University of Warsaw

Faculty of Management

Mudassira Sarfraz

Social Norms and Women‘s Labor Force Participation: Evidence from Pakistan

Ph.D Thesis

In the Discipline of Economics

Dissertation was written under the supervision of

Prof. dr hab Joanna Tyrowicz

University of Warsaw, Faculty of Management

Warsaw, 2021

Studia doktoranckie – ścieżka angielska - **projekt** „Zarządzanie wielokulturowe w erze globalizacji” realizowany przez Wydział Zarządzania Uniwersytetu Warszawskiego na podstawie umowy nr POWR.03.02.00-00-I053/16-00 w ramach Programu Operacyjnego Wiedza Edukacja Rozwój 2014-2020 finansowanego z funduszy strukturalnych Unii Europejskiej

Oświadczenie kierującego pracą

Oświadczam, że niniejsza praca została przygotowana pod moim kierunkiem i stwierdzam, że spełnia ona warunki do przedstawienia jej w postępowaniu o nadanie stopnia doktora.

Data Podpis kierującego pracą

Statement of the Supervisor on Submission of the Dissertation

I hereby certify that the thesis submitted has been prepared under my supervision and I declare that it satisfies the requirements of submission in the proceedings for the award of a doctoral degree.

Date Signature of the Supervisor

Oświadczenie autora pracy

Świadom odpowiedzialności prawnej oświadczam, że niniejsza praca doktorska została napisana przeze mnie samodzielnie i nie zawiera treści uzyskanych w sposób niezgodny z obowiązującymi przepisami.

Oświadczam również, że przedstawiona praca nie była wcześniej przedmiotem procedur związanych z uzyskaniem tytułu zawodowego w wyższej uczelni.

Oświadczam ponadto, że niniejsza wersja pracy jest identyczna z załączoną wersją elektroniczną.

Data Podpis autora pracy

Statement of the Author on Submission of the Dissertation

Aware of legal liability I certify that the thesis submitted has been prepared by myself and does not include information gathered contrary to the law**.**

I also declare that the thesis submitted has not been the subject of proceeding in the award of a university degree.

Furthermore I certify that the submitted version of the thesis is identical with its attached electronic version.

Date Signature of the Author

Zgoda autora pracy

Wyrażam zgodę ma udostępnianie mojej rozprawy doktorskiej dla celów naukowo-badawczych.

Data Podpis autora pracy

Author's consent dissertation

I agree to make my dissertation available for research purposes

Date Signature of the Author

Abstract

The thesis identifies underlying bi-directional causal mechanisms that link equitable gender role beliefs (EGRB), operationalized as social norms affirmative of women's empowerment, with women's labor force participation (LFP) by utilizing microdata from Pakistan. There are two independent empirical studies. The first study uses neoclassical labor supply and modernization theories to investigate the hypothesis that more EGRB cause an increase in LFP among women. This hypothesis is addressed by utilizing the Instrumental Variable (IV) method on four rounds of the Pakistan Social and Living Standard Measurement (PSLM) survey. The second study uses the gender stratification theory to investigate the hypothesis that higher LFP among women causes gender role beliefs to become more equitable. This study exploits a quasi-natural experiment from Benazir Income Support Program (BISP) implemented in Pakistan to address this hypothesis. Overall, the results from both the studies confirm the presence of bi-directional causality between equitable gender role beliefs and women's LFP in Pakistan, suggesting a feedback loop and virtuous cycle between these two processes. The thesis also offers insights into self-reporting bias gender norm surveys, as well as implications for policy tools needed to ameliorate social stigma associated with women's labor force participation in countries with inequitable gender role beliefs.

Key words

Social norms, equitable gender role beliefs, women's LFP, causal identification, instrumental variables, PSLM, BISP, extreme weather events, Pakistan

Normy społeczne a aktyywnośc zawodowa kobiet: przypadek Pakistanu

Streszczenie

Celem rozprawy jest ustalenie przyczynowości w związkach pomiędzy aktywnością zawodową kobiet oraz normami społecznymi. Na gruncie teorii, przyczynowość może przebiegać w obu kierunkach. Wykorzystując dane jednostkowe z poziomu gospodarstw domowych i osób w Pakistanie, w rozprawie poddano weryfikacji czy i jakim stopniu występuja zależności przyczynowo-skutkowe pomiędzy tymi kategoriami. W pracy przeprowadzono dwa niezależne badania. W pierszym z nich wykorzystujemy neoklasyczną teorię podaży pracy oraz teorię modernizacji, by postulować, że bardziej równościowe normy społeczne zwiększają uczestnictwo kobiet w rynku pracy. Do weryfikacji tej tezy wykorzystano nowozaproponowane zmienne instrumentalne oraz dane z badania Pakistan Social and Living Standard Measurement (PSLM). survey. Drugie badanie, na gruncie teorii stratyfikacji, formułuje tezę, że większa aktywnosć zawodowa kobiet sprzyja bardizej równościowym normom społecznym. W tym badaniu bazujemy na quasi-naturalnym eksperymencie z program transferów społecznych Benazir Income Support Program. Wyniki obu badań sugerują występowanie wzajmnie wzmacniających się mechanizmów prorównościowych: większa aktywność zawodowa sprzyja bardziej równym normom społecznym, podczas gdy prorównościowe normy zwiekszają aktywność zawodową. Dodatkowymi wynikami rozprawy są także wnioski dotyczące obciążeniach w badaniach norm równościowych a także instrumetów polityki pozwalających skutecznie zmniejszać stygmatyzację społeczną aktywności zawodowej kobiet.

Słowa kluczowe

Normy społeczne, przekonania dotyczące równości płci, aktywność zawodowa kobiet, Pakistan

TABLE OF CONTENTS

ACKNOWLEDGMENT

I praise and thank Allah SWT for His greatness and giving me the strength and courage to complete this thesis.

I would like to express my gratitude to a few persons who inspired and helped throughout my research work. First and foremost, I present my sincere acknowledgment to my supervisor, prof. dr hab. Joanna Tyrowicz. Her competent supervision and thorough guidance significantly contributed to the successful completion of this thesis. Her encouragement, constructive comments, and guidance enhanced my research and greatly influenced my life experience. I am also thankful to the academic staff and my Ph.D. fellows at the Faculty of Management, University of Warsaw, who have always helped and supported me throughout my Ph.D. duration.

My research would have been impossible without the constant motivation from my parents, Sarfraz Khan and Shahjahan Begum, whose prayers and best wishes have enabled me to carry out this work and specially my father and wanted me to be an empowered woman. Words are not enough to express my sincere thanks to my husband, Kamran, who has always facilitated me. His love, devotion, and belief in me have enabled me to complete the most demanding academic task. Last but not least, I express my deepest love to my daughter, Maryam, and my son Hassan, who has given me extra strength and motivation and to get all the things done.

CHAPTER 1 INTRODUCTION

Many studies have recognized significant economic gains from women's ability to fully utilize their labor market potential. The most crucial potential progress at the macro level is the prospect of higher economic growth because of higher productivity as women join the labor market (Aguirre et al., 2012), which reduces poverty (International Labor Organization, 2016). Women's labor force participation (LFP) enhances the quality of life and provides additional sources to earn livelihoods. Besides this, it serves as an essential driver and outcome of economic growth and promotes human resource development. At the micro-level, many studies show that participation in the labor market enhances opportunities for the current and future generations (Luke and Munshi, 2011). Against this background, United Nations reports document that female labor force participation (FLFP) hoovers close 25% in South Asia, North Africa, and West Asia (United Nations, Human Development Report, 2015). In the background studies, United Nations argues that the major obstacles for women to engage in the labor markets are related to socio-cultural norms and household responsibilities. This line of argumentation is based both on theoretical premises and on empirical evidence.

In terms of theory, several major theories predict the correlation between women's LFP and social norms. The neoclassical theory explains that the supply of women is determined by a rational choice made at the household level considering the utility gained from labor market participation and the utility from performing household activities (Mincer, 1962; Becker, 1965). By contrast, modernization theory which builds on human capital theory establishes that increasing women's LFP results from structural changes in the labor market, growing educational level, and declining fertility rates (Goldin, 1990; Blau, Ferber, and Winkler, 1992). Thus, both the neoclassical and modernization theories point out that the causality runs from social norms to women's LFP: LFP decisions are conditioned by equitable gender role beliefs. In opposition to these approaches, gender stratification theory proposes that women's participation in economic production and control over material resources can increase bargaining power, motivating a shift to more equitable gender norms (Blumberg 1988, and Chafetz 1988). Therefore, gender stratification theory posits that the causality runs from higher women's LFP to equitable gender role beliefs (EGRB). To summarize, while all theories

predict the correlation between women's LFP and social norms, they imply opposite directions of causality.

My purpose in this thesis is to verify empirical support for both these directions for the case of Pakistan. The research objective of this thesis is to establish and identify underlying causal mechanisms that link equitable gender role beliefs (EGRB) and women's labor force participation (LFP). The specific objectives of this research are as follows:

- To examine the causal effect of equitable gender role beliefs on women's LFP
- To investigate the causal impact of women's LFP on equitable gender role beliefs

This research operationalizes equitable gender role beliefs as social norms affirmative of women's empowerment. Accordingly, following empirical hypotheses are formulated.

H1: More equitable gender role beliefs cause an increase in LFP among women.
H2: LFP among women causes gender role beliefs to become more equitable.

These two hypotheses are addressed in two independent empirical studies, each with its own, independent identification strategy. Given the formulation of these two hypotheses, four potential outcomes are possible. If we find evidence corroborating both hypotheses, we will prove that correlations are biased estimators of the relationship, and reverse causality invalidates most correlational studies on women's LFP and equitability of gender role beliefs in Pakistan in the period covered by my studies. However, if the data reject either of the hypotheses, the reverse causality circle would be broken, and correlations could be informative of the underlying links between social norms and FLFP. Finally, suppose we find no support for both hypotheses. In that case, our research can serve as evidence that better data and more robust identification strategies are needed before mechanisms relating to EGRB and women's LFP are uncovered.

1.1 The case of Pakistan

Pakistan is the fifth most populous country globally, located in South Asia, with an estimated population of 207.7 million. Regarding labor market statistics, Pakistan's total labor force is 65.5 million, 50.74 million are men, and 14.76 million are women

(Pakistan Bureau of Statistics, 2020). Concerning its geographic location, Pakistan is among the countries that are heavily exposed to erratic weather events. The Global Climate Risk Index has placed Pakistan on the fifth spot on the list of countries most vulnerable to climate change in its annual report for 2020 (Eckstein, Kunzel, Schafer & Winges, 2020). While Pakistan ranks as one of the poorest countries in the world, with GDP per capita of roughly 5000$ per year in PPP terms, it has experienced two decades of tremendous economic growth, with GDP per capita essentially tripling between 2000 and 2020.

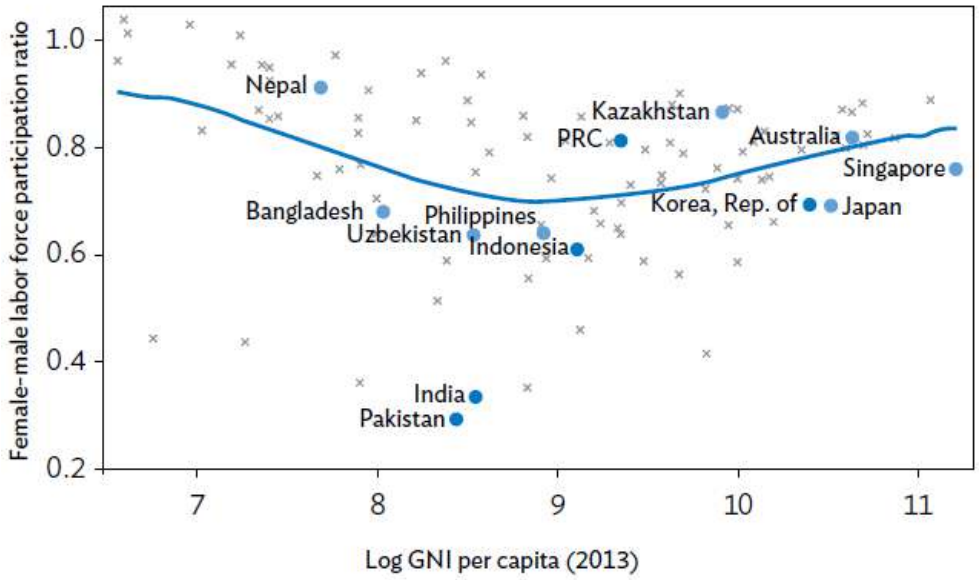


Figure 3.1 Female Labor Force Participation and GNI per capita Worldwide
Source: (Tanaka & Muzones, 2016). Page Number 2, Figure 1.

During this period of rapid economic growth, women's LFP has increased moderately: from 14% in 2001 to approximately 20% at the end of the 2010s (Pakistan Bureau of Statistics, 2018). Indeed, Pakistan is an outlier among developing countries regarding FLFP: it is the lowest among countries in the South Asian region (Tanaka & Muzones, 2016). FLFP in Pakistan is lower than implied by its level of development (Figure 1.1). Moreover, women's LFP rate is higher in urban areas than in rural areas. The rural LFP is inflated due to unpaid/contributing family workers approximately 55% of the total women's employment in 2015 (Pakistan Bureau of Statistics, 2015). However, this percentage reduced to 52 in 2018. At the provincial level, FLFP ranges from 26.5% (Punjab), 12.1% (Sindh), 11.3% (KP), and 7.9% (Balochistan) (Pakistan Bureau of

Statistics, 2018). According to a study conducted by Tanaka & Muzones (2016), 40% of women in Pakistan report that the main reason they do not work is the restriction by their male family members to work outside the home. Several interrelated aspects limit women's mobility outside the home, among them (i) social, cultural, and religious norms; (ii) safety and crime; and (iii) the quality of available transport services (Tanaka & Muzones, 2016). 25% of women with university degrees work outside the home, which means even if women earn a tertiary degree, their labor force participation is low. Such circumstances make Pakistan part of what Kandiyoti (1988) calls the "belt of classic patriarchy," which includes North Africa, the Muslim Middle East (including Turkey and Iran), and South and East Asia (Pakistan, Afghanistan, Northern India, and rural China).

