimator proposed in de Chaisemartin and D'Haultfœuille (2020) that gives valid results even if the treatment effect is heterogenous over time and across groups. I plot the estimation results for main results in Figure 9.[23] As expected, the estimates are very similar to my main results.[24]

Sun and Abraham (2020) demonstrates that in event study designs with variation in treatment timing across units, the coefficient on a given lag or lead can be contaminated from other periods. Pretrends can arise from heterogeneity of the treatment effects as well. As a robustness check for my event study results, I perform the alternative estimation method proposed in Sun and Abraham (2020) that is free from such contamination.[25] The event study results from this method is reported in Figure 10 and very similar to my main event study results.

[20]I implement the decomposition following Goodman-Bacon (2021) and based on bacondecomp STATA package.

[21]My calculations are based on twowayfeweights STATA package.

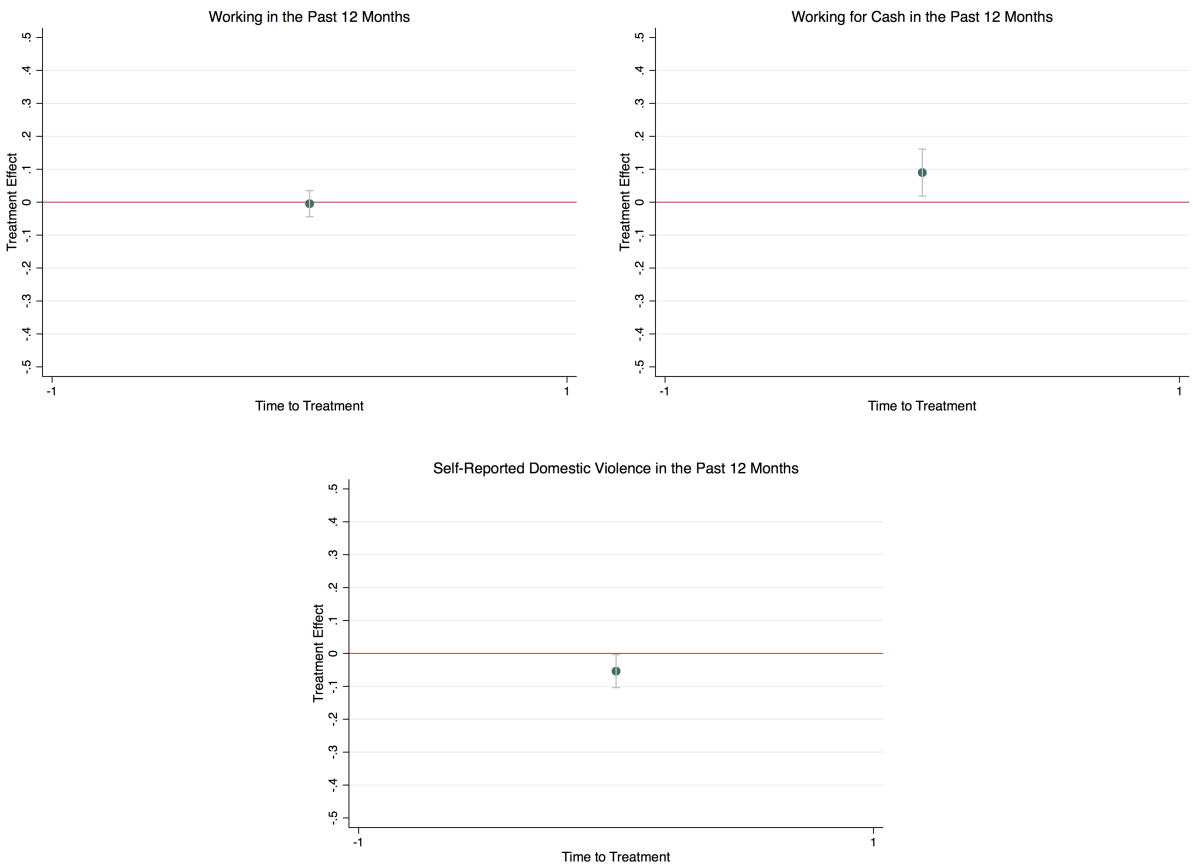
[22]If that ratio is close to 0, that estimate and the average treatment effect can be of opposite signs even under a small and plausible amount of treatment effect heterogeneity. This is the case where, treatment effect heterogeneity would be a serious concern for the validity of that estimate.

[23]The results are based on didmultiplegt STATA package.

[24]The weights and the results in Figure 10 are based on within district approach. The results are also robust and similar when donut approach is used.

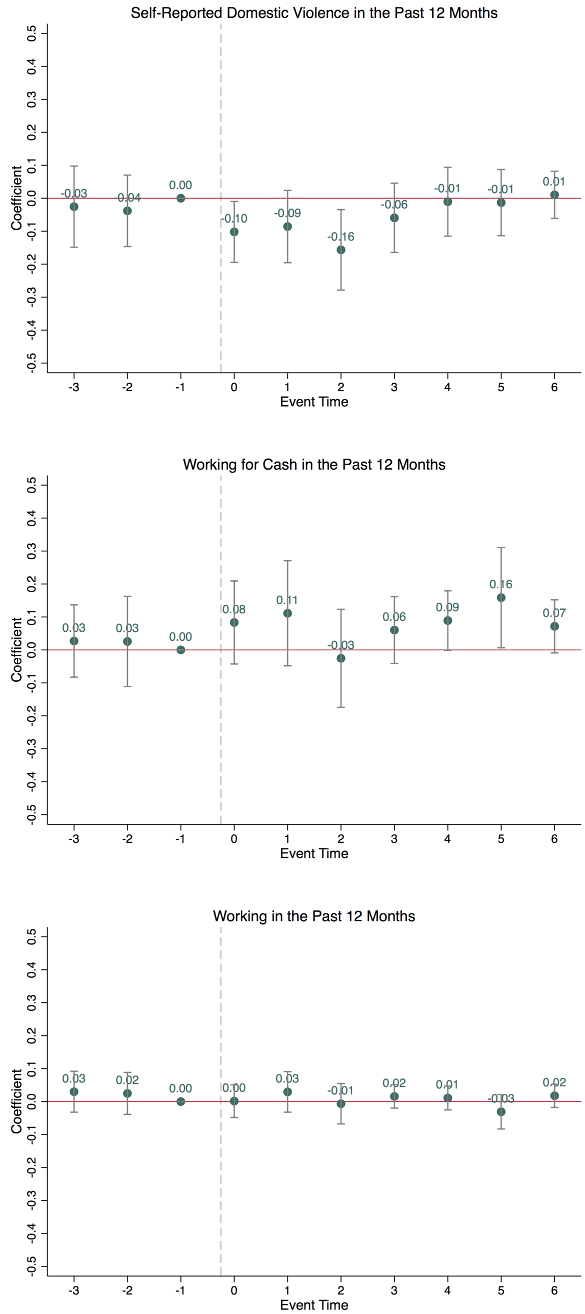
[25]The results are based on eventstudyinteract STATA package.

Figure 9: Impact of a Mill Opening on Women's Employment, Type of Earnings and Self-Reported Domestic Violence in the Past 12 Months using de Chaisemartin and D'Haultfœuille (2020)



Note: Sample consists of partnered women who married before the expansion of the mills. All variables are measured for the past 12 months. 4 km catchment area is used. Within district approach is used.

Figure 10: Dynamic Impact of a Mill Opening on Women's Employment, Type of Earnings and Self-Reported Domestic Violence in the Past 12 Months using Sun and Abraham (2020)



Note: Sample consists of partnered women who married before the expansion of the mills. All variables are measured for the past 12 months. 4 km catchment area is used. Within district approach is used.

7.3 Creating a Mill Suitability Index using Machine Learning

My main identification assumption is that conditional on mill suitability conditions, exposure to mill is random. In the main specification, in order to control for mill suitability conditions, I use log of coffee trees in 1999 and FAO coffee suitability index. To better control for mill suitability, I predict a mill suitability index using a machine learning algorithm and use it as a control variable in my main empirical specification (interacted the time trends with the newly constructed mill suitability index rather than the log of coffee trees in 1999).

When PEARL project is first launched, American agronomists developed a model to identify the mill suitable areas for future mill placement in Rwanda (Schilling and McConnell, 2004).[26] The conditions are sufficient number of coffee trees, proximity to a spring water source, being on the road network and being outside of a national park boundary. In light of these conditions, I predict a mill suitability index. To create the index, as in Macchiavello and Morjaria (2020), I divide the country into 1 km x 1 km grid cells. For each grid, I create variables that provide information on grid's distance to a river, main road, whether the grid is in a national park and the number of coffee trees in the sector the grid lies in. The are in total 28,901 grids in the country and 213 of them have the mills. Such variation is the basis for the mill suitability index. Figure A.3 visualizes the exercise.

Schilling and McConnell (2004) made specific suggestions on the number of coffee trees ($\geq$30,000 trees), distance to river ($\leq$3 km) and distance to road ($\leq$1 km) for optimal mill placement. Macchiavello and Morjaria (2020) predicted a mill suitability index based on these specific conditions using a probit model. I do not restrict myself to the exact conditions Schilling and McConnell (2004) proposed. I create dummy variables for different numbers of coffee trees in 1999, each possible distance to river and main road and whether the grid is in a national park. I predict a suitability index at the grid level based on these dummy variables and its interactions. As the set of these variables is potentially large, I use a machine learning technique, Least Absolute Shrinkage and Selection Operator (LASSO), to predict the mill suitability index. After predicting suitability at the grid level, I aggregate the index at the sector level. The index summaries the mill suitability of a sector before the rapid expansion of the mills in the country. I interact the index with linear time trends in my main empirical specification (rather than interacting the time trends with log coffee trees in 1999). Table 13 reports the results. Although the estimates are a bit smaller and less precise, they are similar to my main results.

[26] The model was unable to be implemented due to lack of GIS data at the time and thus subsequent mill placement is not based on these conditions (Macchiavello and Morjaria, 2020).