This thesis will utilize two large-scale representative household surveys: cross-sectional Pakistan Social and Living Standard Measurement Survey (PSLM) and panel Benazir Income Support Program (BISP) to investigate the bidirectional causality between EGRB and women's LFP. It operationalizes EGRB as social norms affirmative of women's empowerment. The research will also exploit the fact that Pakistan experiences the negative consequences of climate change heavily. This allows constructing an instrument for EGRB using geographical data of high granularity on extreme weather events over the past several decades.[1] This identification strategy, described in detail later, will address how EGRB cause LFP among women. Subsequently, this thesis exploits a quasi-natural experiment from an income support program (BISP) implemented in Pakistan in the same period to provide an empirical evaluation of the causal link that LFP among women causes males' gender role beliefs more equitable. These causal directions will be tested using an econometric method of causal identification, i.e., instrumental variables (IV). Additionally, the research will also apply the regression discontinuity design (RDD) approach on BISP data to check the causal impact of this cash transfer program on women's LFP.

1.2 Theoretical background

The neoclassical theory explains that female labor supply is determined by the rational choice made at the domestic level, considering the utility gained from labor market

[1] Data on frequency of extreme weather events comes from EM-DAT, launched by Center for Research on the Epidemiology of Disasters (CRED)

participation and the utility from performing household activities. Although the labor supply theory considers labor supply as a choice between work and leisure, the rise in the quantity of women workers has prompted the discussion around the role of household production. Jacob Mincer (1962) discovered the basics of the neoclassical model by combining the theoretical developments in econometrics and labor supply. Subsequently, Garry S. Becker (1965) appeared with the household production model in addition to female time allocation that substituted the conventional theory of labor supply.

This theory suggests that women would be willing to spend time in the labor market if the associated utility is more than the utility of time doing household work. Several factors affect the utility obtained from the market and non-market activities. Significant determinants for assessing these utilities include tastes and preferences, housework requirements, infants' presence, and income availability from other sources like spouse income. Marital status decides if other alternate sources of income are there, and the need for care increases with children in the house (Blau et al., 2014). Moreover, neoclassical theory is based on the notion that household preferences can be characterized by a single utility function and assumes that household members have the same choices and pool their resources to maximize the single utility function (Becker 1981). The neoclassical theory would seem to hold in Pakistan because the prevailing traditional cultural restrictions on women often position the male head as the household decision-maker (Hakim and Aziz 1998).

The research also utilizes Becker’s marriage market hypothesis (Becker, 1981) to explain how variation in gender ratio affects women’s LFP. Marriage is a long-term agreement between two persons that results in valued yet partly intangible yields: love, security, children, companionship, salary from the job, and domestic properties from household production (Becker; 1973, 1981). One spouse is often dedicated to market work in a marital relationship, whereas the other fulfills the household responsibilities. The share of resources needed for managing the home is assigned to the partner performing labor at home. If the number of marriageable women is higher than the number of marriageable men, women's bargaining power would be lower in the marriage market (Angrist; 2002, Chiappori et al.; 2002). This condition reduces the ‘price’ that a breadwinner partner is obliged to pay as compensation to the partner performing household labor, shifting family resources in favor of men. Women having

fewer resources within marital relationships will have lower reservation wages for participation in the labor market, making them more likely to work for pay (Grossbard-Shechtman,1984). This hypothesis and neoclassical labor supply theory are used as theoretical foundations to explore whether exogenous variation in gender ratio due to climate-induced migration cause women's LFP to increase, thereby making traditional social norms costly to persist.

The central idea of this research is also based on modernization theory to explain women's LFP in Pakistan. Modernization theory, with its foundation on human capital theory, describes educational expansion, economic development, a decline in fertility, structural changes in the labor market, and related processes as indicators of modernization in a society. It argues that all such indicators favor women's labor market activity and make women's employment in the formal labor market more profitable (e.g., Blau et al. 2014; Goldin 1990). Apart from these indicators, urbanization is also an indicator of modernization in society. In Pakistan, the urbanization rate is 3% per annum (Kugelman, 2014); cities have become densely populated, with 35% to 50% of the urban population residing in informal settlements (Kugelman, 2014). Besides urbanization, Pakistan is also experiencing economic growth, educational expansion, low gender inequality, and low fertility rates (Tanaka & Muzones, 2016). However, the country has had almost stagnant women's LFP in the last two decades. Therefore, this research will extend the arguments of neoclassical theory and marriage market hypothesis and explore theoretical reasoning of modernization theory to explain women's LFP.

1.3 Causality between women' LFP and equitable gender role beliefs

One of the reasons for the rise in women's LFP stems from the studies on migration; the idea that is closely linked with modernization theory which Karl Marx and Max Weber originated and further revived during the 1950s and 1960s by Seymour Martin Lipset, Daniel Lerner, Walt Rostow and Karl Deutsch (Inglehart and Norris 2003). Migration may take different forms, i.e., internal (rural-rural or rural-urban) and international (cross border), and there can be various reasons why people migrate. Besides this, there is also a gender dimension of migration. Conventional economic approaches consider migrants as genderless agents who respond in a homogeneous way to macro-level socioeconomic changes, whereas feminist studies suggest that the migrant's gender

identity and the existing social norms determine the migration pattern (Chant 1991; Gisbert, Painter, and Quiton. 1994; Tacoli 1999; Hampshire 2002).

This research intends to explore another strand of the theory that provides the foundations for investigating the causal direction from FLFP to more equitable gender norms. The gender stratification theory by Blumberg (1984) provides the theoretical justification to explore this link. The theory argues that gender stratification is inversely linked to women's access to material resources and the level of women's economic power. An increase in women's access to resources and their ability to participate in the labor market can reduce their ideological and physical oppression in societies. Similar theoretical foundations are provided by the bargaining model, which is contrary to the idea of neoclassical theory's single utility function of the household. The bargaining approach considers that different household members have distinct preferences, allows men and women to have individual agency within a household, and provides evidence of changing bargaining power (Lundberg and Pollak 1996).

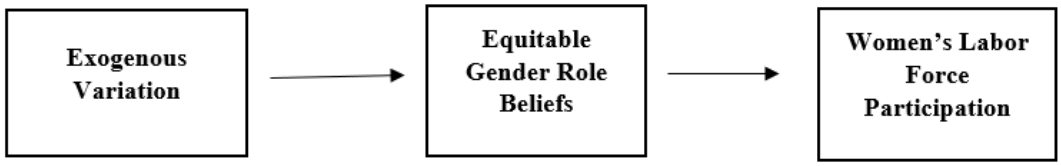


Figure 3.2 Causal impact of equitable gender role beliefs on women' LFP

In the case of Pakistan, urbanization/migration is at an annual rate of 3%. Metropolitan cities have the highest rate of urbanization as the population of Karachi, one of the largest cities in Pakistan, increased from 80 % from 2000-2010, termed as the most significant rise of any municipality in the world (Kotkin & Cox, 2013). The reasons for internal migration were found to be diverse and multiple. Migration depends on several economic factors, such as better employment opportunities in urban areas and low wages in rural areas. Other social factors, such as better health and education facilities and overall good living standards in urban areas, motivate many rural households to migrate. However, the principal reason is the motive to expand livelihood strategies and mitigate financial risks associated with stresses and shocks, mainly climate disasters (Ishfaq, 2019) which, as mentioned in the beginning, are continuously happening in

Pakistan. A similar argument was made in a seminal work by Mueller, Gray, & Kosec (2014), who in their 21- years (1991-2012) longitudinal study concluded that, as compared to other variables, extremely high temperatures, rainfall, and flood intensity are strongly linked with rural-urban migration.

A study by Ishfaq (2019) collects qualitative and quantitative data from rural areas of eight districts of Pakistan: Mardan, Hunza, Nagar, Sargodha, Chakwal, Dera Ghazi Khan, Faisalabad, and Rawalpindi. They carried out Focus Group Discussions, face-to-face interviews, and questionnaires to gather information from 1019 non-migrant and migrant households. Their objective was to explore the link between migration and climate change, such as the changes in rainfall and temperature and change in frequency of extreme climate events (floods, heat waves, droughts, etc.). They found migration to be a common strategy to mitigate climatic and non-climatic shocks. However, migration was generally male-dominated as a livelihood strategy, with women often excluded (Ishfaq, 2019).

The study's findings mentioned above can be interpreted to reveal the phenomenon of climate-induced male out-migration in Pakistan, which could potentially affect the gender ratio in areas that are heavily exposed to extreme weather events. The idea is that when a severe weather event hits a region or district, it destroys the means of income generation for people living in those districts. Consequently, as an adaptation strategy to such weather events, people are forced to relocate to other unaffected districts because it is implausible that these inhabitants will continue to live in those districts under high uncertainty of harsh climatic conditions. As this migration is in general male-dominated (Ishfaq, 2019), this would result in a lower gender ratio, i.e., male to female ratio, in those disaster-hit districts. Women who are left behind would, in turn, “substitute missing males” (in words of Cardoso and Morin, 2019) and eventually become part of the labor market. Hence, this male migration could unintentionally improve labor market conditions for women.

This means that certain circumstances threaten the sustainability of social norms. Bertrand, Kamenica & Pan (2015) investigate the similar argument that “how are gender identity norms evolving in the face of market forces that are making those norms increasingly costly?” (Bertrand et al., 2015, p. 612). A similar analysis has been conducted by Cardoso and Morin (2018). They address the question “can economic

pressures overcome social norms?" by investigating the potential channels that drive female labor force participation to rise in response to unbalanced sex ratios in the presence of strong social norms against female employment in Portugal. They mention that the timing of this early rise in FLFP coincided with a significant military draft to fight the war in the African colonies against pro-independence movements and massive emigration of males. They exploit exogenous variation in sex ratio driven by war casualties and male emigration to investigate the implication of gender imbalances on female labor market outcomes.

Within this research, gender ratio can be thought of as an endogenous variable that cannot be ignored as it leads to inconsistency of the estimated parameters; therefore, understanding the valid causal identification is a crucial matter. Since we argue that gender ratio drives women's LFP, however, there is a potential issue of reverse causality in the sense that as we observe that LFP increases in one period, in the next period, there are chances of variation in gender ratio in the next generation probably because of women relocating themselves in the wake of availing labor market opportunity. There is another possibility of omitted variable bias. Although there are all possible demographic, household, and individual level explanatory variables in the model, the data for these variables are taken from the Pakistan Social and Living Standard Measurement (PSLM) survey. Still, some omitted district-specific and socio-economic conditions could influence both the gender ratio and outcome variable FLFP. So, given these two possibilities of endogeneity, we would instrument gender ratio with something which is entirely exogenous, i.e., exposure to extreme weather events, which is why this thesis will utilize this district-level exogenous variation in gender ratio to explain women's LFP in the presence of strict social norms in Pakistan.

To investigate the second hypothesis, i.e., the causal effect of women's LFP on EGRB, this research utilizes a nationwide cash transfer program called Benazir Income Support Program (BISP). It is the largest unconditional cash transfer (UCT) program of the government of Pakistan, having ever-married women in the household as the target beneficiaries. A poverty score is used as a threshold for the households to be eligible for the BISP cash transfer. This poverty score is constructed from 23 poverty covariates by using Proxy Means Test (PMT) method. Households with a poverty score of less than equal to 16.17 are considered eligible and thus part of the treatment group. While households above 16.17 are deemed ineligible, and therefore they are the control group.

However, there are few exceptions to this placement of households into treatment and control groups. Households have the option to appeal and become program beneficiaries if the poverty score is between 16.17 and 21.17 conditioned on the presence of at least one disabled person, at least one senior citizen (65 years of age or older) and less than three household members, and households with four or more than four children under 12 years (Ambler and Brauw, 2019).

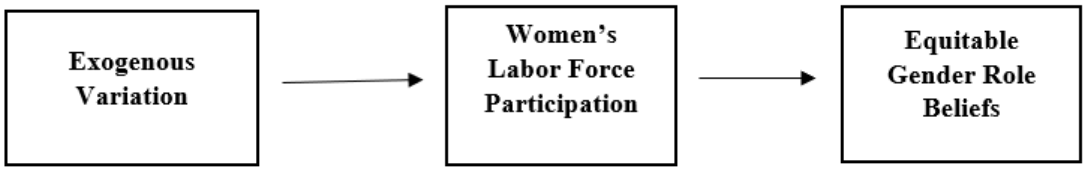


Figure 3.3 Causal Impact of women's LFP on equitable gender role beliefs

Even if these exceptions are present, the household cannot manipulate their poverty score to receive cash. Therefore, it is essential to discuss how poverty scores are applied in Pakistan; it is difficult for households to influence their scores. Since the poverty score is constructed with 23 poverty indicators that are impossible to hide at the time of the survey, households are neither aware of how the poverty score was calculated nor the eligibility cutoff. Therefore, there is a solid reason to believe that the treatment status is exogenous because it is determined based on poverty score calculated from covariates which are hard to manipulate by the households. Moreover, it is noticed that BISP has not updated the PMT since 2011, which raises concerns about the identification of program beneficiaries in subsequent years. Therefore, this thesis also attempts to replicate the PMT model used by BISP. So, this research would exploit the exogenously determined BISP status and replicated poverty score and their interaction to instrument endogenous women's LFP. The exogenous variation in women's LFP will help explain the underlying causal mechanism of how women's LFP may contribute to more equitable gender norms in Pakistan by using instrument variable (IV) regression.

The research also attempts to validate the findings from IV by looking at how the BISP cash transfer program has affected women's LFP by using regression discontinuity design (RDD). Under RDD, households just above the poverty score threshold of 16.17 are a comparison group for households just below the threshold. Therefore, a comparison of BISP beneficiaries and non-beneficiaries would help in the causal impact

of cash transfer on FLFP and verify whether there is a discontinuity in women's LFP around the threshold. As the BISP program is not random, discontinuity is not sharp, which means some ineligible households are in the treatment group, and some eligible households are in the control group. Therefore, a fuzzy RDD is utilized.