Table 13: Effect of Mill Exposure on Main Outcome Variables controlling for LASSO-predicted Mill Suitability Index

	Within District			Donut		
	(1) Work	(2) Cash	(3) Violence	(4) Work	(5) Cash	(6) Violence
Mill	-0.00	0.05**	-0.06*	-0.00	0.06***	-0.07*
	(0.01)	(0.02)	(0.03)	(0.01)	(0.02)	(0.04)
Observations	11237	10064	4080	5418	4902	1740
Dependent variable mean	0.88	0.39	0.35	0.88	0.44	0.37

Note: Robust standard errors clustered at the sector level are in paranthesis. All estimations include individual controls, cohort fixed effects, year of marriage fixed effects, sector fixed effects, district-by-year fixed effects, linear time trends interacted with LASSO predicted, standardized mill suitability index. Sample consists of partnered women who married before the expansion of the mills. All dependent variables are measured for the past 12 months. 4 km catchment area is used. *** p<.01, ** p<.05, * p<.1

7.4 DHS Clusters within the Catchment Area

I show that women within the 4 km catchment area are statistically the same. I redefine the treatment and control groups. I construct the treatment group as women within a 2 km buffer around a mill. Women in the control group are within the donut area between 2 and 4 km from a mill. Thus, all of the observations are within the 4 km buffer around a mill. If the 4 km buffer area correctly captures mill exposure, then, there should not be statistically significant differences in the main outcome variables across women within the 4 km buffer. Table 14 reports the results. There are no statistical differences in the main outcome variables between women within a 2 km buffer around a mill and the women within the donut area between 2 and 4 km from a mill. Results also hold for the subsample of women who is working in agriculture (women who reported agriculture as their occupation).

Throughout the paper, I use a 4 km radius buffer around a mill to construct the treatment group. I also experiment with multiple radii to construct the treatment group using within the district approach for the control group. I show that as the radius of the catchment area increases, meaning more DHS clusters who are actually not treated become a part of the treatment group, my results fade out. Tables A.21, A.22 and Figures A.5, A.6 presents and plots the results using DHS and HMIS (hospital) data.

Table 14: Effect of Mill Exposure on Main Outcome Variables: Within the 4 km Buffer

	All Sample			Women in Agriculture		
	(1) Work	(2) Cash	(3) Violence	(4) Work	(5) Cash	(6) Violence
Mill	-0.00	0.03	-0.05	0.01	0.03	-0.07
	(0.01)	(0.02)	(0.05)	(0.00)	(0.02)	(0.05)
Observations	2468	2468	1001	2164	2164	891
Dependent variable mean	0.98	0.52	0.47	0.99	0.57	0.47

Note: Robust standard errors clustered at the sector level are in paranthesis. All estimations include individual controls, cohort fixed effects, year of marriage fixed effects, sector fixed effects, district-by-year fixed effects, linear time trends interacted with baseline sector level characteristics. Sample consists of partnered women who married before the expansion of the mills. All dependent variables are measured for the past 12 months. The treatment group is a 2 km buffer around a mill and the control group is the donut area between 2 and 4 km from a mill. Thus, all of the observations are within the 4 km buffer around a mill. *** p<.01, ** p<.05, * p<.1

8 Conclusion

This paper studies whether providing paid work opportunities to women decreases violence against women in the household. I use the rapid expansion of the coffee mills in Rwanda in the 2000s that provided paid employment opportunities to millions of women for the first time in their lives. I exploit the differential timing of and spatial variation in exposure to a mill opening. A woman who is exposed to a mill is more likely to work for cash and less likely to self-report experience domestic violence in the past 12 months, while the probability of working remained unchanged. I also find that mill exposure increases women's earnings as well as their husbands'. As a unique feature of the paper, I complement my domestic violence results with the universe of monthly hospital records for domestic violence in Rwanda. Using the monthly hospital records, I take advantage of the timing of the harvest season (March-July), the only period mills operate. I show that it is less likely for a hospital in the catchment area of a mill to have a domestic violence patient when mills are operating compared to one month before the mills' month of operation. I also uncover the mechanisms behind the results. Using data on household decision-making and occupations, I first provide evidence that mill exposure increases women's bargaining power in the household. Then I show that a decline in domestic violence is observed even among couples who do not work together, ruling out exposure reduction as a mechanism. In that subsample, the increase in household resources is via the increase in woman's income only (not the husband's), which suggests that increase in husband's earnings is not the mechanism behind the results. My analysis makes me confident that the decline in domestic violence is driven by women's paid employment. Women's paid employment plausibly decrease domestic violence due to an increase in women's bargaining power and a decrease in financial stress in the household.

I also support my empirical results with a simple theoretical framework. My results are consistent with a household bargaining model where an exogenous increase in women's earnings improves her outside option and decreases the domestic violence she experiences. My results provide insights to policymakers that aim to combat domestic violence. Based on my results, providing paid employment opportunities to women has the potential to alleviate domestic violence in developing countries.

References

AgriLogic (2018). Value Chain Analysis for the Coffee Sector in Rwanda. Technical Report July.

Aizer, A. (2010). The gender wage gap and domestic violence. *American Economic Review 100*(4), 1847–1859.

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

Aizer, A. and P. Dal Bó (2009, apr). Love, hate and murder: Commitment devices in violent relationships. *Journal of Public Economics 93*(3-4), 412–428.

Alesina, A., B. Brioschi, and E. La Ferrara (2020). Violence Against Women: A Cross-cultural Analysis for Africa. *Economica*.

Anderberg, D., H. Rainer, J. Wadsworth, and T. Wilson (2016). Unemployment and Domestic Violence: Theory and Evidence. *Economic Journal 126*(597), 1947–1979.

Anderson, S. and G. Genicot (2015). Suicide and property rights in India. *Journal of Development Economics 114*, 64–78.

Angelucci, M. (2008). Love on the rocks: Domestic violence and alcohol abuse in rural Mexico. *B.E. Journal of Economic Analysis and Policy 8*(1).

Arenas-Arroyo, E., D. Fernandez-Kranz, and N. Nollenberger (2021, feb). Intimate partner violence under forced cohabitation and economic stress: Evidence from the COVID-19 pandemic. *Journal of Public Economics 194*.

Ashraf, N., E. Field, and J. Lee (2014). Household bargaining and excess fertility: An experimental study in zambia. *American Economic Review 104*(7), 2210–2237.

Bayisenge, J. (2010). Access to Paid Work and Women's Empowerment in Rwanda. *19*(1), 84–106.

Bhalotra, S., D. G. Britto, P. Pinotti, and B. Sampaio (2021). Job Displacement, Unemployment Benefits and Domestic Violence. *SSRN*.

Bhalotra, S., U. Kambhampati, S. Rawlings, and Z. Siddique (2019). Intimate Partner Violence: The Influence of Job Opportunities for Men and Women. *The World Bank Economic Review*.

Blau, F. D. and L. M. Kahn (2017, sep). The gender wage gap: Extent, trends, & explanations. *Journal of Economic Literature 55*(3), 789–865.

Bloch, F. and V. Rao (2002). Terror as a bargaining instrument: A case study of dowry violence in rural India. *American Economic Review 92*(4), 1029–1043.

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

Borusyak, K. and X. Jaravel (2018, mar). Revisiting Event Study Designs. *SSRN Electronic Journal*.

Boudreaux, K. C. (2011). Economic liberalization in Rwanda's coffee sector: A better brew for success. In *Yes Africa Can*, pp. 185.

Brassiolo, P. (2016). Domestic violence and divorce law: When divorce threats become credible. *Journal of Labor Economics 34*(2), 443–477.

Bursztyn, L., A. L. González, and D. Yanagizawa-Drott (2020, oct). Misperceived social norms: women working outside the home in saudi arabia. *American Economic Review 110*(10), 2997–3029.

Callaway, B. and P. H. Sant'Anna (2020). Difference-in-Differences with multiple time periods. *Journal of Econometrics*.

Calvi, R. and A. Keskar (2021). 'Til Dowry Do Us Part: Bargaining and Violence in Indian Families.

Chiappori, P. A. and M. Mazzocco (2017). Static & intertemporal household decisions. *Journal of Economic Literature 55*(3), 985–1045.

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

Chin, Y. M. and S. Cunningham (2019). Revisiting the effect of warrantless domestic violence arrest laws on intimate partner homicides. *Journal of Public Economics*.

de Chaisemartin, C. and X. D'Haultfœuille (2020, sep). Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects. *American Economic Review 110*(9), 2964–2996.

Doepke, M. and M. Tertilt (2019). Does female empowerment promote economic development? *Journal of Economic Growth 24*(4).

Duflo, E. (2012). Women empowerment and economic development. *Journal of Economic Literature 50*(4), 1051–1079.

Dugan, L., D. S. Nagin, and R. Rosenfeld (1999). Explaining the decline in intimate partner homicide: The effects of changing domesticity, women's status, and domestic violence resources. *Homicide Studies 3*(3), 187–214.

Erten, B. and P. Keskin (2018). For better or for worse? Education and the prevalence of domestic violence in Turkey. *American Economic Journal: Applied Economics 10*(1), 64–105.

Erten, B. and P. Keskin (2020). Breaking the Cycle? Education and the Intergenerational Transmission of Violence. *The Review of Economics and Statistics 102*(2), 252–268.

Erten, B. and P. Keskin (2021a). Female employment and intimate partner violence: Evidence from Syrian Refugee inflows to Turkey. *Journal of Development Economics 150*.

Erten, B. and P. Keskin (2021b). No Trade-offs? The Impact of WTO Accession on Intimate Partner Violence in Cambodia.

Eswaran, M. and N. Malhotra (2011). Domestic violence and women's autonomy in developing countries: theory and evidence. *Canadian Journal of Economics/Revue canadienne d'économique 44*(4), 1222–1263.

FAO (2010). Women's work. Technical report.

Farmer, A. and J. Tiefenthaler (1997). An economic analysis of domestic violence. *Review of Social Economy 55*(3), 337–358.

Field, E., R. Pande, N. Rigol, S. Schaner, and C. T. Moore (2021, jul). On her own account: How strengthening women's financial control impacts labor supply and gender norms. *American Economic Review 111*(7), 2342–2375.

García-Ramos, A. (2021, may). Divorce laws and intimate partner violence: Evidence from Mexico. *Journal of Development Economics 150*.

Gelles, R. (1974). *The Violent Home: A Study of Physical Aggression Between Husbands and Wives*. Sage Publications.

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*.

Green, D. P., A. M. Wilke, and J. Cooper (2020). Countering Violence Against Women by Encouraging Disclosure: A Mass Media Experiment in Rural Uganda. *Comparative Political Studies 53*(14).

Guarnieri, E. and H. Rainer (2021, jun). Colonialism and female empowerment: A two-sided legacy. *Journal of Development Economics 151*.

Haushofer, J., C. Ringdal, J. P. Shapiro, and X. Y. Wang (2019). Income Changes and Intimate Partner Violence: Evidence from Unconditional Cash Transfers in Kenya. *NBER Working paper*.

Heath, R. (2014, may). Women's Access to Labor Market Opportunities, Control of Household Resources, and Domestic Violence: Evidence from Bangladesh. *World Development 57*, 32–46.

Heath, R., M. Hidrobo, and S. Roy (2020, mar). Cash transfers, polygamy, and intimate partner violence: Experimental evidence from Mali. *Journal of Development Economics 143*.

Heath, R. and S. Jayachandran (2017, jan). The causes and consequences of increased female education and labor force participation in developing countries. In *The Oxford Handbook of Women and the Economy*, pp. 345–367. Oxford University Press.

Heath, R. and A. M. Mobarak (2015, jul). Manufacturing growth and the lives of Bangladeshi women. *Journal of Development Economics 115*, 1–15.

Hidrobo, M., A. Peterman, and L. Heise (2016). The effect of cash, vouchers, and food transfers on intimate partner violence: Evidence from a randomized experiment in Northern Ecuador. *American Economic Journal: Applied Economics 8*(3), 284–303.

ICRW (2011). Evolving Men: Gender Equality Survey. Technical report.

IMF (2018). The Future of Work in Sub-Saharan Africa. Technical Report 18.

Kruk, M. R., W. Meelis, J. Halász, and J. Haller (2004, oct). Fast positive feedback between the adrenocortical stress response and a brain mechanism involved in aggressive behavior. *Behavioral Neuroscience 118*(5), 1062–1070.

La Mattina, G. (2017). Civil conflict, domestic violence and intra-household bargaining in post-genocide Rwanda. *Journal of Development Economics 124*, 168–198.

Luke, N. and K. Munshi (2011, jan). Women as agents of change: Female income and mobility in India. *Journal of Development Economics 94*(1), 1–17.

Lundberg, S. and R. A. Pollak (1993, dec). Separate Spheres Bargaining and the Marriage Market. *Journal of Political Economy 101*(6), 988–1010.

Macchiavello, R. and A. Morjaria (2020). Competition and Relational Contracts in the Rwanda Coffee Chain. *The Quarterly Journal of Economics*.

Macmillan, R. and R. Gartner (1999, nov). When She Brings Home the Bacon: Labor-Force Participation and the Risk of Spousal Violence against Women. *Journal of Marriage and the Family 61*(4), 947.

Manser, M. and M. Brown (1980). Marriage and Household Decision-Making: A Bargaining Analysis. *International Economic Review 21*(1), 31.

McElroy, M. B. and M. J. Horney (1981). Nash-Bargained Household Decisions: Toward a Generalization of the Theory of Demand. *International Economic Review 22*(2), 333–349.

Miller, A. R. and C. Segal (2019). Do Female Officers Improve Law Enforcement Quality? Effects on Crime Reporting and Domestic Violence. *Review of Economic Studies*.

MINAGRI and MINICOM (2008). Rwanda National Coffee Strategy. Technical report.

Mulligan, C. B. and Y. Rubinstein (2008, aug). Selection, investment, and women's relative wages over time. *Quarterly Journal of Economics 123*(3), 1061–1110.

Perez-Heydrich, C., J. L. Warren, C. R. Burget, and M. E. Emch (2013). Guidelines on the Use of DHS GPS Data. Spatial Analysis Reports No. 8. Technical report.

Qian, N. (2008, aug). Missing women and the price of tea in China: The effect of sex-specific earnings on sex imbalance. *Quarterly Journal of Economics 123*(3), 1251–1285.

Rogall, T. (2021, jan). Mobilizing the masses for genocide. *American Economic Review 111*(1), 41–72.

Sanin, D. (2021). Do Domestic Violence Laws Protect Women From Domestic Violence ? Evidence From Rwanda.

SCAA Sustainability Council (2015). The Blueprint for Gender Equality in the Coffeelands. Technical report.

Schilling, M. and T. McConnell (2004). Model for Siting Coffee Washing Stations in Rwanda.

Scholer, M. (2008). Women In Coffee. *International Trade Forum* (3/4), 32–33.

Stevenson, B. and J. Wolfers (2006). Bargaining in the shadow of the law: Divorce laws and family distress. *Quarterly Journal of Economics 121*(1), 267–288.

Sun, L. and S. Abraham (2020). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*.

Sviatschi, M. M. and I. Trako (2021). Gender Violence, Enforcement, and Human Capital: Evidence from Women's Justice Centers in Peru.

Tauchen, H. V., A. D. Witte, and S. K. Long (1991). Domestic Violence: A Nonrandom Affair. *International Economic Review*, 491–511.

Tur-Prats, A. (2021, may). Unemployment and intimate partner violence: A Cultural approach. *Journal of Economic Behavior and Organization 185*, 27–49.

WHO (2013). Global and regional estimates of violence against women: prevalence and health effects of intimate partner violence and non-partner sexual violence. Technical report.

World Bank (2015). Women, Business and the Law 2016: Getting to Equal. Technical report, World Bank.

Ya-Bititi, G. M., P. Lebailly, D. S. Mbonyinkebe, and P. Burny (2019, aug). 'Coffee has given us power to act': Coffee cooperatives and women's empowerment in Rwanda's rural areas: A case study of Karaba coffee cooperative. In *Cooperatives and the World of Work*, pp. 107–118. Taylor and Francis.

Appendix

A Additional Tables and Figures

Table A.1: Summary Statistics of Mills

	Mean	Std. Dev.
Panel A: Mill Level Characteristics		
Owner: Cooperative	0.5	0.5
Owner: NGO	0.2	0.4
Opened before 2005	0.1	0.3
Opened between 2005-2010 (Rapid Expansion period)	0.8	0.4
Panel B: Sector Level Characteristics		
Number of coffee farmers in 1999	702.2	464.2
Log Coffee Trees in 1999	11.5	0.6
Log Coffee Trees in 2003	11.6	0.6
Log Coffee Trees in 2009	12.9	0.7
Log Coffee Trees in 2015	13.1	0.9
FAO-GAEZ Coffee Suitability Index for Coffee: Marginal	0.3	0.5
FAO-GAEZ Coffee Suitability Index for Coffee: Moderate	0.4	0.5
FAO-GAEZ Coffee Suitability Index for Coffee: Medium	0.2	0.4
FAO-GAEZ Coffee Suitability Index for Coffee: Good	0.0	0.2
Panel A: Spatial Characteristics		
Sector Area in km2	52.4	26.3
District Area in km2	834.2	274.7

Note: Mill level characteristics are from the Rwanda GeoPortal and Macchiavello and Morjaria (2020). Data on coffee trees and farmers are from the coffee censuses. FAO-GAEZ Suitability Index categories are based on FAO-GAEZ's suitability categories. Aggregate crop suitability is (1) Very high when index $>$ 8500; (2) High when 7000 $<$ index $\leq$ 8500; (3) Good when 5500 $<$ index $\leq$ 7000; (4) Medium when 4000 $<$ index $\leq$ 5500; (5) Moderate when 2500 $<$ index $\leq$ 4000; (6) Marginal when 1000 $<$ index$\leq$ 2500; (7) Very marginal when 0 $<$ index $\leq$ 1000; and (8) Not suitable when index $=$ 0. The continuous index is at 9 km^2 resolution. I aggregate the index at the sector level and classify the sector level index based on FAO's suitability categories. Spatial characteristics are the author's calculations using spatial maps by the Rwanda GeoPortal via ArcGIS software.

Table A.2: Summary Statistics for Women: DHS Women's Recode

	All		Exposed to a mill		Not exposed to a mill	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Panel A: Main dependent variables						
Worked in the past 12 months	0.88	0.32	0.88	0.32	0.88	0.32
Worked for cash in the past 12 months	0.39	0.49	0.47	0.50	0.37	0.48
Experienced domestic violence in the past 12 months	0.34	0.47	0.37	0.48	0.33	0.47
Panel B: Controls						
Husband lives in the house	0.88	0.32	0.87	0.34	0.89	0.32
Husband's age	42.42	9.85	43.11	9.72	42.20	9.88
Husband's Occupation: Agricultural	0.72	0.45	0.73	0.44	0.71	0.45
Husband's education in years	4.32	3.76	4.32	3.60	4.32	3.81
Occupation: Agricultural	0.77	0.42	0.79	0.41	0.76	0.43
Marital status: Married	0.73	0.44	0.80	0.40	0.71	0.45
Monogamy (No other wives)	0.89	0.31	0.92	0.27	0.88	0.33
Number of unions: One	0.84	0.37	0.85	0.36	0.84	0.37
Age at first marriage	19.84	3.33	20.16	3.39	19.74	3.30
Years since marriage	16.45	6.44	16.87	6.28	16.31	6.48
Education in years	3.99	3.54	4.26	3.42	3.90	3.58
Muslim	0.02	0.13	0.01	0.11	0.02	0.14
Christian	0.96	0.19	0.97	0.17	0.96	0.19
Has children aged 5 and under	0.76	0.43	0.73	0.44	0.77	0.42
Type of residence: Rural	0.84	0.37	0.86	0.35	0.83	0.38
Household has a radio	0.63	0.48	0.64	0.48	0.62	0.48
Household's main floor material is cement	0.17	0.37	0.16	0.37	0.17	0.38
Household has electricity	0.12	0.33	0.13	0.34	0.12	0.32
Household wealth is above the median	0.51	0.50	0.49	0.50	0.51	0.50
Observations	11652		2853		8799	

Note: Sample consists of partnered women who married before the expansion of the mills. "Exposed to the mill" represents being in the cathment areas of a mill. 4 km catchment area is used.

Table A.3: Summary Statistics for Husbands: DHS Couple's Recode

	All		Exposed to the mill		Not exposed to the mill	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Panel A: Main dependent variables						
Worked in the past 12 months	0.87	0.33	0.89	0.31	0.87	0.34
Worked for cash in the past 12 months	0.80	0.40	0.82	0.38	0.79	0.41
Panel B: Controls						
Husband lives in the house	1.00	0.07	0.99	0.08	1.00	0.07
Wife's Age	36.20	6.85	36.97	6.70	35.96	6.87
Wife's Occupation: Agricultural	0.76	0.43	0.79	0.41	0.75	0.43
Wife's Education in years	3.91	3.44	4.14	3.32	3.84	3.47
Occupation: Agricultural	0.62	0.49	0.66	0.48	0.61	0.49
Marital status: Married	0.77	0.42	0.84	0.37	0.74	0.44
Monogamy (No other wives)	0.95	0.22	0.96	0.20	0.94	0.23
Number of unions: One	0.76	0.43	0.77	0.42	0.76	0.43
Age at first marriage	23.71	4.42	24.12	4.41	23.58	4.42
Years since marriage	17.16	7.25	17.59	7.09	17.03	7.29
Education in years	4.33	3.69	4.39	3.61	4.31	3.71
Muslim	0.02	0.14	0.01	0.12	0.02	0.15
Christian	0.95	0.21	0.96	0.20	0.95	0.22
Has children aged 5 and under	0.78	0.42	0.76	0.43	0.78	0.41
Type of residence: Rural	0.84	0.37	0.87	0.33	0.83	0.37
Household has a radio	0.65	0.48	0.66	0.47	0.65	0.48
Household's main floor material is cement	0.16	0.37	0.16	0.37	0.16	0.37
Household has electricity	0.12	0.32	0.13	0.34	0.11	0.32
Household wealth is above the median	0.53	0.50	0.53	0.50	0.53	0.50
Observations	4816		1159		3657	

Note: Sample consists of the partners of women in Table A.3. "Exposed to the mill" represents being in the cathment area of a mill. 4 km catchment area is used.