The BISP data is particularly suitable to test the causal direction from FLFP to equitable gender norms as this program targets women as its ultimate beneficiaries. This data will help investigate modernization theory – access to cash by being beneficiary of this program can help increase women's LFP – and gender stratification theory – that participation in the labor market makes gender role beliefs more equitable. Modernization theory says that if income increases, people prefer to work because they are modern, and being modern means abandoning part of the traditional gender norms. For that to be true, the research seeks to identify those women who would not have worked without this program and observes their gender norms which means we are identifying those women who change their LFP status because of extra cash they receive. So, if the modernization theory is true, their gender norms would change.

On the other hand, if modernization theory does not seem to hold, more income in the household would reduce women's LFP, and gender norms remain unchanged or become more traditional. Therefore, this thesis looks at women's emancipation and their husbands' support for her to be part of the labor force. If there is no support, it will be hard for a woman to achieve empowerment. Since Pakistani households tend to have a single utility function, deciding on women's LFP is a family strategy. So, it would be entirely misleading to ignore this critical aspect. These arguments based on modernization theory and gender stratification theory will be tested in this research by using cash transfer program data. Figure 1.2 presents a theoretical model derived from these theories.

1.4 Research Gap

Existing research provides several factors that determine women's LFP in Pakistan. Different researchers have acknowledged family system, husband education, family size, area of residence, literacy status, and socio-economic norms as the elements that influence woman labor force participation (Chaudhary and Nosheen, 2009; Faridi, Chaudhry & Anwar, 2009). Among various studies, Ejaz (2011) suggests a negative

relationship between the gender wage gap, fertility, and FLFP. It is observed that the education of both partners positively influences the likelihood of female labor force participation. Furthermore, to analyze the earning occupations of married women, Ahmad and Hafeez (2007) focused on parental education in their study conducted in 2002 at Mandi-Bahaudin district of Punjab, Pakistan.

In contrast, Faridi et al. (2009) argue an insignificant relationship between parents' education and female labor market participation. Nevertheless, a positive link between education and FLFP is proved in studies based on the data from 164 women respondents aged 15- 64 years. The study was conducted in 2007-2008 in district Bahawalpur in south Punjab. Faridi et al. (2009) also support the idea that FLFP is positively affected by partner's education, i.e., higher education of husbands increases wives' employment prospects provided they are well educated.

Regarding culture as a determinant of women's LFP, Shah (1986) examined that wearing a veil negatively impacts women's participation in the job market. Moreover, (Kazi et al., 1988) measured changes in the pattern and level of women's employment based on the agricultural census (1972, 1980) and labor force survey (1987 – 88). They argue that females face cultural limitations and domestic responsibilities, and they have a lower level of skills and education which creates a disadvantaged position for them. For this reason, women are given low-paid jobs, unskilled, and casual occupations in the informal economic sector. Ibraz, Fatima & Aziz's (1993) findings indicate that females' contribution is insignificant in rural areas. Although women provide support in taking care of cattle and are involved in agriculture, which contributes to the village economy, prevailing cultural philosophies challenge their productive role. Another study by Ejaz (2011) highlights the possible issues caused by endogeneity. The endogenous variables are fertility, possession of home appliances, and joint family system that is explained by 1) the average number of home-based equipment possessed in the vicinity, 2) the sex of firstborn child, 3) identical or different gender of first two children, 4) presence of a clinic nearby, 5) use of birth control device and 6) type of accommodation.

Several studies have investigated the impact of BISP cash transfer on adult labor supply, child labor, and women's empowerment in Pakistan. These studies include Ambler and Brauw (2019), who, in their research, found no effect of BISP cash transfers on household labor supply in Pakistan. They also found inadequate evidence women would

participate in the labor market in response to being cash recipients. However, in another study by Ambler and Brauw (2016), they used fuzzy RDD methods to identify the impacts and found that BISP transfers positively affect women's decision-making power and empowerment. Studies by Nayab & Farooq (2014) and Zuneira, Usman & Farooq (2018) have measured BISP's impact on poverty, and the analysis is cross-sectional.

Iqbal, Padda & Farooq (2019) recently provided evidence of improvement in women's political empowerment and freedom of mobility over time due to BISP cash transfer. Moreover, results from Iqbal, Farooq & Padda (2020) reveal that after two years of BISP intervention, there was no significant impact on women's empowerment; however, after 5 and 8 years, there was a substantial impact, especially on women's mobility, their involvement in decision-making, and voting behavior. These studies have analyzed the effects of cash transfer on female labor supply and different indicators of women empowerment by using RDD and the difference in difference approach. A study by Jalal (2017) is unique in that it evaluates the targeting performance of this program and found that there are significant exclusion and inclusion errors. The difference in different results of the study indicates that BISP improves women empowerment indicators. It is noticed that all these studies have used women's self-reported responses on empowerment indicators from survey questionnaires. However, using women's self-reported responses to their empowerment and decision-making in a household is criticized in recent studies due to social desirability bias such as Peterman, Schwab, Roy, Hidrobo, and Gilligan (2020) and Glennerster, Walsh, and Diaz-Martin (2018).

A brief review of these studies implies that several significant factors explain women's LFP in the context of Pakistan. In prior empirical research, if causality was studied, the effort was put in proving the causality in one given direction confirmed through observational or experimental data. This research intends to fill the existing research gap in several ways by providing a comprehensive study of one specific society: Pakistan, where favorable conditions and data availability permit verifying the validity of bidirectional causality at the same time. First, it exploits the fact that Pakistan experiences the negative consequences of climate change heavily. An instrument is constructed using geographical data of high granularity on extreme weather events for several decades. This instrument will be used to address the causal effect of social norms on LFP by using a large-scale representative cross-sectional household survey: Pakistan Social and Living Standard Measurement Survey.

Secondly, this research exploits a quasi-natural experiment from an income support program implemented in Pakistan in the same period to provide an empirical evaluation of the causal link that higher women's LFP causes men gender role beliefs to be more equitable by using panel data from a large-scale survey of Benazir Income Support Program, BISP. The program uses poverty score cutoff to identify the beneficiary and non-beneficiary households. As discussed before, BISP has failed to update its poverty score since 2011; therefore, this thesis attempts to replicate the BISP's PMT model using baseline 2011 BISP data. Therefore, this thesis is a first attempt to use BISP treatment status as one of the instruments to identify the causal impact of women's LFP on gender role beliefs using the instrumental variable approach. Prior research on BISP data has mainly investigated the causal effect of this cash transfer using RDD and the difference and difference approach. However, the idea presented in this thesis is to identify how women may change their LF participation decision because of their beneficiary status. So, an exogenously determined treatment status of households to this program and the replicated poverty score along with their interaction are used as instruments to address the causal effect of women's LFP on equitable gender role beliefs.

Thirdly, the thesis also addresses the existing empirical gap concerning the development of women empowerment indicators. As discussed earlier, prior research has utilized women's self-reported responses on their empowerment, which is criticized due to social desirability bias. Studies on BISP data have used women's responses on various empowerment dimensions through survey questionnaires, i.e., their decision-making in the household, access to money, freedom of mobility, voting behavior, and gender role beliefs. The present research considers the issue of self-reported women's responses and constructs a unique specification concerning gender role beliefs in the household. As discussed in the previous section, working in the labor market in Pakistan is a family strategy; therefore, women's emancipation does not matter for her to decide on participation in the labor market. Thus, this thesis constructs a specification that captures the coincidence of the husband's and wife's gender norms, i.e., both husband and wife share traditional or non-traditional gender norms. This specification will identify that if a woman has non-traditional gender norms whereas her husband is more traditional, then there are prospects that if the wife receives some extra cash, it will influence her husband's gender norms to be more equitable.

Similarly, if the wife is of more traditional views and her husband is not, then again, there are possibilities that the wife would change her gender norms opinions in response to being a cash recipient. Whereas, if both husband and wife share the gender norms, i.e., they have either traditional or non-traditional norms, it can be anticipated that it will be hard to change their norms even if a wife is a program beneficiary. So, this specification will provide a more holistic approach to capture the effect of cash transfer on women's empowerment. Consequently, it will help in explaining women's participation in the labor market in response to being program beneficiaries and whether women's LFP causes a coincidence of husband's and wife's gender norms or not.

Lastly, the research employs multiple correspondence analysis (MCA) as a data reduction technique to develop an index of husband's and wife's gender norms coincidence. This research is a first attempt to use MCA to create an index from several binary gender norms questions from survey data in the case of Pakistan. The study also applies MCA on given women empowerment domains from the BISP data, namely, women's decision making in household, women's freedom of mobility, voting behavior, and women's gender norms, to compare how the results from these original specifications differ from the new specification of husband's and wife' gender norms coincidence.

The datasets used in the existing studies are either cross-sections limited to certain districts in Pakistan or collected from potentially non-random surveys conducted by authors from one or a few districts. This research exploits previously underutilized individual-level data for Pakistan Social and Living Standard Measurement Survey (PSLM) from the early 2000s and 2010s. These are high-quality individual-level data, representative of the population, and rich in personal details. This extensive data allows for in-depth individual-level analysis to figure out what factors explain FLFP in Pakistan. Lastly, however, Pakistan's Labor Force Survey dataset provides information on migration from one administrative district to another, identifying several possible reasons for migration. But migration data due to climatic events are not available. Therefore, this research utilizes district-specific variation in gender ratio in the district affected by extreme weather events to determine the gender dimension of climate-induced migration

1.5 Relevance of the Thesis

This research aims to answer questions that are of vital importance for policymakers from all over the world. First, the issue of women's LFP is one of the significant ones across the globe. As Klasen (2019) observed, although women's LFP is rising globally, it is still below men's. Second, climate change is a phenomenon affecting the world likewise irrespective of what stage a country stands on a development ladder. Third, social norms have been proved to be persistent, and from a historical point of view, they have remained unchanged over a more extended period (Alesina, Giuliano, and Nunn 2013). Therefore, the questions addressed in this research are equally crucial for developed and developing countries as the problem of low women LFP, extreme climatic events, and stagnant social norms are worldwide issues.

In this research, we are focusing on Pakistan as a case study. Pakistan provides a particularly suitable case to study this relationship because all three of the above issues compound in this society and economy. In addition to being a low-income country, undergoing the process of modernization, and facing particularly harsh consequences of climate change, Pakistan is also a culturally and socially patriarchal society with forcefully identified gender norms. This makes Pakistan relevant from a research perspective and suitable as a playground for testing various methods of causal identification and quantifying the strength of multiple operating influence channels. There is also a significant research gap in that, in general, countries from South and Southeast Asia are often white spots in terms of causally identified studies of social norms and women's LFP. This research attempts to verify the causal links using different econometric methodologies, i.e., instrumental variable and regression discontinuity.

1.6 Structure of the Thesis

Following the introduction, the rest of the thesis is divided into four chapters. In Chapter 2, detailed literature is reviewed on the relationship between gender role beliefs and FLFP, gender ratio and FLFP, how cash transfer programs affect FLFP, and how climate change plays its role in affecting internal migration. The two studies of this thesis employ the econometric methodology of causal identification, i.e., Instrumental Variables estimated through Two Stage Least Square (TSLS) and Regression

Discontinuity Design. The details on these methodologies are discussed in Chapter 3. Chapter 4 provides information on sources of data, operationalization of key variables, and their descriptive statistics. Chapter 5 presents the interpretation and analysis of the results derived from the econometric methodologies used in the two studies. Lastly, chapter 6 concludes the research with recommendations for policymakers, limitations of the study, suggestions for future research, and thesis contribution.

CHAPTER 2 REVIEW OF LITERATURE

The literature review is divided into sections in line with the research objectives and research questions presented in chapter 1. To begin with, in first section 2.1, the research on the relationship between social norms and women's labor force participation is discussed in detail. In addition, section 2.2 focuses on studies that explicitly address how variations in gender ratio influence LFP of women. Section 2.3 explicitly explains existing literature on how climatic events affect migration patterns. Section 2.4 addresses studies on the link between cash transfer programs and women's LFP. The literature specifically focused on women's LFP in Pakistan is presented in section 2.5, and lastly, the summary of the chapter is discussed in section 2.6.

Before moving on to the literature review, it is first essential to operationalize for this thesis the understanding of the term: social norms. We follow the approach that social norms are rules of action shared by people in each society or group; they define what is considered normal and acceptable behavior for the members of that group (Cislaghi and Heise 2018a). In literature, norms are also thought to be different from personal attitudes, and norms and personal attitudes often oppose each other (Fishbein and Ajzen 1975, as cited in Cislaghi and Heise, 2020). Personal attitudes are beliefs and perceptions of a person about a specific action; however, norms are views of what is acceptable in society. On the other hand, the notion of gender norms is primarily discussed in the literature with specific reference to the nature of gender. The idea used mainly by feminist sociologists is conceptualized as a social system that allocates roles, resources, and power according to whether a practice is considered male/masculine or female/feminine (Ridgeway and Correll, 2004). Within the context of this research, social norms related to women's labor force participation are discussed; therefore, the thesis is specifically interested in analyzing personal gender role beliefs. Thus, the terminologies of 'social norms' and 'gender role beliefs' are used interchangeably.

2.1 Social Norms and Women's Labor Force Participation

Several studies investigated the link between social norms and women's LFP. The majority of the work involves western societies such as Stam, Verbakel, & de Graaf. (2013) for the Netherlands, Stähli, Le Goff, Levy, & Widmer (2009) for Switzerland

and Kanji (2011) for Great Britain. However, research for non-western countries is scarce, such as Contreras and Plaza (2010) for Chile. These investigations involve a cross-sectional research design and investigate the impact of attitudes on employment decisions regarding specific gender roles and found that women in traditional gender roles are less likely to be employed than those with a less conventional role. Regarding panel data, studies use country-fixed effects to control for time-invariant variations across countries. For instance, Fortin (2005) used three waves of the World Value Survey (WVS) to examine the link amongst women's employment rates and attitudes in 25 OECD nations. She used the probit model to study the impact of work values and gender role beliefs on occupation status with the sample of 25 – 54-year-old females. The findings propose low employment possibilities for females with traditional gender role attitudes. Nordenmark (2004) used a similar approach, which considered data of ten countries from the 1994 International Social Survey Program (ISSP). A relatable conclusion has been drawn from the study that used regression analysis to measure gender ideology by including individual gender attitudes with country-fixed effects.