Table A.4: Summary Statistics for Hospitals: HMIS

	All		Exposed to a mill		Not exposed to a mill	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Panel A: Main dependent variables						
Has a GBV Patient: Women aged older than 18	0.79	0.41	0.70	0.46	0.85	0.36
Has a GBV Patient: Women aged 10-18 yrs	0.84	0.36	0.77	0.42	0.89	0.31
Has a GBV Patient: Men aged older than 18	0.24	0.43	0.19	0.39	0.27	0.45
Has a GBV Patient: Men aged 10-18 yrs	0.13	0.34	0.11	0.31	0.15	0.36
Death due to GBV: Women aged older than 18	0.02	0.16	0.03	0.17	0.02	0.15
Death due to GBV: Women aged 10-18 yrs	0.01	0.08	0.01	0.08	0.01	0.08
Death due to GBV: Men aged older than 18	0.02	0.13	0.02	0.15	0.02	0.12
Death due to GBV: Men aged 10-18 yrs	0.01	0.07	0.01	0.10	0.00	0.06
Panel B: Breakdown of dependent variables in numbers						
Patients w. physical GBV symptoms: Women aged older than 18	2.52	4.33	1.56	2.19	3.07	5.07
Patients w. sexual GBV symptoms: Women aged older than 18	1.70	2.37	1.16	1.70	2.02	2.63
Patients w. physical GBV symptoms.: Women aged 10-18 yrs	0.24	0.80	0.16	0.49	0.29	0.92
Patients w. sexual GBV symptoms: Women aged 10-18 yrs	5.07	5.24	4.36	4.67	5.58	5.49
Panel C: Controls						
Patient is referred to the hospital by police	0.83	0.38	0.77	0.42	0.86	0.35
Patient is referred to the hospital by community health worker	0.28	0.45	0.20	0.40	0.33	0.47
Patient is referred for care to higher level health facility	0.08	0.28	0.08	0.27	0.09	0.28
Observations	1499		497		966	

Note: GBV stands for gender based violence. Variables in Panel A and C are dummy variables. Panel B reports the monthly hospitaliations for GBV. Variables reported in Panel B are used to construct the "has a GBV patient" dummy variable for a specific age group. "Patient" in Panel C is a gender based violence patient. Sample consists of hospitals. HMIS is a panel data. There are 42 hospitals which are observed for every month for 3 years. Thus, the number of observations is a product of the number of hospitals, 12 and 3. "Exposed to the mill" represents being in the cathment area of a mill. 4 km catchment area is used.

Table A.5: Summary Statistics for Women and Their Husbands: EICV

	All		Exposed to a mill		Not exposed to a mill	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Panel A: Labor income variables						
Women's daily labor income, in Rwandan francs	430.20	2823.20	441.37	2887.79	358.06	2363.41
Husbands' daily labor income, in Rwandan francs	1153.84	5718.61	1183.43	5824.34	962.80	4979.23
Women-to-Husband daily labor income ratio	0.63	13.82	0.65	14.78	0.54	0.92
Panel B: Controls						
Husband is not absent in the past 12 months	0.78	0.42	0.77	0.42	0.79	0.41
Husband's age	37.04	8.94	37.11	8.93	36.64	9.04
Husband's Occupation: Agricultural'	0.16	0.37	0.16	0.36	0.20	0.40
Husband's Education: No education	0.10	0.30	0.10	0.30	0.11	0.32
Husband's Education: Incomplete primary	0.59	0.49	0.59	0.49	0.55	0.50
Husband's Education: Primary completed	0.25	0.43	0.25	0.43	0.25	0.43
Husband's Education: More than primary	0.15	0.35	0.15	0.35	0.15	0.36
Household size	5.34	2.01	5.32	2.00	5.50	2.09
Occupation: Agricultural	0.45	0.50	0.42	0.49	0.59	0.49
Marital status: Married monogamous	0.81	0.39	0.82	0.38	0.74	0.44
Marital status: Married polygamous	0.03	0.17	0.03	0.17	0.04	0.19
Education: No education	0.11	0.31	0.10	0.30	0.16	0.37
Education: Incomplete primary	0.60	0.49	0.60	0.49	0.55	0.50
Education: Primary completed	0.23	0.42	0.24	0.43	0.18	0.38
Education: More than primary	0.12	0.33	0.13	0.33	0.11	0.31
Children in the HH aged 5 and under	0.80	0.40	0.80	0.40	0.83	0.37
Type of residence: Rural	0.83	0.37	0.84	0.37	0.83	0.38
Household's main floor material is cement	0.18	0.38	0.18	0.39	0.16	0.36
Household has electricity	0.14	0.34	0.14	0.35	0.11	0.32
Observations	18600		16105		2495	

Note: Sample consists of partnered women who presumably married before the expansion of the mills. Year of marriage is not collected in EICV. In order to construct the sample of women who partnered before the expansion of the mills (as in DHS), I took the scohorts of women I observe in DHS. EICV is not geocoded. Thus, "Exposed to the mill" represents having a mill in the district.

Table A.6: Effect of Mill Exposure on Women's Occupation: Within District Approach

	(1) Managers	(2) Sales	(3) Agricultural Self-Employed	(4) Agricultural Employee	(5) Manual Skilled & Unskilled
Mill	0.00	-0.01	-0.01	-0.00	0.02
	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)
Observations	8762	8762	8762	8762	8762
Dependent variable mean	0.03	0.07	0.82	0.04	0.03

Note: Robust standard errors clustered at the sector level are in paranthesis. All estimations include individual controls, cohort fixed effects, year of marriage fixed effects, sector fixed effects, district-by-year fixed effects, linear time trends interacted with baseline sector level characteristics. Sample consists of partnered women who married before the expansion of the mills. All dependent variables are measured for the past 12 months. 4 km catchment area is used. Within district approach is used for the control group. *** p<.01, ** p<.05, * p<.1

Table A.7: Effect of Mill Exposure on Husbands' Occupation: Within District Approach

	(1) Managers	(2) Sales	(3) Agricultural Self-Employed	(4) Agricultural Employee	(5) Manual Skilled & Unskilled
Mill	0.00	0.03**	-0.04	0.00	0.01
	(0.01)	(0.02)	(0.03)	(0.01)	(0.03)
Observations	3547	3547	3547	3547	3547
Dependent variable mean	0.05	0.06	0.66	0.04	0.17

Note: Robust standard errors clustered at the sector level are in parantheses. All estimations include individual controls, cohort fixed effects, year of marriage fixed effects, sector fixed effects, district-by-year fixed effects, linear time trends interacted with baseline sector level characteristics. Sample consists of partners of women in the first specification. All dependent variables are measured for the past 12 months. 4 km catchment area is used. Within district approach is used for the control group. *** $p<.01$, ** $p<.05$, * $p<.1$

Table A.8: Effect of Mill Exposure on Women's OccupationL Donut Approach

	(1) Managers	(2) Sales	(3) Agricultural Self-Employed	(4) Agricultural Employee	(5) Manual Skilled & Unskilled
Mill	0.01	-0.01	-0.01	-0.01	0.02
	(0.01)	(0.01)	(0.02)	(0.01)	(0.02)
Observations	4628	4628	4628	4628	4628
Dependent variable mean	0.03	0.08	0.80	0.03	0.04

Note: Robust standard errors clustered at the sector level are in parantheses. All estimations include individual controls, cohort fixed effects, year of marriage fixed effects, sector fixed effects, district-by-year fixed effects, linear time trends interacted with baseline sector level characteristics. Sample consists of partnered women who married before the expansion of the mills. All dependent variables are measured for the past 12 months. 4 km catchment area is used. Donut approach is used for the control group. *** p<.01, ** p<.05, * p<.1

Table A.9: Effect of Mill Exposure on Husbands' Occupation: Donut Approach

	(1) Managers	(2) Sales	(3) Agricultural Self-Employed	(4) Agricultural Employee	(5) Manual Skilled & Unskilled
Mill	0.01	0.04**	-0.05	-0.01	0.01
	(0.01)	(0.02)	(0.04)	(0.02)	(0.03)
Observations	1958	1958	1958	1958	1958
Dependent variable mean	0.05	0.06	0.67	0.04	0.16

Note: Robust standard errors clustered at the sector level are in parantheses. All estimations include individual controls, cohort fixed effects, year of marriage fixed effects, sector fixed effects, district-by-year fixed effects, linear time trends interacted with baseline sector level characteristics. Sample consists of partners of women in the first specification. All dependent variables are measured for the past 12 months. 4 km catchment area is used. Donut approach is used for the control group. *** p<.01, ** p<.05, * p<.1

Table A.10: Effect of Mill Exposure on Women-to-Husband Log Daily Labor Income Ratio

	All Sample	Occupation: Agriculture
	(1)	(2)
	Woman/Husband	Woman/Husband
Mill	0.28***	1.98***
	(0.05)	(0.27)
Observations	5236	1487

Note: Robust standard errors clustered at the district level are in parentheses. All estimations include individual controls, cohort fixed effects, district fixed effects, district-by-year fixed effects, linear time trends interacted with baseline district level characteristics. Since EICV is not geocoded, the mill variable is mill per district area in km^2 in a given year. In Columns 3-4, the sample consists of women and their husbands who reported their occupation as agricultural. *** p<.01, ** p<.05, * p<.1

Table A.11: Effect of Mill Exposure on Women-to-Husband Log Daily Labor Income Ratio, Women's and Their Husbands' Log Daily Labor Income using Inverse Hypberbolic Sin Transformation for the Dependent Variables

	All Sample			Occupation: Agriculture		
	(1)	(2)	(3)	(4)	(5)	(6)
	W/H	W	H	W/H	W	H
Mill	0.211***	7.256***	6.673***	0.286***	8.297***	7.023***
	(0.02)	(0.12)	(0.09)	(0.02)	(0.08)	(0.09)
Observations	4257	5218	8329	2320	4157	3511

Note: Robust standard errors clustered at the district level are in paranthesis. All estimations include individual controls, cohort fixed effects, district fixed effects, district-by-year fixed effects, linear time trends interacted with baseline district level characteristics. Since EICV is not geocoded, the mill variable is mill per district area in km^2 in a given year. In Columns 4-6, the sample consists of women and theit husbands who reported their occupation as agricultural. *** p<.01, ** p<.05, * p<.1

Table A.12: Effect of Mill Exposure on Women-to-Husband Log Daily Labor Income Ratio using Heckman Selection Model

	All Sample
	(1) W/H
Women-to-Husband Log Labor Income Ratio	
Mill	0.032***
	(0.00)
Observations	5218

Note: Robust standard errors clustered at the district level are in parentheses. Estimates show the second stage results. In both specifications, total number of children under the age 5 is used as the identifying variable and used in the first stage, selection equation. It affects women's labor market participation (reservation wage), but not the wages (wage offer). All estimations include individual controls, cohort fixed effects, district fixed effects, district-by-year fixed effects, linear time trends interacted with baseline district level characteristics. In the selection equation, province fixed effects are used to avoid collinearity. Since EICV is not geocoded, the mill variable is mill per district area in km^2 in a given year. In Columns 4-6, the sample consists of women and their husbands who reported their occupation as agricultural. *** p<.01, ** p<.05, * p<.1

Table A.13: Effect of Mill Exposure on Household's Log Education Expenditures for Elementary School-Aged Children

	(1) Household's Log Education Expenditures for Children Aged 6-13
Mill	4.98***
	(0.26)
Observations	3063

Note: Robust standard errors clustered at the district level are in parentheses. Robust standard errors clustered at the district level are in parentheses. All estimations include individual controls, cohort fixed effects, district fixed effects, district-by-year fixed effects, linear time trends interacted with baseline district level characteristics. Since EICV is not geocoded, the mill variable is mill per district area in km^2 in a given year. Sample consists of households that have elementary school aged children (aged 6-13). *** p<.01, ** p<.05, * p<.1

Table A.14: Effect of Mill Exposure on Hospitalizations for Gender Based Violence

	Women		Men	
	(1) Older than 18	(2) Younger than 18	(3) Older than 18	(4) Younger than 18
Mill x Month 1	0.05	-0.01	0.06	-0.03
	(0.07)	(0.05)	(0.05)	(0.06)
Mill x Month 2	0.00	0.00	0.00	0.00
	(.)	(.)	(.)	(.)
Mill x Month 3	-0.04	-0.12	-0.10	-0.05
	(0.06)	(0.10)	(0.06)	(0.08)
Mill x Month 4	-0.02	0.02	-0.08	0.01
	(0.08)	(0.06)	(0.07)	(0.09)
Mill x Month 5	-0.01	-0.02	0.01	0.03
	(0.08)	(0.11)	(0.07)	(0.10)
Mill x Month 6	-0.18**	-0.07	-0.02	-0.01
	(0.07)	(0.07)	(0.10)	(0.08)
Mill x Month 7	-0.14*	0.01	0.11	0.07
	(0.07)	(0.10)	(0.09)	(0.12)
Mill x Month 8	0.06	-0.00	-0.04	0.04
	(0.07)	(0.10)	(0.09)	(0.10)
Mill x Month 9	0.02	0.01	0.03	-0.05
	(0.07)	(0.11)	(0.09)	(0.09)
Mill x Month 10	0.00	0.03	-0.06	-0.02
	(0.05)	(0.09)	(0.07)	(0.11)
Mill x Month 11	-0.14	-0.06	-0.09	-0.02
	(0.11)	(0.08)	(0.07)	(0.08)
Mill x Month 12	-0.07	0.07	-0.06	0.00
	(0.11)	(0.09)	(0.09)	(0.07)
Observations	1210	1210	1209	1210
Dependent variable mean	0.79	0.84	0.24	0.14

Note: Robust standard errors clustered at the sector level are in parentheses. All estimations include hospital controls, hospital fixed effects, district fixed effects, province-by-year fixed effects, linear time trends interacted with baseline district level characteristics. Omitted categeory is the hospitalizations in Month 2, February. 4 km catchment area is used. *** p<.01, ** p<.05, * p<.1

Table A.15: Effect of Mill Exposure on Hospitalizations for Bone and Joint Disorders (Women)

	Bone and Joint Disorders
	(1) Women Older than 18
Mill x Month 1	0.14
	(0.18)
Mill x Month 2	0.00
	(.)
Mill x Month 3	0.07
	(0.14)
Mill x Month 4	-0.01
	(0.12)
Mill x Month 5	-0.08
	(0.16)
Mill x Month 6	-0.11
	(0.23)
Mill x Month 7	0.03
	(0.16)
Mill x Month 8	-0.02
	(0.19)
Mill x Month 9	-0.03
	(0.24)
Mill x Month 10	-0.04
	(0.22)
Mill x Month 11	-0.01
	(0.20)
Mill x Month 12	-0.03
	(0.19)
Observations	1078
Dependent variable mean	20.37

Note: Fractures are excluded. Examples include osteoarthritis, gout, rheumatoid arthritis, lupus, bursitis. Dependent variable is in logs. Robust standard errors clustered at the sector level are in parentheses. All estimations include hospital controls, hospital fixed effects, district fixed effects, province-by-year fixed effects, linear time trends interacted with baseline district level characteristics. Omitted categeory is the hospitalizations in Month 2, February. 4 km catchment area is used. *** p<.01, ** p<.05, * p<.1

Table A.16: Hospitalizations for Gender Based Violence Among Women Older than 19 across Different Catchment Areas

	Women Older than 19		
	(1)	(2)	(3)
	4 km	5 km	10 km
Mill x Month 1	0.05	0.05	0.11
	(0.07)	(0.06)	(0.10)
Mill x Month 2	0.00	0.00	0.00
	(.)	(.)	(.)
Mill x Month 3	-0.04	-0.04	0.08
	(0.06)	(0.06)	(0.10)
Mill x Month 4	-0.02	-0.01	0.06
	(0.08)	(0.07)	(0.11)
Mill x Month 5	-0.01	0.00	0.10
	(0.08)	(0.07)	(0.07)
Mill x Month 6	-0.18**	-0.16**	-0.12
	(0.07)	(0.07)	(0.08)
Mill x Month 7	-0.14*	-0.13	0.01
	(0.07)	(0.08)	(0.10)
Mill x Month 8	0.06	0.05	-0.01
	(0.07)	(0.07)	(0.09)
Mill x Month 9	0.02	0.03	-0.02
	(0.07)	(0.07)	(0.07)
Mill x Month 10	0.00	0.00	-0.05
	(0.05)	(0.04)	(0.07)
Mill x Month 11	-0.14	-0.11	0.08
	(0.11)	(0.10)	(0.10)
Mill x Month 12	-0.07	-0.04	0.08
	(0.11)	(0.10)	(0.10)
Observations	1210	1210	1210
Dependent variable mean	0.79	0.79	0.79

Note: Robust standard errors clustered at the sector level are in parentheses. All estimations include hospital controls, hospital fixed effects, district fixed effects, province-by-year fixed effects, linear time trends interacted with baseline district level characteristics. Omitted categeory is the hospitalizations in Month 2, February. *** $p<.01$, ** $p<.05$, * $p<.1$

Table A.17: Effect of Mill Exposure on Women's Employment and Type of Earnings

	Work Variables	
	(1)	(2)
	Work	Cash
Mill x Event Time -3	0.02	0.05
	(0.03)	(0.06)
Mill x Event Time -2	0.02	0.05
	(0.03)	(0.07)
Mill x Event Time -1	0.00	0.00
	(.)	(.)
Mill x Event Time 0	-0.01	0.10*
	(0.02)	(0.06)
Mill x Event Time 1	0.00	0.06
	(0.03)	(0.08)
Mill x Event Time 2	0.00	0.07
	(0.03)	(0.08)
Mill x Event Time 3	0.00	0.06
	(0.02)	(0.06)
Mill x Event Time 4	0.00	0.09**
	(0.02)	(0.04)
Mill x Event Time 5	-0.03	0.17***
	(0.03)	(0.06)
Mill x Event Time 6	0.01	0.10**
	(0.02)	(0.04)
Observations	7827	8163
Dependent variable mean	0.89	0.41

Note: Robust standard errors clustered at the sector level are in parentheses. All estimations include individual controls, cohort fixed effects, year of marriage fixed effects, sector fixed effects, district-by-year fixed effects, linear time trends interacted with baseline sector level characteristics. Sample consists of partnered women who married before the expansion of the mills. All variables are measured for the past 12 months. Ommitted categeory is Event Time -1, 1 year before a mill opening. 4 km catchment area is used. *** p<.01, ** p<.05, * p<.1

Table A.18: Effect of Mill Exposure on Self-Reported Domestic Violence

	(1) Domestic Violence
Mill x Event Time -3	0.01
	(0.07)
Mill x Event Time -2	-0.06
	(0.08)
Mill x Event Time -1	0.00
	(.)
Mill x Event Time 0	-0.16***
	(0.06)
Mill x Event Time 1	-0.17**
	(0.09)
Mill x Event Time 2	-0.12
	(0.09)
Mill x Event Time 3	-0.14**
	(0.06)
Mill x Event Time 4	-0.05
	(0.05)
Observations	3528
Dependent variable mean	0.35

Note: Robust standard errors clustered at the sector level are in parentheses. All estimations include individual controls, cohort fixed effects, year of marriage fixed effects, sector fixed effects, district-by-year fixed effects, linear time trends interacted with baseline sector level characteristics. Sample consists of partnered women who married before the expansion of the mills. All variables are measured for the past 12 months. Ommitted categeory is Event Time -1, 1 year before a mill opening. 4 km catchment area is used. *** p<.01, ** p<.05, * p<.1

Table A.19: Effect of Mill Exposure on Women's Employment and Type of Earnings across Different Catchment Areas

	4 km		5 km		10 km	
	(1) Work	(2) Cash	(3) Work	(4) Cash	(5) Work	(6) Cash
Mill x Event Time -3	0.01	0.03	0.00	-0.03	-0.02	-0.03
	(0.03)	(0.05)	(0.02)	(0.06)	(0.02)	(0.04)
Mill x Event Time -2	0.02	0.07	0.01	-0.00	-0.04	0.04
	(0.02)	(0.05)	(0.03)	(0.04)	(0.03)	(0.05)
Mill x Event Time -1	0.00	0.00	0.00	0.00	0.00	0.00
	(.)	(.)	(.)	(.)	(.)	(.)
Mill x Event Time 0	-0.01	0.04	-0.02	0.04	-0.03	0.04
	(0.02)	(0.06)	(0.02)	(0.04)	(0.02)	(0.04)
Mill x Event Time 1	0.01	0.06	0.01	0.06	0.01	0.03
	(0.03)	(0.08)	(0.02)	(0.06)	(0.03)	(0.06)
Mill x Event Time 2	-0.00	0.05	0.00	0.11*	-0.01	0.01
	(0.03)	(0.07)	(0.03)	(0.07)	(0.02)	(0.05)
Mill x Event Time 3	-0.01	0.07	-0.03	-0.01	-0.05*	-0.02
	(0.02)	(0.06)	(0.02)	(0.04)	(0.03)	(0.05)
Mill x Event Time 4	0.02	0.09**	-0.03	-0.01	-0.01	0.06
	(0.02)	(0.04)	(0.02)	(0.04)	(0.02)	(0.04)
Mill x Event Time 5	-0.04*	0.14**	-0.05**	0.01	0.00	0.09**
	(0.02)	(0.06)	(0.02)	(0.06)	(0.02)	(0.04)
Mill x Event Time 6	0.01	0.08**	-0.01	-0.00	-0.01	-0.03
	(0.02)	(0.04)	(0.02)	(0.04)	(0.02)	(0.04)
Observations	9823	10081	9823	10081	9823	10081
Dependent variable mean	0.88	0.39	0.88	0.39	0.88	0.39