Algan and Cahuc (2005) used a different strategy by using macro indicators to analyze determinants of women's LFP at the country level. The relationship between the female employment rate and aggregate attitudes is assessed using three waves of WVS at the country level. The findings suggest a relationship through regression analysis while controlling family institutions and the labor market, e.g., public spending share in childcare and employment protection legislation for 19 OECD countries. As expected, family culture considerably influences women's LFP. Giavazzi, Schiantarelli & Serafinelli (2009) also used a similar strategy and instrumented individual and aggregate gender role attitudes with average religiosity and found a negative impact of traditional gender role attitudes on women's LFP.

There is also evidence of theoretical and empirical efforts made by researchers. For example, Fernandez (2013) developed a theoretical model that jointly determines culture and female LFP. Empirical studies such as Fortin (2009) investigated the relationship between women's gender role attitudes and work decisions. The relationship between men's views on gender role questions and women's LFP has also been studied by Charles & Bradely (2009). Studies have identified possible driving forces affecting LFP such as the AIDS crisis (Fortin 2009), innovations in contraception (Goldin and Katz 2000), individual's cultural background (Fernandez and Fogli 2009), and early

childhood experience (Fernandez, Fogli, & Olivetti, 2004), and traditional agricultural practices (Alesina et al., 2013).

On the studies identifying the relationship between the factors affecting gender identity and women's LFP decision, the study results by Fernandez et al. (2004) explain that men's gender role preferences developed from their early childhood experiences can significantly affect women's participation in the labor market. In their seminal paper, Fernandez et al. substantiate intergenerational connections amongst the working behavior of daughter-in-law and mother-in-law. Study results propose that working mothers who have nurtured men have a greater possibility of allowing their wives to work. Moreover, to explain causality, they use the variation in mobilization rate in the U.S. during WWII. To support the statistical linkage discovered by Fernandez et al. (2004), Morrill and Morrill (2013) found a causal relationship between the working behavior of mother-in-law and daughter-in-law may result from "assortative mating. Morrill and Morrill (2013) investigated intergeneration preference formation from mothers to daughters, while Fernadez et al. (2004) focused on the intergeneration link between mothers and sons.

Fernandez et al. (2004) motivated other researchers to test the "mothers and sons" story using different countries' datasets. Kawaguchi and Miyazaki (2009) used the Japanese General Social Surveys (JGSS) data to replicate Fernandez et al. (2004) and found statistically insignificant results. Additionally, they examined the link between the gender role attitude of sons and their mother's work status. They discovered that men raised by full-time employed women have a higher possibility of preferring working wives. This study adds to Fernandez et al. (2004) as it directly analyzed the phenomenon. Another study by Butikofer (2013), using the Swiss Household Panel 2005, replicated Fernandez et al. (2004)'s cross-section analysis, and the results are statistically significant. Antecol (2003) found that if the male partner displayed a more liberal attitude towards gender equality, the female partner is more likely to participate in the labor force.

Social norms have also been considered a significant factor in influencing women's marriage and labor market participation (Bertrand et al., 2015). Both men and women face restrictions and societal pressures to adopt a specific role. Society prescribes that women have a responsibility to take care of the house while men should assume the role

of breadwinner for the family. Another doctrine suggests “a man should earn more than his wife.” According to World Value Survey, 16.8% of Brazilian respondents agree that ‘when jobs are limited, men should be given preference over women.’ While 33.5% of Brazilians accept that “if a woman earns more than her partner it will result in several problems” (Codazzi, Pero and Anna, 2018). Bertrand et al. (2015) suggested that social norms influence gender identity, which is apparent from the widely accepted belief that “a woman should not earn more than her husband.” Based on this notion and given the economic identity theory by Akerlof and Kranton (2000), Bertrand et al. (2015) suggest an inventive approach by using data on relative income from U.S. households.

Moreover, this study has been further expanded by considering its impact on a married couple. The authors developed a hypothesis by predicting that the wife’s labor supply declines when the wife’s earnings are more than the husband's. According to Bertrand et al. (2015), when married women earn higher than male partners, they are less likely to participate in intensive and extensive margins in the labor market. The identity framework can clarify these findings despite being counter-intuitive according to the standard utility model. In addition, the chances of divorce increase when the husband is earning less than the wife and the household production of the wife grows to compensate.

After having the discussion made above, it can be noted that studies cannot confirm the direction of causality either one way or two ways, i.e., from gender role beliefs to the women’s LFP or vice versa. A few researchers attempt to solve this problem using longitudinal data (Himmelweit and Sigala, 2004; Cunningham, 2008). These studies evaluate gender role attitudes over time and identify any variations with the change of labor market status by using multidimensional data. These studies find a recursive relationship between employment status and gender role attitude, which means that variation in attitudes results in a change in employment status and vice versa. At large, gender role attitudes are variable over time and seem to be affected by employment status (Cunningham 2008).

Research that deals with the cultural predictors of women's participation in the labor market have problems with the research designs. The cross-sectional nature of the method implies that women's participation can also impact cultural factors. For instance, more participation by women can lead to a change in the social norms of conservative

nature and vice versa. In other words, methodological flaws prevent the researchers from making causal inferences. (Seguino, 2007) explored whether the participation of women in the labor market can alter societal norms and found that it influences social norms but with a time lag. Drawing on an epidemiological approach, Guner and Uysals (2014) studied the interplay between domestic migration and female migrant labor supply. They used the female employment rate of migrants' origin province as a proxy for the cultural values of migrated women. They found that the women's employment rate in the area of origin when migrant women were born can positively influence women's labor supply behavior. This previous generation's attitude towards women's employment is used to address the issue of endogeneity.

Social conservatism as a constraint to women's labor participation has also been conceptualized in a study by Dildar (2015) whereby religiosity and patriarchal norms were found to negatively impact the extent of women's involvement in the labor supply after controlling for endogeneity. The author points out that the women's participation in the labor market may lead to changed attitudes, and reverse causality is also possible, i.e., changed attitudes leading to more women's LFP. The issue of endogeneity is dealt with by using an instrumental variable derived from the long-run effects of family structure on pre-adult socialization. Therefore, after using family conservatism as an instrument for patriarchal norms, the study finds the negative impacts on female labor force participation.

2.2 Gender Ratio and Women's Labor Force Participation

The sociological literature produced during the 1960s to 1970s was based on cross-sectional data collected from various countries. It was characterized by mixed evidence of the correlation between population sex ratio and women's labor force participation. Ward and Pampel (1985a, b) examined the dataset from sixteen developed countries and concluded that during 1955 and 1975, the population sex ratio mainly account for the women labor force participation. In another study, Pampel and Tanaka (1986) studied from 1965 and 1970 from seventy developed countries and showed that the higher the population of women, the higher the women labor force participation. Two years later, in a study contradicting the prior evidence, South and Trent (1988) inferred from their data from 58 countries from the developed world that there is no correlation between population sex ratio and women labor force participation.

As per Guttentag and Secord (1983), sex ratios affect outcomes of imbalances existing in the marriage market and supply of labor. In the Canadian context, it was observed that the marriage market condition does not affect the female labor supply. Age also does not affect the rise in women's participation in the workforce (Galarneau, 1994). In comparison with the context of the USA, Fukuda (2006) found that the correlation between cohort effects and sex ratio difference for female LFP is relatively weak in the context of Japan. It is also indicated by Angrist (2002), based on a study conducted in the context of the second generation of immigrants in America, that the lower sex ratio can account for more participation of women immigrants in the labor supply. A lower sex ratio implies fewer men in comparison with women. Evidence for the increase of women in the workforce in the aftermath of WWII was presented by Acemoglu, Autor, & Lyle (2004), who utilized the regional variation in the military mobilization of the US army for the war for this purpose. Their study found that states where the mobilization rate is higher, have higher post-war FLFP.

Whereas the post-war sex ratio decrease was temporary in the US, unlike Germany, where the sex ratio decline was permanent, it had a long-lasting effect on FLFP. It was found in the study by Fernández et al. (2004) that women who became part of the labor market because of missing men, their sons were found to have married with working women afterward. This led sociology researchers to conclude that experience of being sons of working mothers results in more acceptance for wives who would be employed. Another example of an unbalanced sex ratio comes from the trans-Atlantic slave trade in Africa by Teso (2014). This study reveals that the slave trade left the African region with a more female-biased sex ratio as males were exported in the slave trade. Teso argues that the legacy of the slave trade in Africa has had a lasting impact on the gender roles as women, who belong to ethnic groups which were severely affected by the slave trade, are more likely to be in the labor force. In line with this result, a recent study by Grosjean and Khattar (2019) found that areas with historically more male-biased population sex ratios are more conservative regarding the female labor supply in Australia.

Given a vast literature on sex ratio and FLFP, the problems of omitted variables and reverse causality prevent researchers from establishing causality related to sex ratio. The idea is that a person can move domestically or internationally due to various reasons related to marriage or other labor market opportunities. These reasons may correlate

with the outcome of interest. Therefore, researchers can turn to variation in sex ratios beyond the control of an individual to establish causal identification. Like Angrist (2002), another researcher Abramitzky, Delavande, & Vasconcelos (2011), utilized the variation in sex ratio in France due to WWI to assess the mating patterns. Moreover, Brainerd (2017) uses both sex ratio and the outcome indicators measured in the census of 1959 after the war, showed that the low level of sex ratio in Russian regions and among various cohorts reflected in the drop of marriages and rise in the birth of children outside marriage. In the German context, the variation in sex ratios in the aftermath of war enabled Kesternich, Siflinger, Smith & Steckenleiter (2020) to assess its long-term impact on fertility and marriage patterns.

Grant, Kesternich, Steckenleiter, & Winter (2018) exploited random variation in sex ratio in Germany after WWII to examine the lasting effects of sex ratio imbalance on the extent to which women participate in the labor market. Another outcome of interest has been the representation of women in the German political landscape. The sex ratio of 1946 in Germany impacted the number of nominations that female politicians could secure in the election of 1990, which shows that the historical imbalance of sex ratio continues to reflect in the present scenario of women's LFP. This phenomenon could be explained by the absence of men from the labor market owing to war which opened space for women in the labor market. And it brought the realization that occupations that were thought to be not meant for women underwent a change in perception regarding their suitability for women.

On the other hand, various other frameworks such as Ward and Pampel (1985a, b) could also explain the impact of sex ratio and cohort effects. Ward and Pampel (1985a, b) argued that instead of supply-side factors such as opportunities for marriage and opportunities in the labor market, competition matters. The absence of men leads employers to recruit more women. The reason behind why women's LFP differs across cohorts could be found in the relative income hypothesis proposed by Easterlin (1980). The baby boomer generation entered the labor market in the postwar scenario when wages were low compared to their consumption levels.

Similarly, Canadian women, LFP was also found to have been affected by marriage market condition during 1971 and 2001 (Emery and Ferrer, 2009), which is indicative of the fact that sex ratio effects are in line with the predictions of the theory, yet they are

not statistically significant. Another study by Cardoso and Morin (2018) also emphasizes the demand-side and supply-side effects of an unbalanced sex ratio in post-war Portugal. They are also of the view that in the absence of men, women entered male-dominated occupations as demand for them increased by firms, and hence the gender pay gap declined. They also highlighted that impact of the sex ratio on the marriage market points to the 'muted' supply channel.

The relation between emigration and female employment patterns has attracted many scholars focusing on the impact of male migration on women's LFP (Binzel and Assaas,2011; Desai and Banerji, 2008; Lokshin and Glinskaya, 2009). The literature shows that labor market participation can be influenced by international migration through two channels, i.e., migrant remittances and reallocation of migration-induced labor (Amuedo- Dorantes and Pozo, 2006). These channels directly impact households and have indirect relations mediated through other variables such as productive investment (Binzel and Assad, 2011) rural-urban mobility (de Haas. 2008).

The migration of a household member affects the reallocation of either labor among those left behind or in the income itself. This result of the migration decision shows a household strategy of utility maximization conceptualized by Stark (1991) in New Economics of Labor Migration. In this regard, Sorensen (2004) presented an argument that in rural Morocco, harvesting, which was used to be considered as a typical task of males during the 1950s, has been feminized because of the emigration of males. Eventually, it became a women's task as young men have shunned the harvesting. Furthermore, it is stated that international migration has increased land and livestock ownership, but it has also increased female participation in unpaid work (Steinmann, 1993). In simple words, women must compensate for the lost labor due to emigration while continuing their domestic chores.

The effect, as mentioned earlier, is likely to occur where only men are already playing their role in the agriculture sector. In the nuclear household, this increase in women's unpaid or extra workload has not been found in Todgha valley in Morocco because women have the charge to take care of their agricultural produces even before their parents' migration. If the migration happened, they would ask or seek help from other men, or they would like to hire laborers for the agricultural task (de Hass and Van Rooj, 2010). It reveals that migration-induced reallocation of labor is limited in societies with

strict gender roles. In the same vein, Morocco's response to labor market migration is somewhat like Egypt, where the labor supply of females is indifferent to changes in male labor supply (Binzel and Assaad, 2011). They made an argument that if the changes in FLFP occur, it will affect unpaid domestic work because these are subjected to less restrictive social norms and can be combined with domestic chores and childcare duties in rural areas, which is why male migration is likely to affect women's labor supply in rural areas.