Note: Robust standard errors clustered at the sector level are in paranthesis. All estimations include individual controls, cohort fixed effects, year of marriage fixed effects, sector fixed effects, district-by-year fixed effects, linear time trends interacted with baseline sector level characteristics. Sample consists of partnered women who married before the expansion of the mills. All variables are measured for the past 12 months. Ommitted categeory is Event Time -1, 1 year before a mill opening. *** p<.01, ** p<.05, * p<.1

Table A.20: Effect of Mill Exposure on Self-Reported Domestic Violence across Different Catchment Areas

	Domestic Violence		
	(1) 4 km	(2) 5 km	(3) 10 km
Mill x Event Time -3	0.00	-0.00	0.00
	(0.07)	(0.06)	(0.07)
Mill x Event Time -2	-0.06	-0.09	0.01
	(0.08)	(0.07)	(0.09)
Mill x Event Time -1	0.00	0.00	0.00
	(.)	(.)	(.)
Mill x Event Time 0	-0.16**	-0.09	-0.02
	(0.07)	(0.06)	(0.06)
Mill x Event Time 1	-0.14*	-0.08	-0.03
	(0.08)	(0.09)	(0.08)
Mill x Event Time 2	-0.13	-0.04	-0.05
	(0.09)	(0.08)	(0.07)
Mill x Event Time 3	-0.14**	-0.09	0.01
	(0.06)	(0.08)	(0.08)
Mill x Event Time 4	-0.05	-0.01	0.01
	(0.05)	(0.05)	(0.06)
Observations	3583	3583	3583
Dependent variable mean	0.35	0.35	0.35

Note: Robust standard errors clustered at the sector level are in parentheses. All estimations include individual controls, cohort fixed effects, year of marriage fixed effects, sector fixed effects, district-by-year fixed effects, linear time trends interacted with baseline sector level characteristics. Sample consists of partnered women who married before the expansion of the mills. All variables are measured for the past 12 months. Ommitted categeory is Event Time -1, 1 year before a mill opening. *** p<.01, ** p<.05, * p<.1

Table A.21: Effect of Mill Exposure on Women's Employment, Type of Earnings and Self-Reported Domestic Violence in the Past 12 Months for Different Catchment Area Sizes

	4 km			5 km			10 km		
	(1) Work	(2) Cash	(3) Violence	(4) Work	(5) Cash	(6) Violence	(7) Work	(8) Cash	(9) Violence
Mill	-0.00	0.07***	-0.09***	-0.01	0.04**	-0.07*	-0.01	-0.02	0.02
	(0.01)	(0.02)	(0.03)	(0.01)	(0.02)	(0.04)	(0.01)	(0.03)	(0.04)
Observations	9823	8766	3583	9823	8766	3583	9823	8766	3583
Dependent variable mean	0.88	0.40	0.35	0.88	0.40	0.35	0.88	0.40	0.35

Note: Robust standard errors clustered at the sector level are in parentheses. All estimations include individual controls, cohort fixed effects, year of marriage fixed effects, sector fixed effects, district-by-year fixed effects, linear time trends interacted with baseline sector level characteristics. Sample consists of partnered women who married before the expansion of the mills. All dependent variables are measured for the past 12 months. Catchment areas are constructed by buffers around the mills where 4, 5 and 10 km represent buffer radius. *** p<.01, ** p<.05, * p<.1

Table A.22: Effect of Mill Exposure on Household Decision Making: Woman Alone or Jointly with Husband

	4 km			5 km			10 km		
	(1) Large HH Purchases	(2) Own Health	(3) Family Visit	(4) Large HH Purchases	(5) Own Health	(6) Family Visit	(7) Large HH Purchases	(8) Own Health	(9) Family Visit
Mill	0.05**	0.02	0.02	0.04*	0.01	0.01	0.02	0.01	0.01
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Observations	9823	9823	9823	9823	9823	9823	9823	9823	9823
Dependent variable mean	0.69	0.73	0.81	0.69	0.73	0.81	0.69	0.73	0.81

Note: Robust standard errors clustered at the sector level are in parentheses. All estimations include individual controls, cohort fixed effects, year of marriage fixed effects, sector fixed effects, district-by-year fixed effects, linear time trends interacted with baseline sector level characteristics. Sample consists of partnered women who married before the expansion of the mills. All dependent variables are coded as 1 if the decision is taken by women alone or jointly with their husbands and zero otherwise. Otherwise category includes husband only or someone else. The survey question on household decisions changed significantly over the data cycles. in 2005, the question asks "Who has the final say on decision x?" After 2005, the question is "Who usually decides on x". The answer option "Respondent and other" is also discontinued in 2010 and 2014 rounds. To make a comparable variable, I create the variable based on whether the woman had a say in the decision, either alone or jointly. Results should be interpreted with caution. Cathment areas are constructed by buffers around the mills where 4, 5 and 10 km represent buffer radius. *** $p<.01$, ** $p<.05$, * $p<.1$

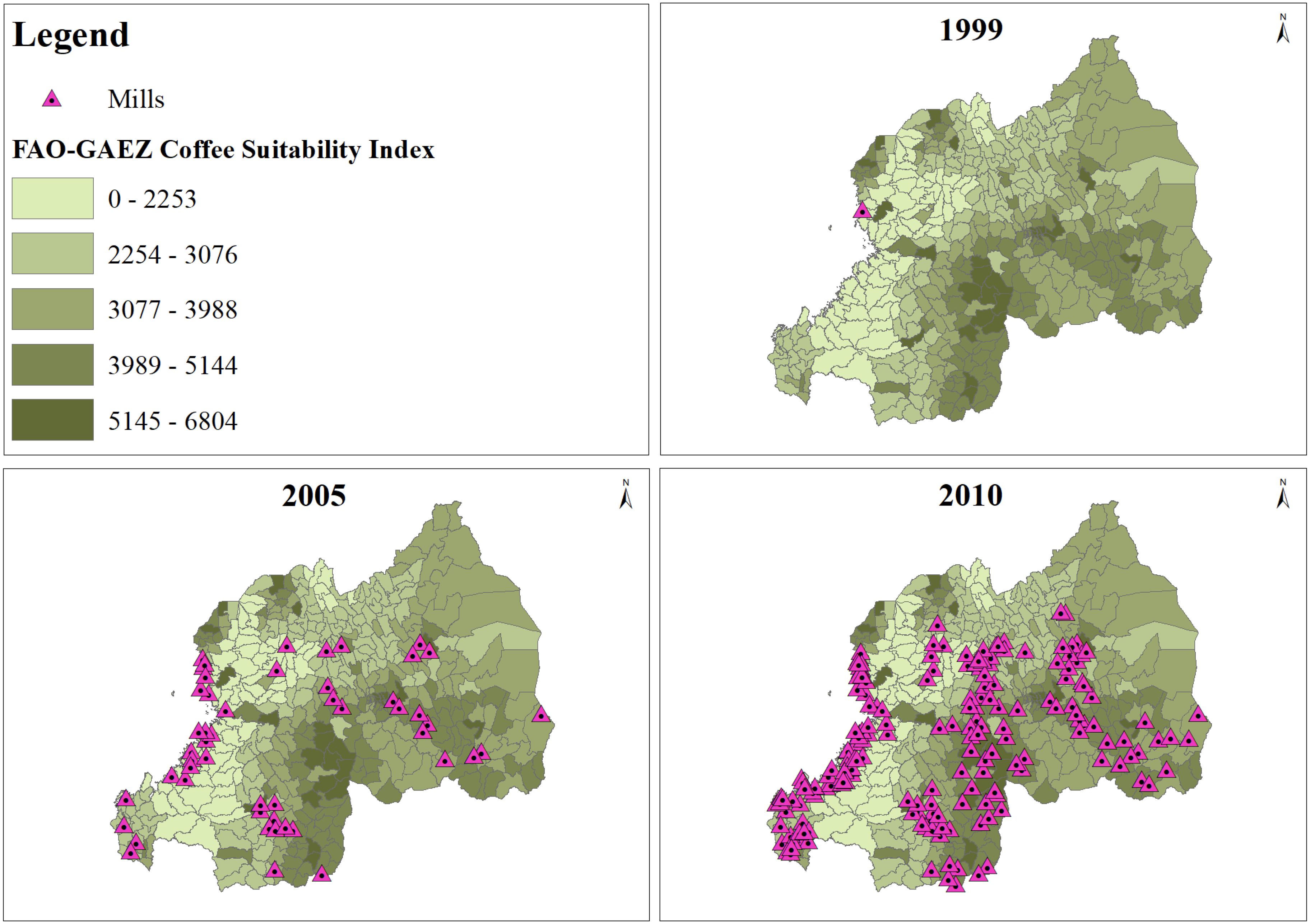


Figure A.1: FAO-GAEZ Coffee Suitability Index and Expansion of Mills in Rwanda

Note: The maps are constructed by combining data on mills, FAO-GAEZ Coffee Coffee Suitability Index, The Rwandan GeoPortal spatial data on sector boundries.