2.3 Link between Migration and Climate Change

The previous section has discussed various factors of human migration. This section will present studies that specifically discuss climate-induced migration. Scientific evidence has revealed that changes in the environment have affected human settlements and their mobility since the Holocene (Richerson, Boyd, & Bettinger 2001; Morgan, 2009). The effect on resource productivity because of environmental changes forced people to migrate and settle in clusters to avoid exposure risk (Halstead, 1989 and Morgan, 2009). Currently, human migration and dislocations continue to happen because of climate change. These climate changes at times may be slow-onset or abrupt. Biodiversity, soil degradation, changes in water flow, variation in agriculture production are considered slow-onset environmental changes (Parry, Parry, Canziani, Plautikof, Van der Linden, Hanson, 2007). Abrupt or sudden changes in the environment consist of extreme rainfall, an outburst of a glacial lake, flash flood, drought, heat waves, etc. Subjected to these changes, people either migrate voluntarily or by compulsion (Kalin, 2010).

Climate change is linked with socio-economic, political, and ecological factors that induce human migration ((Black, Adger, Arnell, Dercon, Geddes, & Thomas, 2011; McLeman and Smit, 2006). Other significant causes of migration include inequality, poverty, and institution failure (Barnett and Adger, 2007). According to scientific research, migration is a mitigation strategy to respond to climate change (Massy, Axinn, & Ghimire, 2010; McLeman and Hunter, 2010). People move from high-risk areas to seek livelihood opportunities and reduce risk (Tacoli, 2009). Developing countries such as Bangladesh, Haiti, Male, and Eritrea have included migration and displacement as an adaptive mechanism in response to climate change in their national adaptation programs of action (Banerjee, Gerlitz, & Hoermann, 2011). But according to Warner (2010), migration occurs because of the failure to manage climate change. In such cases, people

are displaced from rural to urban areas (transformative adaptation) instead of situ adaptation (staying in the same vicinity).

There is a shortage of quantitative research on migration concerning climate adaptation because of the uncertain and complex patterns of climate-induced migration and the variation of people's individual and collective responses towards climate change. Furthermore, the phenomenon of climate change itself is so uncertain to precisely predict its impact on human settlements (Black et al., 2011; Banerjee et al., 2011). The nature of prediction about the above-stated phenomena includes questions like, how many people will migrate—how they will migrate-what impact it will create in the new settlement. It is even more challenging when there is nonlinear relation between migration and climate change and depends on various factors (McLeman and Smit,2006).

According to the Asian Development Bank (2012), climate-induced migration is one of the causes of urbanization faced by developing countries today. These nations spend most of their developmental budget in urban areas, which may be a pull factor coupled with the push factor of climate change to persuade people to migrate within the country. As a result, urban areas act as a magnet to attract rural migrants (Hussain, 2014). According to Adamo (2010), climate change is an essential factor for increasing migration flow to cities either temporarily or permanently, stressing policymakers to take measures to compensate for the migrant influx. A few scientific studies put efforts into developing relations between interlinked factors. They maintained that the productivity of agriculture sectors could be affected when variation in climate variables occurs. These variations in climate variables are temperature and average precipitation. For instance, 5-25 percent of wheat production and staple food could be affected by climate change (Sulatna et al., 2009). Majid and Zahir's (2014) study had shown the same results that climate change affects farm productivity and socioeconomic vulnerabilities in the province of Sindh and Punjab in Pakistan. According to their research, drought is a vital climate change variable that badly affects the production of sugarcane, cotton, rice, and wheat. This decline in crop production negatively impacts the tenant farmers and sometimes forces them to search for alternative livelihoods or even compel them to migrate to urban settings. Mueller et al., (2014) seminal work have found the same results as they described that heat stresses are another cause of internal migration in Pakistan.

Farooq et al. (2005), in their study from Faisalabad, Pakistan, noticed that 13% and 80% of the respondents migrated due to poor educational and economic opportunities, respectively. Other studies presented a similar finding that in Pakistan, people migrate mainly because of differences in labor productivity between rural and urban areas, technological advancement, and urban investment in cities (Khan, Shehnaz, & Ahmed, 2000; Mahmud, Musaddiq, Said, & Sabir, 2010). Several studies have explored the concept of internal migration independently within Pakistan. Rural to urban migration consist of 40 percent of total internal migration (Arif, 2005). Ranking the rural-urban migration, Sindh and Baluchistan have the lowest migration, followed by KPK and Punjab (Mahmud et al., 2010). The majority of these migrants move to already densely populated districts. For instance, 56 percent tend to move to the provinces or the capital city of Islamabad because of the social networks or higher developmental programs in stated regions.

The problem of rural-urban migration has been addressed in the study of Mueller et al. (2014). Their study was based on the longitudinal survey conducted from 1991-2012 in the rural setting of Pakistan. This study has explored the linkage between weather and long-term migration. Their study analyzed critical weather variables at the village level, such as rainfall during the monsoon season and average temperature during winter. They found that extremely high temperatures during winter have the strongest correlation with migration concerning other variables.

Additionally, high temperature destroys one-third of agriculture production. This negative shock to agricultural output significantly impacts households' subsistence, and as a result, their mobility is mainly directed towards other cities. Mueller et al. (2014) have argued that the poor have more flexibility in migration because they do not have land or other assets that compel them to stay at home. Hence, it has been summed up that the benefits of migration for the poor outweigh the mobility cost during harsh and extreme climate conditions, which induces all kinds of migration.

2.4 Cash Transfer Programs and Women Labor Force Participation

The issue of different hierarchies of gender exists in society, which translates into inequality. It is further aggravated by the way gender is socially constructed, which devalues women and reinforces gender roles in the labor market. The ideological conception of gender justifies the injustice in the distribution of resources and power in

society. Seguino (2007) believes that gender stereotypes provide how society differentiates between men and women and then assign roles for them in the labor market. Moreover, the author mentions that the gender norms have set the appropriate behaviors to be displayed by both men and women. This factor further develops the division of labor between men and women. Challenging these norms has social and well-being consequences for those who defy these norms, for instance, social stigma and anxiety.

Seguino (2007) argues that being financially empowered will reduce the gap in terms of gender income inequality. Second, it encourages women to take on more challenging and rewarding occupations to weaken socially constructed gender roles. Burnell and Ratzel (2005) presented evidence that cultural norms influence women's sense of agency and work as a mediating factor in analyzing the effects of wages on bargaining power in India. Therefore, it is imperative for those working in behavioral economics to study the impact of financially empowering women on norms in society to help policy makers develop more effective policy interventions.

Concerning empowering women financially, direct cash transfer schemes have gained attraction among governments of developing countries, which are meant to provide financial support to households in the lower strata of society. There have been reservations regarding its unintended consequences, such as whether programs adversely affect the labor market participation on the part of the recipients. Theoretically speaking, based on the standard consumer model, the increase in the non-labor income results in the rise in the value of leisure, which leads to the reduction of the supply of labor in the market (Becker, 1965).

The framework mentioned above can be used to analyze how these cash-transfer schemes play their role in influencing the contribution to the labor market by the recipient individuals and households. Interestingly, there are several reasons why there would be an increase in labor market participation by recipients. It lies within the realm of possibilities that recipients could pursue riskier and better jobs after receiving cash transfers. This pursuit for difficult and better jobs could be traced to the recipient household bargaining power, which increases due to the cash transfer (Chiappori, 1992). Thus, it could be asserted that the impact of these cash transfers on household

contribution to the labor market is a mix of positive and negative (Baird, McKenzie, and Ozler 2018).

Based on data collected from one of the largest such cash transfer schemes, *'Progresa'* in Mexico, Parker and Todd (2017) found that the program had no impact on neither labor supply nor leisure. In Brazil, the program called *'Bolsa Familia'* differentially impacted the labor supply of the informal sector and formal sector, i.e., recipients were drawn more to the informal sector and less to the formal sector (De Brauw, Gilligan, Hoddinott, & Roy, 2015). In another study, Banerjee, Hanna, Kreindler, & Olken (2017) re-analyzed the results of randomized control trials of conditional cash transfer programs and found that cash transfer does not impact labor supply or hours worked.

However, concerning unconditional cash transfers, studies have found differences in recipients' and non-recipients labor supply. After receiving the cash transfers, workers in Zambia and Malawi devoted their time working on their farms rather than others (Ervin et al., 2017). It has also been found that households with women pensioners contribute less to the labor supply (Bertrand, Mullainathan, and Miller, 2003). Past research indicates that upon receiving a pension, women are set to make more income gains as males make less contribution to the labor market when they become pensioners, whereas women don't (Ambler, 2016). It is still unclear in the South African context what effect the presence of pensioners in the household has on the participation by the prime-age adult.

2.4.1 Benazir Income Support Program (BISP) in Pakistan

This sub-section extends the discussion of cash transfer programs and their impact on labor supply, specifically in Pakistan. BISP is a flagship program introduced by the government in 2008 as means for the social safety net for those at the bottom of the society's financial hierarchy. The program served as many as 3.4 million women with monthly cash transfers amounting to PKR 1,000, equivalent to twelve dollars (Gazdar, 2011). Although conceived as a program to be operated monthly, it served quarterly. Therefore, each recipient had twelve thousand rupees per year. During the initial two years, members of the National Assembly (MNA) had the mandate to find suitable and deserving recipients and nominate them for the scheme (Khan and Qutub, 2010). Each MNA was allocated the quota of 8000 recipients from among their constituency voters based on the demographic characteristics. Income more than 6000, foreign nationality,

ownership of land more than 3 acres, and the employment of family members in the public sector would make an individual disqualified from becoming a recipient. The particulars of the nominees were thoroughly cross-checked by National Database and Registration Authority (NADRA) (Haseeb and Vyborny 2016; Nayab and Farooq 2014).

Since so many households qualified for becoming recipients, MNAs exercised their discretion in nominating the members of the low-income families. And consequently, it was found by a study conducted by the Pakistan Institute of Development Economics (PIDE) that 16.1 percent of the recipients were not eligible for the BISP program. PIDE used the accurate Pakistan Panel Household Survey (PPHS) of 2010 to arrive at these results (Nayab and Farooq, 2014). Objections were raised, and concerns were shown regarding sustainability and transparency in Universal Cash Transfer (World Bank 2013). To address these concerns, the government of Pakistan collaborated with the World Bank to develop a poverty scorecard that was a benchmark for targeting potential beneficiaries (Hou, 2009).

The Proxy Means Test (PMT) based method was accepted to target beneficiaries in 2010, replacing parliamentarians-based targeting mechanism. The data of 27 million households were collected and named as National Socio-Economic Registry (NSER) in 2011. Through NSER, household-level data was collected on 23 poverty indicators identified by PMT. This poverty scorecard census was aimed at identifying the poorest twenty percent of households in Pakistan. The cut-off was 16.17 out of a total of 100. The individuals below this threshold are all qualified to apply for BISP, whereas those above 16.17 stand disqualified (Cheema et al., 2014). So far, 5.8 million families (ever-married women) are enrolled in the unconditional cash transfers, and the target is 7 million. The quarterly payment of this UCT was increased over time; it was Rs. 3000 in 2008. The amount was raised to PKR 3600 in July 2013, PKR 4500 in July 2014, PKR 4700 in July 2015, PKR 4834 in July 2016, PKR 5000 in July 2018, and PKR 6000 in January 2020 (Iqbal, Farooq, and Padda, 2020). The amount is distributed via the Pakistan Postal Service, BISP debit, smart card, or mobile money transfers. Gradually, the payment system was updated, and in 2016 biometric verification system (BVS) was introduced. With this new payment system, a beneficiary would personally receive the payment after biometric verification at specified ATMs and point of sales (POS). In a

phased manner, all the beneficiaries have been shifted to the BVS system during 2016–2019 periods in a phased manner (Iqbal et al. 2020).

At its core, BISP has two objectives. One of them is to provide cash relief and protect the poorest of the poor from inflationary setbacks in the short run, while in the long run, the objective is to empower them to invest in health and education. Both purposes will culminate in uplifting the households from poverty (Cheema et al. 2016). Irrespective of its merits and demerits, what is uniquely fascinating about the BISP is that it is the first initiative of its kind in which Pakistan is directly reaching out to women rather than males of the households. This step underscores the recognition on the part of the state regarding the shift in the way in which the state used to conceptualize the household as being ruled by the male. As per the program, if the same household has a widow, a wife, and a divorced sister, it will be eligible for cash transfers as heads of three families. Around eight percent of the total households have two recipients (Cheema et al., 2014).

Regarding the impact of BISP cash transfer on women's labor supply and their decision-making in the household, Ambler and Brauw (2019) found no overall effect of the BISP transfers on household labor supply in Pakistan. At the individual level, they found weak evidence that women would like more involvement in the labor force by increasing women considered unemployed. However, another study by Ambler and Brauw (2016) found that BISP transfers have substantial, positive impacts on variables measuring women's decision-making power and empowerment.

2.5 Determinants of Women's Labor Force Participation in Pakistan

There are quite a few studies on factors affecting women's LFP in Pakistan. The most prominent factors are the traditional factors such as age, marital status, education, etc. Shah (1986) has pointed out several socio-economic and demographic factors that determine the extent of women's involvement in the labor market of Pakistan. The study found a negative relationship between nuclear family type and child-women ratio and the degree to which the women are involved in the labor market. However, the study found a positive effect of dependency ratio, literacy rate, and marital status. In a similar vein, the study conducted in Karachi had similar findings. It highlighted that households with a male head figure demonstrated less likelihood of women participation in the labor market (Chishti, Lodhi, & Rashid, 1989).

Interestingly, the households with other women as family members demonstrated more likelihood of women's involvement in the labor market. In a field study focused on understanding demographic and socio-economic factors influencing the participation of educated-married women in the labor market, Hafeez and Ahmad (2002) applied Probit and Logit model. They collected data from Mandi-Bahauddin, Punjab, from 1998 to 1999. They asserted that the spouse's education level, per month household income, number of workers, and financial assets are all negatively related to the degree to which women participate in the labor market. However, parental education is found to impact FLFP positively. These authors found other factors that can affect the women's decision to participate in the labor market, such as net wealth, spouse's income, distance from city centers, and the number of dependents.

Contrary to Hafeez and Ahmad (2002), Faridi et al. (2009) revealed that parents' education is insignificant. They admitted positive relation between women's LFP and education by analyzing the data of 164 women participants aged (15-64) from the rural area of Bahawalpur during the year 2007-08. Though, Safana, Masood, Muhammad, & Amir (2012) raised a contradicting argument by stating that the education, primary and matric level, of the household head is negatively related to the employment decision of women. Nevertheless, educated women have more probability of participating in the labor market.

Culture plays a vital role in LFP, especially in Pakistan. Shah (1986) revealed that women's adherence to the veil restricts women's contribution to the labor market while analyzing the trends in women's LFP from 1951 to 1981. Kazi et al. (1988) noted the changing pattern of women's employment over the periods by utilizing the Agriculture Census (1972-80) and Labor force survey (1987-88). They stated that cultural restrictions, low level of education, and household responsibilities make women economically underprivileged, which is why they accept low-paid, unskilled, and casual jobs for their subsistence.

On the role of women in agriculture, Ibraz, Fatima, & Aziz (1993) stated that in the village, women's role is taken as insignificant. However, women have significant contributions, apart from agriculture, to keeping and raring livestock, which is vital for the village economy, but the dominant patriarchal norms weaken their productive role. His study was based on a 1989-90 survey conducted in Rajpur village in Punjab. He

looked at the household-related factors which are essential for female participation in economic activities. Naqvi et al. (2002) stated that the presence of a male in a house and the employment status of the household's head have a positive relationship. His study was based on the Pakistan Integrated Household survey (1998-99) using cross-sectional data of women aged 15-49 and established that infants have an inverse effect on women's choice to participate in economic activity. Ejaz (2007) measures the participation of females in the Pakistani labor force by using cross-sectional data of PSLM 2004-05. She utilized the Probit and logit model and revealed that those females with agricultural land ownership and vehicle access are more likely to contribute to the labor market.

Azid, Khan, & Alamasi (2010) found that literacy and the number of school-going children are the essential household elements that impact female contributions to the labor market. They used data from the survey of four thousand households with women between 16 to 60 during the year 2004-05 from the district of Punjab. Ejaz (2011) paid attention to the endogeneity problem where the author included fertility and home appliances as endogenous covariates. Furthermore, the gender wage gap was being utilized to measure empowerment qualitatively. The results are derived from Probit and instrumental variable methods from PSLM 2006-07. The findings revealed an inverse and significant relation between fertility and the gender wage gap and female participation in the labor market. Additionally, they found a direct and significant relationship between ownership of home appliances and FLFP.

2.6 Summary

This chapter presented a detailed review of existing literature on social norms and women's labor force participation, and then the research gap was identified. The chapter discussed studies on how an unbalanced gender ratio affects FLFP to get valuable insights from literature to address the first hypothesis, e.g., studies by Grosjean and Khatter (2019), Cardoso and Morin (2018), and Teso (2014) present excellent empirical evidence in this regard. Since causes of the unbalanced gender ratio presented in these studies are not relevant in Pakistan's context, which in turn led to explore what factors could contribute to variations in gender ratio in Pakistan. After evaluating the literature in section 2.3, it was understood that extreme weather events trigger rural-urban migration in Pakistan, which is mainly male-dominated (Ishfaq, 2019), and could be the

reason for variation in gender ratio, which is entirely exogenous. After linking the phenomenon of climate-induced out-migration with gender ratio, this research suggests these circumstances make gender norms "increasingly costly," the notion presented in (Bertrand et al., 2015) affecting women's labor market participation.

The second hypothesis of this research intends to explore the causal effect of women's LFP on EGRB. The idea is that access to financial resources increases women's bargaining power in the household, likely affecting their decision to participate in the labor market, making gender role beliefs more equitable. As discussed in section 2.4, cash transfers have a mix of positive and negative impacts on household labor supply (Baird, McKenzie, and Ozler 2018). Pakistan's BISP cash transfer is also found to have mixed results concerning women's labor force participation. BISP cash is disbursed only to women in the household, and the eligibility of women for cash transfer depends on their poverty score. Therefore, this research will explore whether access to this cash affects the empowerment or decision-making of beneficiaries and non-beneficiaries of this program. To explore this link, this research utilizes data from Pakistan's largest cash transfer program Benazir Income Support Program (BISP), named after the first and only women Prime Minister of Pakistan, Benazir Bhutto.

Lastly, the discussion in section 2.5 explicitly focuses on empirical evidence from Pakistan based on analyzing determinants of FLFP. It is observed that there is a shortage of literature on FLFP in Pakistan that explicitly explores the phenomenon of causal identification while examining FLFP. This present research, therefore, intends to fill this gap (1) by exploiting district-level exogenous variation in gender ratio (2) by cash transfer to women beneficiaries in analyzing the bidirectional causality between EGRB and women's LFP.

CHAPTER 3 EMPIRICAL STRATEGY

In social science research, causal identification has become an essential part of evaluating the policy effects. With time, advancements in computing methods have facilitated researchers to analyze the causal relationship between variables. While randomized control trials (RCTs) are considered gold standards in evaluating causal relationships, researchers also determine causality using observational or quasi-experimental data. Applied microeconomists have developed causal inference tools for quantitative social science research (Angrist and Pischke, 2010) and the tools are now widely used in other disciplines, not only in Economics.

Identifying causal mechanisms and their testing has crucial implications for understanding a natural phenomenon and undertaking scientific research. Causality happens to be an interesting topic and has been extensively researched by many scholars (Cameron and Trivedi 2005). Angrist and Krueger (1999:8) explain identification strategy as a combination of "a clearly labeled source of identifying variation in a causal variable and the use of a particular econometric technique to exploit this information." When examining the causal analysis of non-experimental data, care should be given that the design should yield coefficients representing a true causal relationship, not only correlation, which could be misleading.

In the case of the so called quasi-experimental data analysis, the counterfactual argument serves as the foundation for causal examination. In a natural experiment, individuals are randomly placed in the treatment and control groups. If we look at the situation from the viewpoint of the treatment group, a counterfactual condition is "what would we have observed on *y* for the individuals in the treatment group had they not received the treatment?" Alternatively, if we first consider individuals in the control group, the counterfactual condition would be, "what would we have observed on *y* for the individuals in the control group had they received the treatment?" Comparing two situations of the real world, i.e., actual versus counterfactual, permits causal analysis. It is done in the randomized experiment, obtained by randomly allocating participants in treatment and control groups, a guaranteed method to tackle endogeneity.

Causation can be explored using observational data, and a descriptive analysis can elaborate patterns, trends, and correlations among the variables. Nevertheless, causality

cannot be determined based on correlation patterns because of potential possibilities of reverse causality, omitted variable, and measurement error. In the presence of these problems, ordinary least square estimates (OLS) are biased and inconsistent and lead to the problem of *endogeneity*; much of what we will discuss in this chapter focuses on explaining what it is and how to deal with it. This endogeneity can be removed if we can identify an instrumental variable (IV): an exogenous variable that explains variation in endogenous variable and is not correlated with the error of the model. The results obtained with IV will identify the causal effect of an endogenous variable on the outcome variable. In case of ignoring the problem of endogeneity, the estimate is inconsistent

3.1 Instrumental Variables

This research uses the instrumental variable method on pooled cross-section data (PSLM) and panel data (BISP). In this section, I describe the IV and its related assumptions in detail and methods to test the required assumptions of the IV. Endogeneity is the situation in which an explanatory variable and error term are correlated with each other (Wooldridge, 2013). In this situation, a non-random error term leads to inconsistent estimation, implying that an explanatory variable's point estimate fails to converge to the population estimates with an increase in sample size. In the case of positive correlation between an explanatory variable and error term, the coefficient captures the impact of that explanatory variable on the outcome variable as well as the effect of other unobserved factors that may explain the variation in the outcome variable, thus making the interpretation useless (Antonakis, Bendahan, Jacquart, & Lalive, 2010). If the effect of unobserved variables is not accounted for, the impact of the explanatory variable is likely to be biased, called endogeneity bias. , (Chintagunta, Erdem, Rossi, & Wedel, 2006).

To explore the first hypothesis of this research, i.e., the causal effect of EGRB on women's LFP, we utilize the pooled cross-section data of PSLM for four years. The following equation depicts the initial regression equation employed, which is estimate through OLS.

$$flfp_{it} = \beta_0 + \beta_1 GenderRatio_{cdt} + X_{it}\,\delta + \varepsilon_{it} \qquad (3.1)$$

The dependent variable is a binary variable for women's LFP that takes value {0,1} measured for each woman '*i*' in year '*t*' of the four cross-sections surveys of PSLM. '*GenderRatio*' is age cohort-district-year specific, an explanatory variable that is endogenous in the model. For other explanatory variables, X_{it} capture the two indicators of social norms discussed in the previous section, i.e. 'employment decision' and 'permission to work,' they are measured for each woman '*i*' in year '*t*' and a set of other control variables considered for analysis, also calculated for each woman '*i*' in year '*t*.' The list of control variables and their descriptive analysis is presented in Appendix Table 23.

There are three reasons for the condition when exogeneity, i.e., error term uncorrelated with explanatory variable, is violated. Wooldridge (2013) mentions them as the omission of variables, errors-in-variables, and simultaneous causality. Within the context of this research, the 'gender ratio' is an endogenous variable and one of the sources of which is an omitted variable. The logic is that since gender ratio is an age cohort-district-year specific measure, there may be some unobserved heterogeneity in gender ratio because of the presence of a time-variant socioeconomic condition of a particular district in a specific year. Another source of endogeneity is simultaneous/reverse causality, which happens when a dependent variable or one or more independent variables are jointly measured, i.e., when dependent and independent variables cause each other simultaneously (Wooldridge, 2013). This source of endogeneity is assumed to be in the gender ratio. As discussed in the first chapter, gender ratio and women's LFP can simultaneously affect each other as changes in women's LFP can affect the composition of gender ratio and vice versa.

To address the issue of endogeneity and have causal inference, the two-stage least squares (TSLS), or instrumental-variable estimation, allows for consistent regression analysis with endogenous explanatory variables. TSLS is a treatment to endogeneity that arises because of reverse causality, omitted variable bias, and measurement error (Cameron & Trivedi, 2005; Greene, 2008). Since endogeneity arises because of the error term (*e*) correlating with explanatory variable (*x*), what TSLS does is removes the part of the variance in the independent variable (endogenous variable) that is correlated with the error term. For this purpose, the TSLS estimator relies on instrumental variable (z), a purely exogenous regressor of endogenous variable. Wooldridge (2002) mentions that an instrument should fulfill the following two conditions.

Cov (z, u) = 0 Instrument exogeneity condition

Cov (z, x) ≠ 0 Instrument relevance condition

The instrumental variable is uncorrelated with the error term. Therefore, it is used in the first stage of TSLS to get predicted values of the endogenous variable, which is uncorrelated with the error term (Antonakis et al., 2010). In the second stage of TSLS, this predicted value of the endogenous variable will be used to estimate the dependent variable. Instrumental variable removes the variance of endogenous variable correlated with the error term to obtain an unbiased and consistent estimate of endogenous variable. However, this consistency is received at the cost loss of efficiency because of less available information

This research utilizes district-level variation in *exposure to extreme weather events* as an instrument for gender ratio to clean the endogenous variable (gender ratio) from endogeneity. Using variables measured at the regional level as an instrument for endogenous variables appearing in individual-level equations is recommended from empirical evidence. Wooldridge (2002, p. 89) describes that "economists often use regional variation in prices or taxes as instruments for endogenous explanatory variables appearing in individual-level equations." With regards to fulfilling the two conditions of being an instrument, it is described in the previous section that *exposure to extreme weather events* is measured by utilizing data on extreme weather events occurring in a particular district, which means that occurring of any extreme weather is entirely beyond one's control and hence by no means happen to be correlated with the error term. Therefore, our instrument can satisfy the exogeneity assumption, further verified by robustness checks in the next chapter. Concerning the second condition of instrument relevance, the first stage estimate of TSLS will provide evidence of this condition by looking at the statistical significance of the *'gender ratio'.* Theoretically, the 'gender ratio' coefficient is supposed to be statistically significantly different from zero to satisfy this condition.

The first stage of TSLS estimates the following regression.

$$GenderRatio_{cdt} = \alpha_0 + \alpha_1 Exposure_{ydt} + X_{it}\,\delta + \varepsilon_{it} \qquad (3.2)$$

In the second stage regression, women's LFP will be regressed on endogeneity-free gender ratio and the rest of the control variable.

To investigate the second hypothesis, i.e., the causal effect of women's LFP on equitable gender norms, we use four rounds of a panel cash transfer data, BISP. Therefore, the panel instrumental variable method is used. Panel data allows using individual fixed effects to control for individual unobserved heterogeneity and random effect model. The choice between fixed effect and random effect model depends on several considerations, as highlighted in (Wooldridge 2013).

1) If omitted variables are assumed to be correlated with the explanatory variables included in the model, then fixed effect models are chosen, controlling omitted variable bias by including time-invariant individual fixed effects. However, if omitted variables are assumed to be uncorrelated with the explanatory variables in the model, random effect is the preferred model.

2) If there is little or no within-subject variation in the variable over time, the fixed-effect model fails to estimate that variable because of time-differencing the variables, resulting in large standard errors. There must be time-varying variables for them to be estimated using the FE model. In contrast, the random effect model estimates the time-constant variables but at the cost of resulting in biased estimates because of not controlling for omitted variables.

Within the context of this research, variation in the variable of interest (women's LFP) and the chosen instrument has influenced the selection of the panel estimation model. Women's LFP has appeared to have minimal variation over time among beneficiary households of BISP data. 24% of women participated in the labor market in 2011, which dropped slightly to 23% in 2013, increased to 26% in 2014, and dropped to 24% in 2016. So, women's LFP remains roughly the same during these years, which shows that beneficiary women's labor market status did not change substantially. Additionally, women's LFP is a small fraction of the whole population, and deriving results on a small fraction may lead to imprecise estimates. As in this research, we want to identify those women who change their LFP because of being beneficiaries of the program, which will further reduce the variation in this variable.