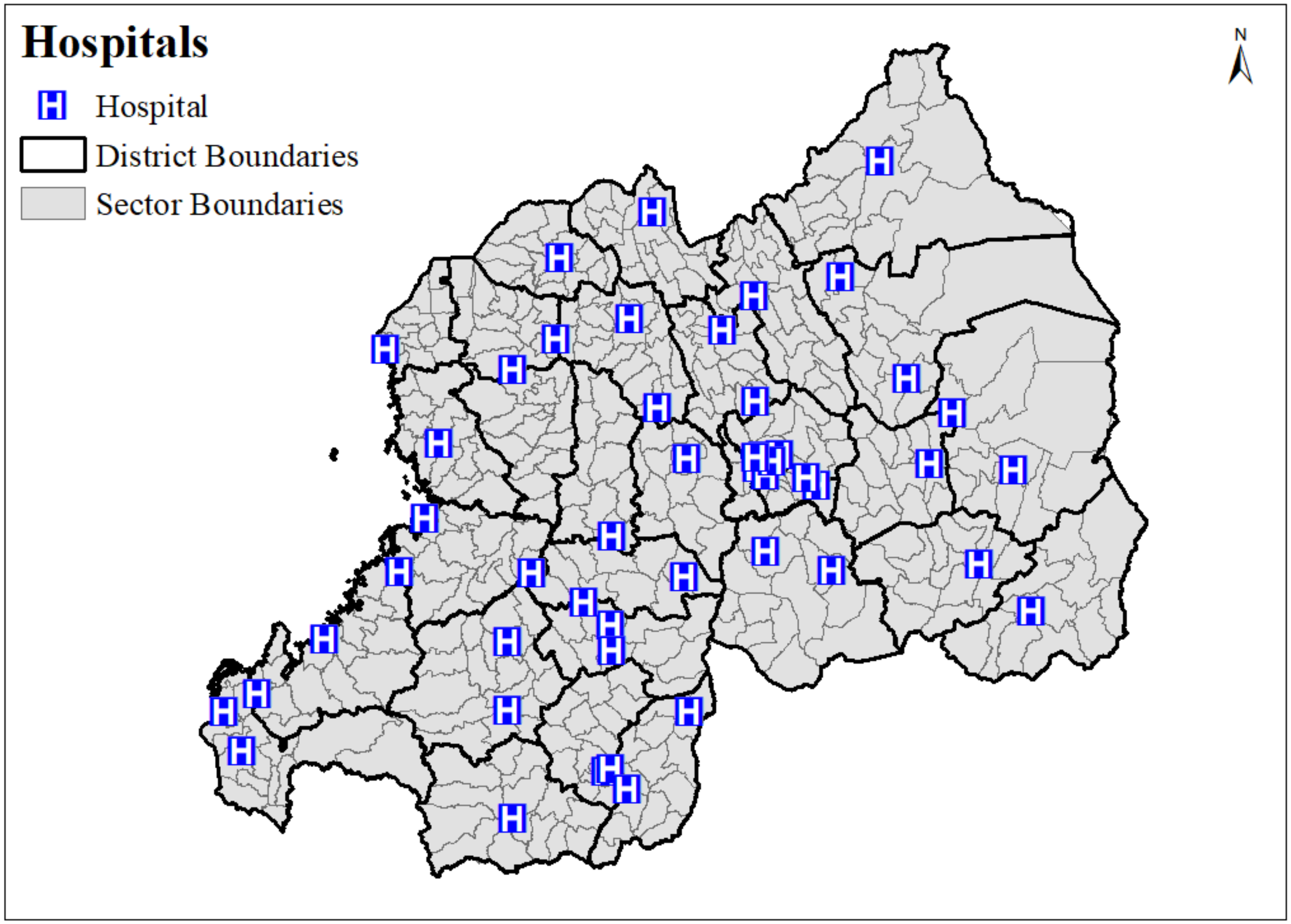


Figure A.2: Hospitals in Rwanda

Note: The map is constructed by combining HMIS data with the GPS coordinates of the hospitals provided by the Rwanda Ministry of Health.

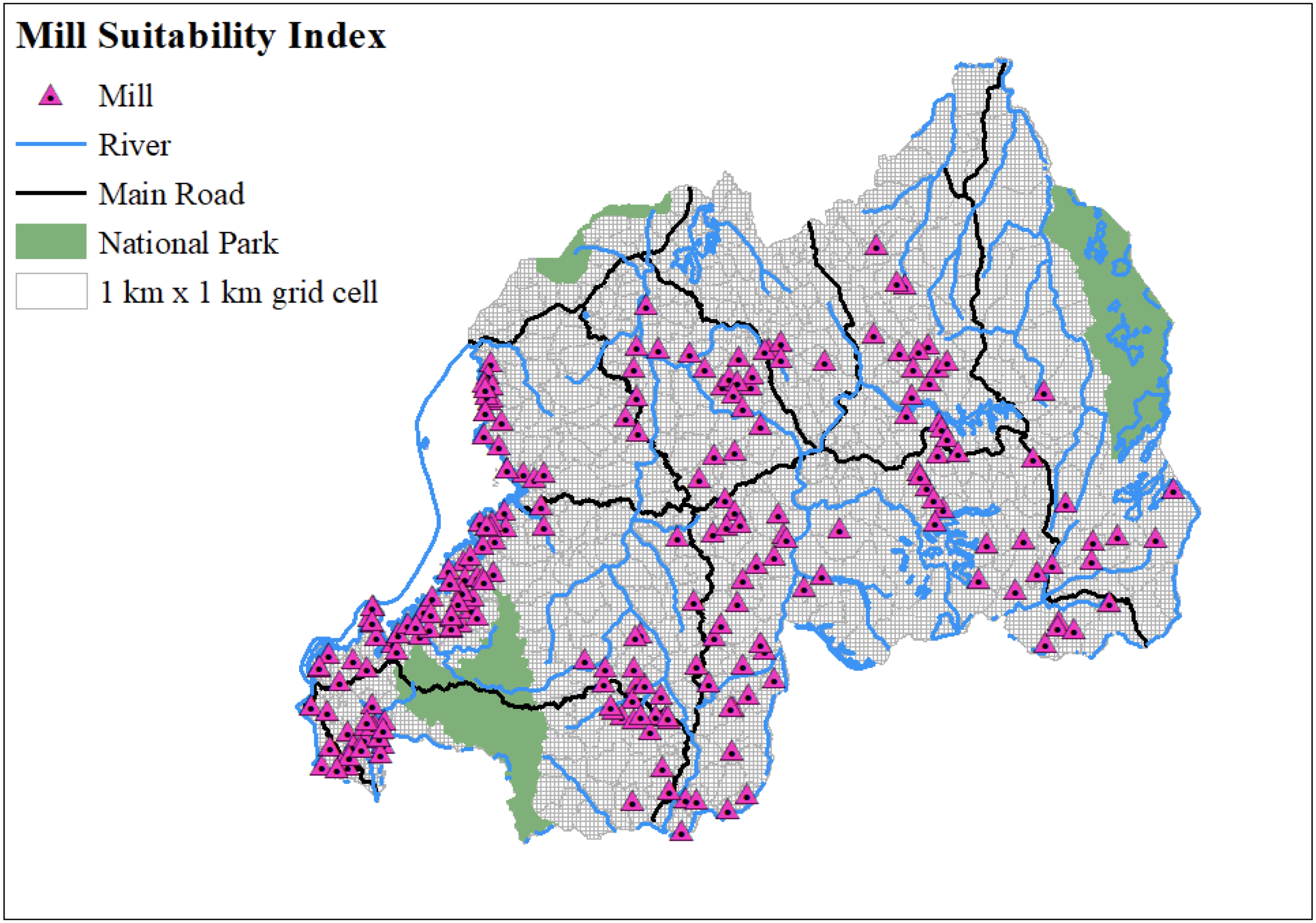
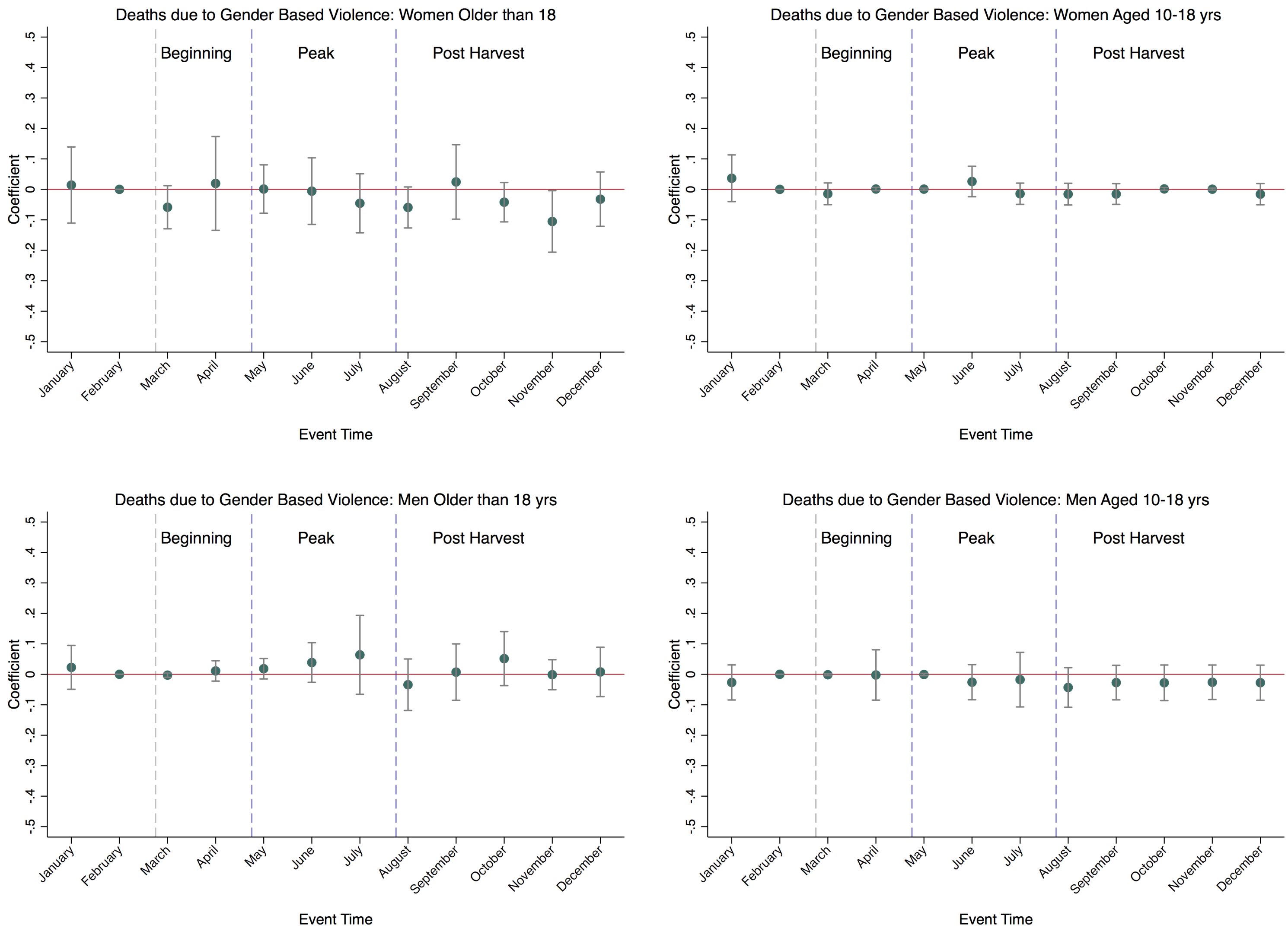


Figure A.3: Predicting a Mill Suitability Index with LASSO

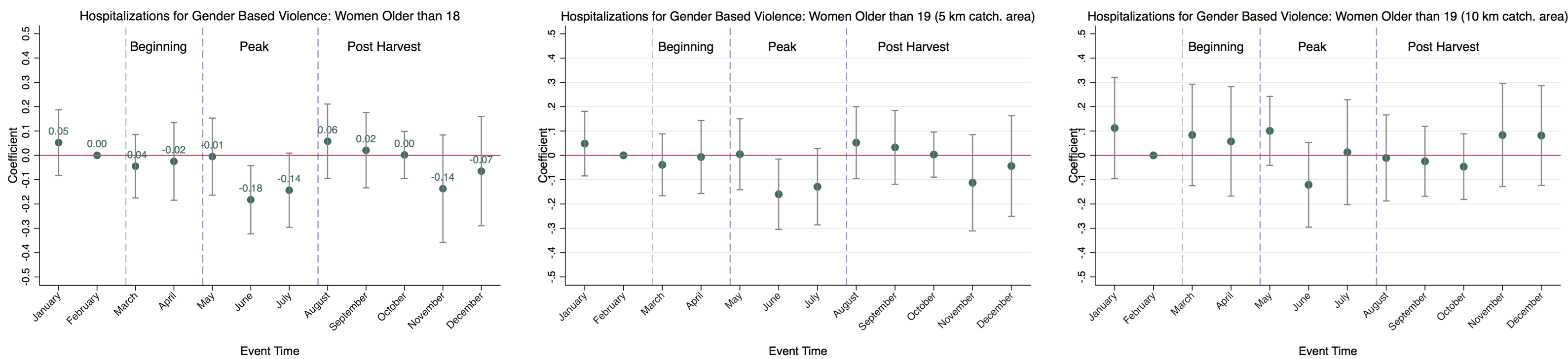
Note: The map is constructed by combining mills data and The Rwandan GeoPortal spatial data.