Moreover, the instrumental variable, official treatment status, is also time-invariant because the BISP program eligibility has not changed since the program's inception. The random effect model is the appropriate estimation method in this research because of the considerations discussed. Although the random effect is based on the assumption that unobserved variables are uncorrelated with observed variables, which is more likely to be wrong, we rely on this estimation method because of the nature of the variable of interest (women's LFP) and the chosen instrument (official treatment status). Additionally, as the standard errors will be quite large in time-invariant variables, which leads to imprecise estimates, the random effect is preferred over FE. As this research is interested in identifying the causal impact of women's LFP on EGRB, there is a need to clean women's LFP of its endogeneity, which is why we employ the random effect IV model for the estimation. The result of the Hausman test on selection between RE and FE is reported in Table 33 in the appendix, which also strongly confirms the RE over FE. The following specification presents the first stage or reduced form equation of TSLS random effect IV.

$$x_{it} = \pi_0 + \pi_1 z_{1it} + \pi_2 z_{2it} + \pi_3 z_{1it}.z_{2it} + w_{it}\delta + \epsilon_{it} \qquad (3.3)$$

In the equation above, *i* denotes cross-sectional units *N* and *t* represents period *T*. z_1, z_2 and z_3 are instrumental variables used in the first stage, i.e., official treatment status, replicated poverty score, and their interaction. Moreover, *x* denotes women's LFP, and *w* is the set of other control variables. The equation is estimated with OLS, and it purges the endogenous variable (women's LFP) of its correlation with the error term because it is endogenous. Then the part of women's LFP that is not correlated with the error term is subsequently used in the second stage regression given below.

$$y_{it} = \beta_0 + \beta_1 x_{it} + w_{it}\delta + v_{it} \qquad (3.4)$$

where

$$v_{it} = \alpha_i + u_{it} \qquad (3.5)$$

As in equation 3.3, In equation 3.4, *i* denotes cross-sectional units *N* and *t* represent period *T*. Moreover, *y* represents gender norms index computed with MCA on husband-wife gender norms coincidence, *x* is the part of women's LFP that is clean from endogeneity, *w* is the set of control variables, *v* is the composite error term. Equation 4.4 is the second stage of the random effect panel instrumental variable method. As discussed above, in RE, the unobserved heterogeneity is assumed to be uncorrelated

with explanatory variables in the model, which is why the RE model includes the unobserved heterogeneity α_i and it becomes part of the idiosyncratic error and makes a composite error v_{it} . As mentioned in Wooldridge (2013) that because α_i is part of the composite error term, the v_{it} are serially correlated across time. To remove this serial correlation, RE models are estimated through Generalized Least Square (GLS).

3.2 Regression Discontinuity Design

The main goal of this research is to utilize Instrumental Variables methods on pooled cross-section data of PSLM and panel data of BISP to address the causal relationship between women's LFP and EGRB. However, the research additionally employs regression discontinuity design (RDD) to validate the effect of BISP cash transfer on LFP of women who are just above the eligibility poverty score cutoff and compare them to those who are just below the cutoff. RD design has been extensively used in the literature during the past five years to answer various questions such as labor supply, disability program, unemployment benefits, the effect of health care programs on health outcomes, and the impact of remedial education programs (Lee and Lemieux, 2010). Jacob, Zhu, Somers, & Bloom (2012) state that "Regression discontinuity (RD) analysis is a rigorous non-experimental approach that can be used to estimate program impacts in situations in which candidates are selected for treatment based on whether their value for a numeric rating exceeds a designated threshold or cut-point."

Regression discontinuity (RD) is an approach that deals with the treatment effect of assignment into treatment and control groups. Thistlethwaite and Campbell (1960) developed this approach to estimate treatment effects. This method is based on a factor that determines how some observations are assigned treatment. The discontinuity in the method is based on the cutoff point for the identified factor, which differentiates between treatment and control groups, therefore, enables researchers to compare observations just above and below the cutoff point (Hartmann, Nair, & Narayanan 2011). Several researchers utilized this approach to exploit threshold rules used by educational institutions to analyze the impact of financial aid on class size (Angrist & Lavy, 1999) and school district boundaries (Black, 1999).

The BISP program is non-random, but its execution allows for analyzing the causal impact of this program on outcome variables. The poverty score cutoff estimated

through the PMT method is used as an identifying factor to determine the eligibility of the program recipients, thereby making the RDD an appropriate technique to undertake causal identification. Since the program is not random, the fuzzy regression discontinuity design is used in studies (Ambler and Brauw, 2016; Ambler and Brauw, 2019). Some households above the cutoff are treated, while some below the cutoff are placed into the control group. To evaluate the impact of BISP cash transfer on women's labor force participation, the following fuzzy RDD model is used:

$$Y_i = \alpha + \tau D_i + f(X_i - c) + \Sigma \beta_i W_i + \varepsilon_i \qquad (3.6)$$

Where in equation 1, Y_i is the outcome of interest, which is women's labor force participation in our case. X_i is the poverty score of household i called a running variable and $f(X_i - c)$ is the function of the distance from the cutoff $D_i = 1(X_i \leq c)$, i.e., $D_i \epsilon \{0,1\}$ is 1 if the household poverty score is below the cutoff c of 16.17 and 0 otherwise τ estimates the impact of the BISP cash transfer program, the treatment effect. Finally, $\Sigma \beta_i W_i$ captures the effect of other control variables.

An essential choice in RDD analysis is selecting the range of observations of the running variable, i.e., the identifying factor to carry out the causal impact of the program. The range of observations is called bandwidth. RDD produces unbiased results within a specific threshold limit. However, as more and more observations are added around the cutoff, the results are susceptible to bias because observations become less comparable. On the other hand, adding observations around the cutoff provided more precise estimates. Therefore, to choose the optimal bandwidth, researchers are faced with the trade-off between the bias of including more observation against the treatment effect variance. Calonico, Cattaneo and Titiunik (2014) developed an approach to select the optimal bandwidth. The authors propose an estimator which utilizes local linear regression on both sides of the cutoff and considers bias-corrected confidence interval. The estimator uses a triangular kernel for data in regressions, as suggested by (Lee and Lemieux, 2010).

Another condition RDD requires is that units cannot strategically sort themselves around the cutoff, so women must not manipulate their poverty score (Lee and Lemieux, 2010). If this condition holds, then the allocation of units into treatment and control groups will be taken as 'random.' This condition is testable following density test by comparing the density of probability distribution of poverty score around both sides of the poverty

score cutoff (McCrary, 2008). Moreover, RD estimates will be validated and falsified by investigating the impact of cash transfer on predetermined covariates and exploiting a placebo cutoff (Calonico etal, 2019).

3.3 Summary

This chapter presents econometric methodologies used for causal identification between women's LFP and EGRB. For the first empirical study, the cross-section PSLM survey data is utilized because the pooled cross-section instrumental variables method is used. The BISP cash transfer data is used for the second study, which is panel data. Therefore, the panel IV method is employed. Concerning panel IV, this research utilizes the random effect IV (RE-IV) estimation method. The reason for choosing RE-IV is supported by a slight variation in the variable of interest women's LFP over time and time-invariant nature of one the IV, official treatment status of households.

Moreover, the Hausman test also strongly fails to reject the null hypothesis of consistency in FE and RE estimators, making random effect the preferred estimation method for the data used. In addition to IV, this research intends to validate the causal impact of BISP cash transfer on women's LFP. As the poverty score threshold determines the BISP program eligibility, we leverage this data feature and utilize the regression discontinuity design (RDD) to illustrate the causal effect of BISP cash transfer on LFP of women who are located just below and just above the threshold. RDD is known to be valid around the threshold, whereas the IV estimator is a Local Average Treatment Effeect estimator (LATE). In this sense, both estimators identify the causal effect *at the margin*, figuratively speaking: for an additional woman joining the labor force.

CHAPTER 4 DATA AND OPERATIONALIZATION OF VARIABLES

This chapter presents sources of different datasets used in the research. It will be followed by the operationalization of key variables of interest and their descriptive statistics.

4.1 Pakistan Social and Living Standard Measurement Survey

This research utilizes a household-level Pakistan Social and Living Standard Measurement Survey (PSLM)[2] in study 1 to investigate the first hypothesis. This survey is collected by two-stage stratified sampling and funded by Public Sector Development Program (PSDP) from July 2004 till June 2015. The survey data helped the government of Pakistan formulate a poverty reduction strategy and investigate the progress of development plans at the provincial and district level, including both rural and urban areas, in the context of Millennium Development Goals (MDGs). Therefore, this survey was designed to monitor the progress of MDG indicators. The first round of this survey was conducted in 2004-05, and it has been undertaken alternatively at district and provincial levels. Until June 2015, ten rounds of PSLM have been completed, and the 10th is the final round of the PSLM project[3]. At the district level, PSLM mainly collects information on multi-dimensional poverty indicators from 80,000 households. In comparison, the provincial level survey collects data on social indicators and household income and expenditure from 26,000 households.

For this research, four rounds of PSLM have been included in the analyses. These rounds were conducted in 2005-06 (round II), 2007-08 (round IV), 2011-12 (round VII), and 2013-14 (round IX), which are conducted at the provincial level. These rounds are selected because information on social indicators, specifically 'women in decision making,' are only available in the rounds mentioned above. Moreover, the rest of the demographic, household-level and employment variables are also taken from this data set. The operationalization of each variable is discussed in section 4.3 of this chapter.

[2] These datasets are available to download from the website of Pakistan Bureau of Statistics (www.pbs.gov.pk).

[3] Post 2015, UN has adopted new development goals called Sustainable Development Goals (SDGs) because of which PSLM questionnaire has been revised to incorporate the new development indicators. Field activities are in progress in order to collect data on revised questionnaire.

So, this study has utilized four cross-sections of this survey, which are then pooled together to be used in the final analysis. A summary of the number of households, observations, and averages of males and females in four rounds of PSLM is presented in table 4.1 below.

Table 4.1 Profile of PSLM Survey 2005-06, 2007-08, 2011-12 and 2013-14

	2005-06	2007-08	2011-12	2013-14
Number of Households	15,453	15,512	15,807	17,978
Gender:				
Men	0.51	0.51	0.51	0.50
Women	0.49	0.49	0.49	0.50
Observations	105608	101521	99800	119018

Source: PSLM Survey 2005-06, 2007-08, 2011-12 and 2013-14, own calculations

4.2. Benazir Income Support Program Database

This thesis utilizes a cash transfer program database named Benazir Income Support Program (BISP) to explore the hypothesis of study 2. BISP contracted Oxford Policy Management (OPM) to conduct an independent impact evaluation of the BISP cash transfer program. OPM has undertaken five impact evaluation surveys, a baseline in 2011, three midline surveys in 2013, 2014, 2016, and the end line survey in 2019. However, this study uses baseline (2011) data and three midline panel surveys (2013, 2014, and 2016). However, an end-line survey of 2019 was not accessible at the time of writing this research. The baseline (2011) and two midlines, 2013 and 2014, covered almost the same households, whereas the 2016 sample covered more households near the poverty score cutoff. Before collecting the 2016 sample, a resampling exercise was done to add new treatment and control households from the existing primary sampling units (PSUs), the central research communities. In their study, Ambler and Brauw (2019) reported that through this exercise, all households matched with BISP management information system (MIS) were retained, given that they had a poverty score of less than 21.17. Secondly, PSUs from the evaluation sample were matched with BISP MIS. Thirdly, nine control and nine treatment households were randomly selected from each PSU.

Two different analyses are undertaken as the number of panel households is larger in 2011, 2013, and 2014 samples compared to 2011, 2013, 2014, and 2016 samples. First,

the panel sample from all four years, i.e., 2011, 2013, 2014, and 2016 is used in the analysis, which will help employ the consistent panel sample from these years, and a 5-year impact of this cash transfer program is estimated. Second, to assess the 3-years effect of this program, a separate analysis is carried on a panel sample from 2011 to 2014. This analysis will leverage the comparatively large number of panel households in these three years. The following tables 4.2,4.3, and 4.4 provide information on the total number of households in four rounds of BISP surveys and the number of panel households from 2011 to 2016 (3497) and 2011 to 2014 (7563) disaggregated by their treatment status, respectively.

Table 4.2 Summary of Households in BISP Data by Treatment Status

	2011	2013	2014	2016
Control	4517	4230	4230	4017
Treated	4158	3991	3991	5300
Observations	8675	8221	8221	9317

Source: BISP Survey, own calculations

Table 4.3 Summary of 4 Years Panel Households in BISP Data by Treatment Status

	2011	2013	2014	2016
Control	1380	1380	1380	1032
Treated	2117	2117	2117	2465
Observations	3497	3497	3497	3497

Source: BISP Survey, own calculations

Table 4.4 Summary of 3 Years Panel Households in BISP Data by Treatment Status

	2011	2013	2014
Control	3932	3932	3932
Treated	3631	3631	3631
Observations	7563	7563	7563

Source: BISP Survey, own calculations

4.3 Emergency Events Database

The Centre for Research on the Epidemiology of Disasters (CRED) initiated the Emergency Events Database (EM-DAT) in 1988. This data was established with the support of the Belgium Government and World Health Organizations to undertake humanitarian action at national and international levels. EM-DAT comprises essential

data on the incidence and effect of more than 22,000 mass disasters that occurred in the world from 1900 until today. The database is compiled from various sources, including UN agencies, non-governmental organizations, insurance companies, research institutes, and press agencies.

Several studies have used the EM-DAT dataset to measure the incidence of weather events, such as windstorms, drought, and other natural disasters, and observe their impact on national income. (Raddatz 2009; Loayza, Olaberria, Rigolini, & Christiaensen, 2012). EM-DAT follows inclusion criteria to include a specific disaster in its database. This criterion is disaster has killed ten or more people, an official emergency state is declared, international assistance is required, or one hundred or more people are affected. However, this approach has a challenge in that it covers only the events with notable economic impact.