Figure A.4: Dynamic Impact of a Mill Opening on Deaths due to Gender Based Violence (Women and Men)



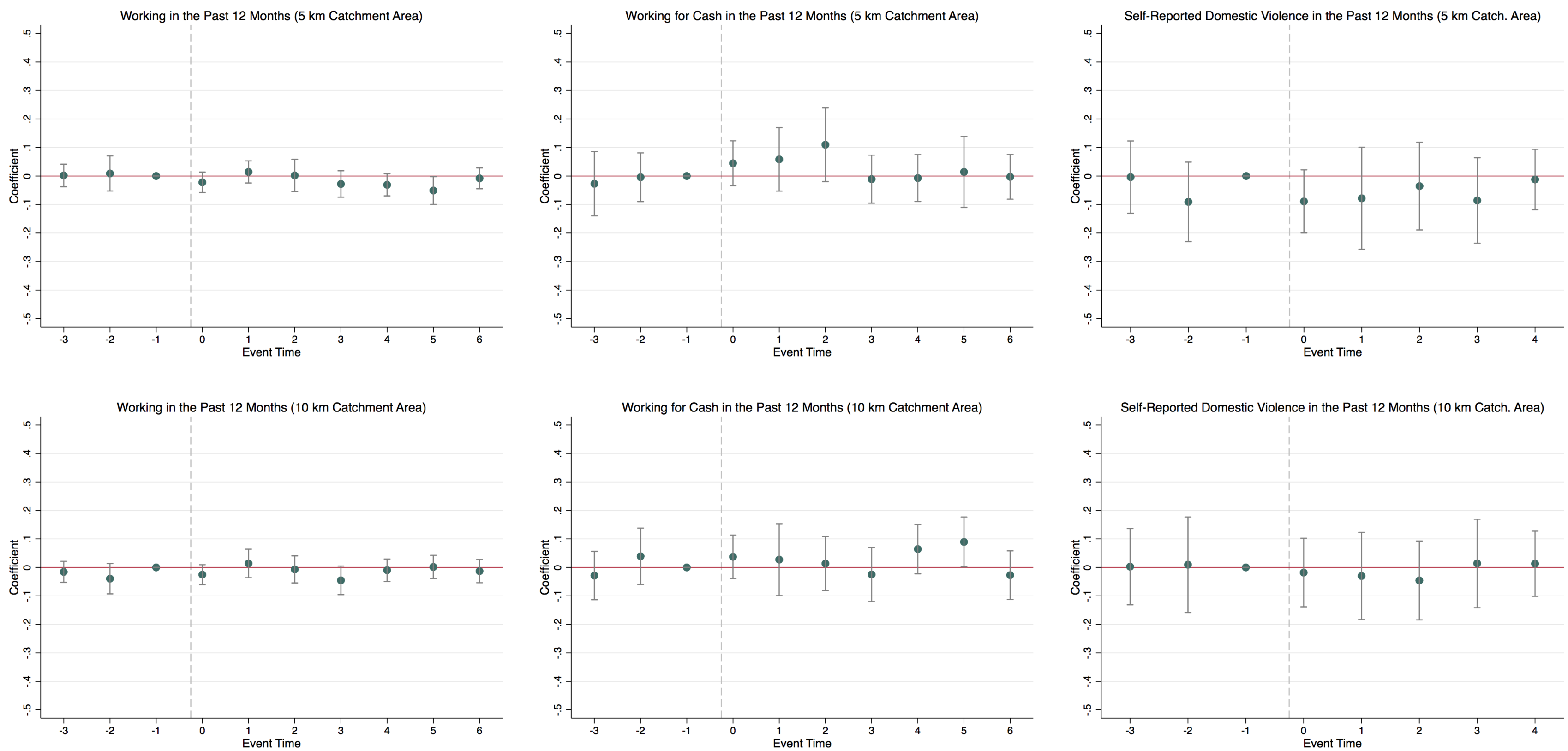
Note: Robust standard errors clustered at the sector level are in parentheses. All estimations include hospital controls, hospital fixed effects, district fixed effects, province-by-year fixed effects, linear time trends interacted with baseline district level characteristics. 4 km catchment area is used. *** p<.01, ** p<.05, * p<.1

Figure A.5: Dynamic Impact of a Mill's Period of Operation on Hospitalizations for Gender Based Violence (Women Aged Older than 19) for Different Catchment Areas



Note: Robust standard errors clustered at the sector level are in paranthesis. All estimations include hospital controls, hospital fixed effects, district fixed effects, province-by-year fixed effects, linear time trends interacted with baseline district level characteristics. 4 km, 5km and 10 km catchment areas are used respectively. *** p<.01, ** p<.05, * p<.1

Figure A.6: Dynamic Impact of a Mill Opening on Women's Employment, Type of Earnings and Self-Reported Domestic Violence in the Past 12 Months for Different Catchment Areas



Note: Robust standard errors clustered at the sector level are in paranthesis. All estimations include individual controls, cohort fixed effects, year of marriage fixed effects, sector fixed effects, district-by-year fixed effects, linear time trends interacted with baseline sector level characteristics. Sample consists of partnered women who married before the expansion of the mills. All variables are measured for the past 12 months. 5 km and 10 km catchment areas are used respectively. *** p<.01, ** p<.05, * p<.1

B Theoretical Appendix

In the theoretical appendix, I show that the derivatives in Section 3 cannot be signed. The probability of the husband choosing violence is

$$F_h\left[(1-y)I_h-(1-x+T)I_w-T\alpha_h(I_h,I_w)+\frac{T[I_w+\alpha_h(I_h,I_w)]}{F_w[(1-x+T)I_w-(1-y)I_h+v]}\right]. \tag{12}$$

B.1 Increase in I_w

$$\frac{\partial F_h}{\partial I_w}=f_w\left[(1-y)I_h-(1-x+T)I_w-T\alpha_h(I_h,I_w)+\frac{T[I_w+\alpha_h(I_h,I_w)]}{F_w[(1-x+T)I_w-(1-y)I_h+v]}\right]\times$$

$$\left[-(1-x+T)-T\frac{\partial\alpha_h(I_h,I_w)}{\partial I_w}+\right.$$

$$\left.\frac{T(1+\frac{\partial\alpha_h(I_h,I_w)}{\partial I_w})F_w[(1-x+T)I_w-(1-y)I_h+v]-f_w[(1-x+T)I_w-(1-y)I_h+v](1-x+T)T[I_w+\alpha_h(I_w,I_h)]}{F_w[(1-x+T)I_w-(1-y)I_h+v]^2}\right].$$

Assume $F_w(.)$ and $f_w(.)\neq 0$. Then the first line of the product is positive. The derivative is negative when $\frac{\partial\alpha_h(I_h,I_w)}{\partial I_w}$ is sufficiently small and the numerator of the fraction in the third line of the product is negative. Suppose $\frac{\partial\alpha_h(I_h,I_w)}{\partial I_w}$ is sufficiently small. Then the condition for for the derivative to be negative is

$$\left(1+\frac{\partial\alpha_h(I_h,I_w)}{\partial I_w}\right)F_w[(2-x)I_w-(1-y)I_h+v]-f_w[(2-x)I_w-(1-y)I_h+v](2-x)(I_w+\alpha_h(I_w,I_h))<0.$$

Given that F_w is a CDF and $F_w(.)\neq 0$ by assumption, $0<F_w(.)\leq 1$. $f_w(.)>0$ by assumption. $(2-x)>0$ due to $x\in(0,1)$. Suppose that the above expression is positive. Then,

$$\left(1+\frac{\partial\alpha_h(I_h,I_w)}{\partial I_w}\right)F_w[(2-x)I_w-(1-y)I_h+v]>(I_w+\alpha_h(I_h,I_w))f_w[(2-x)I_w-(1-y)I_h+v](2-x).$$

$0<F_w(.)\leq 1$. There are values which makes the condition both hold and does not hold (For sufficiently high I_w the condition does not hold, yet for sufficiently small I_w the condition can hold).The derivative cannot be signed.

B.2 Increase in I_h

$$\frac{\partial F_h}{\partial I_h} = f_h\left[(1-y)I_h-(1-x)I_w-T\alpha_h(I_h,I_w)+\frac{T[I_w+\alpha_h(I_h,I_w)]}{F_w[(1-x+T)I_w-(1-y)I_h+v]}\right]\times\left[(1-y)-T\frac{\partial\alpha_h(I_h,I_w)}{\partial I_h}+\right.$$

$$\left.\frac{T\frac{\partial\alpha_h(I_h,I_w)}{\partial I_h}F_w[(1-x+T)I_w-(1-y)I_h+v]-f_w[(1-x+T)I_w-(1-y)I_h+v](1-y)T(I_w+\alpha_h(I_h,I_w))}{F_w[(1-x+T)I_w-(1-y)I_h+v]^2}\right].$$

Assuming that $F_h(.)$ and $f_h(.)>0$, the left hand side is both positive. The right hand side of the product can be both positive or negative because $F_h(.)>0$, $f_h(.)>0$, $(1-y)>0$, $T=1$ and $\frac{\partial\alpha_h(I_h,I_w)}{\partial I_h}<0$. Specifically $(1-y)-T\frac{\partial\alpha_h(I_h,I_w)}{\partial I_h}>0$, but the remaining of the right hand side term is negative.