For the present research, EM-DAT is used to construct a variable for exposure to extreme weather events used as an instrument in Two Stage Least Square (TSLS). The period for extreme weather events is from 1950 to 2014, with some missing years in between. The reasons for missing years are (1) according to the data source, no extreme event took place in Pakistan in that specific year (2) the years when there was an earthquake which this research has not considered. The reason for selecting the starting year as 1950 is because it is the most recent year in the dataset since Pakistan's independence in 1947. Before 1950, the data is available for 1945, which is not considered because Pakistan was not an independent country; instead, it was part of India. The period for extreme weather events data ends in 2014, chosen because the household level data from PSLM used in this research is until 2014. EM-DAT classifies information on disaster into two broad categories, namely, natural and technological. The details of this classification are given in Appendix Table 1.

Since present research is confined to natural disasters that occurred because of climate change, three subgroups of disasters under the disaster group 'natural' have been considered, i.e., meteorological, hydrological, and climatological. Moreover, under these three disaster subgroups, not all the disasters happened in Pakistan. According to the EM-DAT database, Pakistan has experienced a flood, heat wave, storm, landslide, and drought from 1950 to 2014.

Figure 4.1 presents the frequency of extreme weather events in Pakistan, such as flood, drought, extreme temperature, landslide, and storms. It can be observed that there is a sharp rise in the frequency from as low as one event to a maximum of 10 events in a year. The year with a maximum number, i.e., ten extreme weather events, are 2007 and 2019, where flood contributed the most in total occurrence of extreme weather events, which is six times out of ten and eight times out of ten in the year 2007 and 2019, respectively.

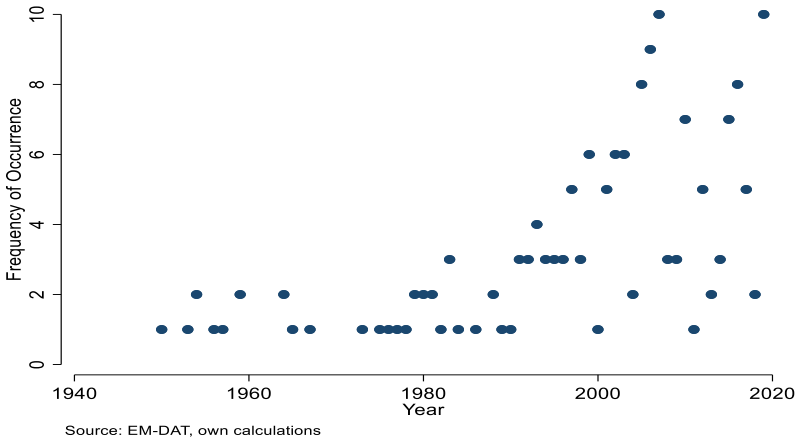
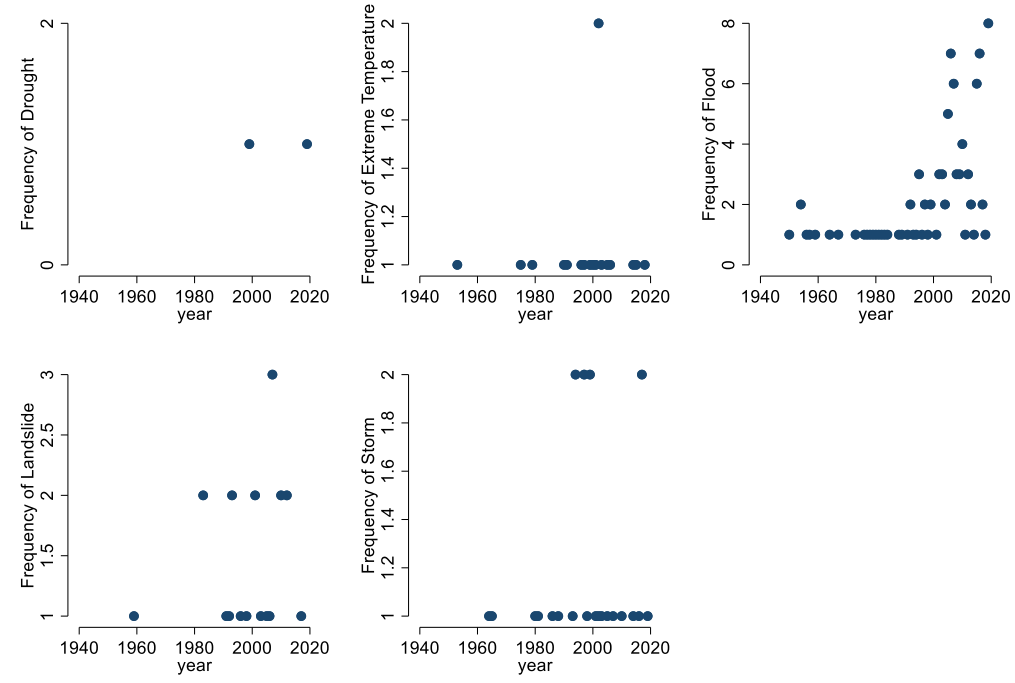


Figure 4.1 Frequency of Occurrence of Extreme Weather Events in Pakistan 1950-2019

Thus, it can be noted that floods occur most frequently in Pakistan compared to other extreme weather events. This observation can further be verified in figure 4.2, which presents a scatter plot of the frequency of occurrence of each event separately. The figure shows that flood has the highest frequency of occurrence among all the events. The country has been experiencing floods almost since its inception, and its frequency has risen sharply after the 1990s. However, drought has a minor occurrence among the other events and is mainly happening in recent years. In contrast, extreme temperatures and storms occur once almost every year since the late 1970s and early 1980s, respectively. Lastly, land sliding is occurring quite moderately throughout the years with a maximum frequency of 3 in 2007.



Source: EM-DAT, own calculations

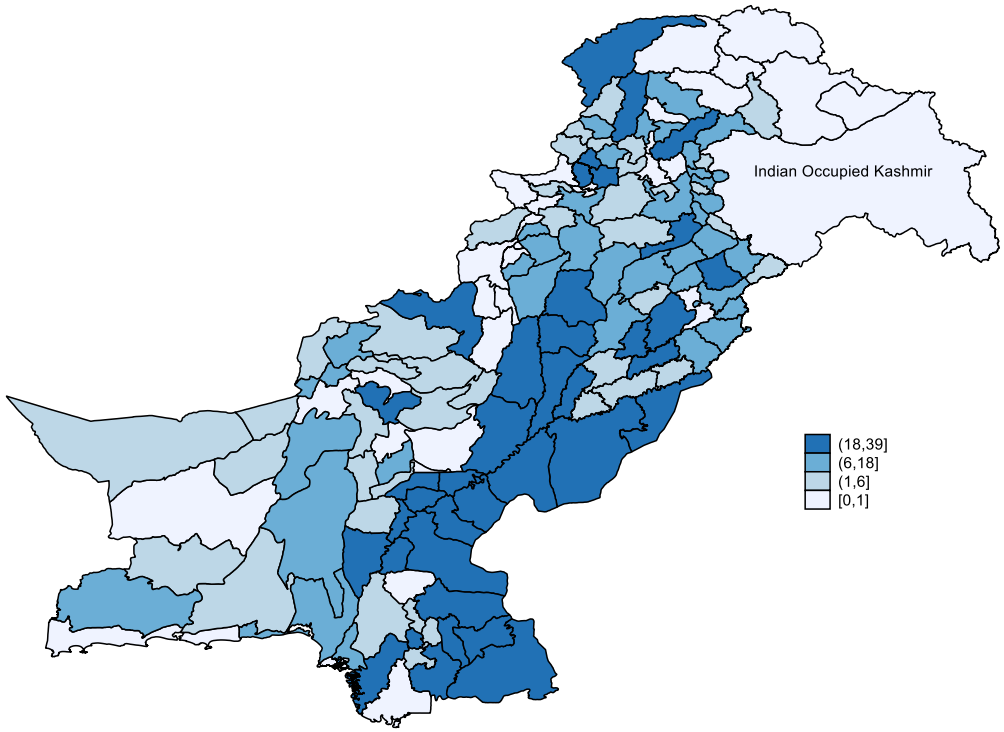
Figure 4.2 Frequency of Occurrence of each Extreme Weather Events in Pakistan 1950-2019

Apart from the frequency of occurrence of extreme weather events, EM-DAT provides information on geographic coordinates, i.e., longitudes and latitudes of the region, number of people affected, number of deaths, and how much total damage (in US dollars) is caused by the disaster. However, in the early years of data, the information on event occurrence is available at the provincial rather than the district level. In contrast, in more recent years, the events data has been provided at the district level. Therefore, to identify the specific districts in each disaster-hit province in the early years and some recent years, the data is taken from two sources (1) Global Active Archive of Large Flood Events, Dartmouth Flood Observatory, University of Colorado[4] , and (2) and Annual Flood Report 2010, Ministry of Water and Power, Government of Pakistan[5].

Figure 4.3 presents the map of Pakistan with the frequency of occurrence of extreme weather events at the district level. This figure is drawn by merging the EM-DAT dataset with the Pakistan Stata-format shapefile[6]. A shapefile is a format for the

[4] G.R.Brakenridge, "Global Active Archive of Large Flood Events", Dartmouth Flood Observatory, University of Colorado. http://floodobservatory.colorado.edu/Archives/index.html.
[5] http://mowr.gov.pk/wp-content/uploads/2018/06/Annual-Flood-Report-2010.pdf
[6] www.arcgis.com

exchange of geographical data, and it contains the geometric information that draws points, lines, or polygons on a map and the data attributes.



Source: EM-DAT and Pakistan Shapefile, own illustration

Figure 4.3 Frequency of Occurrence of Extreme Weather Events at District Level from 1950-2019

On the map, each area with boundary lines shows a district[7] boundary. Whereas area highlighted as "Indian occupied Kashmir" is the disputed territory between India and Pakistan. As mentioned earlier, floods contribute the most among all the extreme weather events in Pakistan. Therefore, the districts highlighted as dark blue are mainly the districts primarily affected by the flood. According to the Annual Flood Report (2017), these are the districts from Punjab and Sindh, where the country's largest river (River Indus) flows along with its tributaries Jehlum and Chenab. Therefore, Punjab and Sindh are affected mainly by river floods, specifically during heavy monsoon rainfall. Whereas hill torrents affect hilly areas of KP, Balochistan, Gilgit Baltistan, and Azad Jammu and Kashmir (AJK), and they are exposed to floods in River Kabul and River Swat. Apart from this, large cities such as Karachi, Lahore, and Rawalpindi are mainly

[7] There are 154 districts in Pakistan. Punjab has 36 districts, KP 34, Balochistan 34, and Sindh 29. Azad Jammu and Kashmir is a self-governing state under Pakistan's control, has 10 districts. Similarly, Gilgit Baltistan, is a region administered by Pakistan as an administrative territory, has 10 districts. Lastly, Islamabad Capital Territory (ICT) is considered as a separate district.

affected by improper drainage systems, especially when it rains heavily during the monsoon.

4.4 Operationalization of Variables

This research utilizes two survey data sets, i.e., PSLM and BISP, and extreme weather events data from EM-DAT discussed in the previous section. These data sets are used to apply causal identification strategy, i.e., Instrumental Variable (IV). This section elaborates in detail the construction of key variables from each of the three datasets mentioned above.

4.4.1 Women's Labor Force Participation

Women's LFP is the key variable in the analysis. The following sections elaborate in detail on the operationalization of FLFP from both survey data sets used in this research, i.e., PSLM and BISP.

Women's LFP from PSLM

As the hypothesis of the first study in this research aims to analyze the causal effect of gender role beliefs on women's LFP, women's LFP is the dependent variable in this first part of the analysis. A person in the labor force is considered engaged in some economic activity and looking for work. Therefore, the labor force is comprised of the number of employed and the number of unemployed. The total labor force indicates the total labor supply that can be utilized to produce goods and services. This definition of LFP is in line with what International Labor Organization (ILO) has proposed. However, contrary to the labor force, an inactive person is neither looking for work nor currently working. Thus, in this research, from PSLM data, female labor force participation is measured by combining two questions asked from the respondents who are ten years of age and older (1)*"Did you do any work for pay, profit or family gain during the last month at least for one hour on any day?"* The respondent could select from the options *'yes'* or *'no.'* Those who answer *'no'* to this question are further asked (2)*"Even if did not work last month, did you have a job or enterprise such as a shop, business, farm, or service establishment (fixed/mobile) during the last month?* Respondents are given *'yes,' 'no, but seeking work'* and *'no, not seeking work.'* After combining the responses of these two questions, women's LFP is coded '0' as 'No' for those who are '*not seeking work'* and '1' as 'Yes' for those who responded *'yes'* to the first question and *'yes'* and *'no but seeking work'* to the second question. Table 4.5

provides a summary of women's LFP in four survey years, and it is seen that the percentage of women who are reported to have 'not seeking work' is higher than women who are working and looking for work.

Table 4.5 Summary: Women's Labor Force Participation from PSLM

	2006	2008	2012	2014
Women's Labor Force Participation:				
No	0.81	0.83	0.83	0.80
Yes	0.19	0.17	0.17	0.20
N	37310	36055	35884	43076

Women's LFP from BISP

To investigate the second study's hypothesis, i.e., the causal impact of women's LFP on equitable gender norms, we utilize four rounds of panel BISP cash transfer data. However, there are two differences in LFP questions in PSLM and BISP surveys. First, in the PSLM survey, respondents answer about their LFP during the preceding *'month'* of the survey, whereas in the BISP survey, respondents answer about their LFP during the prior *'week'* of the survey. Secondly, the BISP survey asks labor force participation questions from respondents 15 years of age and older, and PSLM measures LFP since ten years of age. The selected items from BISP to measure FLFP is (1) "*during the past week, did you do any kind of work for pay, profit or family gain?"* The respondent could select from the options *'yes'* or *'no.'* Those who answer *'no'* to this question are further asked (2)*"Even if did not work last week, did you have a job or enterprise such as farm, business, service, establishment (fixed or mobile)?"* For which respondents are given options *'yes'* and *'no.'* Those who answer *'no'* are further asked (3) "*Did you look for a job in the last four weeks?"* for which respondents again answer either *'yes'* or *'no.'* After combining the responses of these three questions, women's LFP is coded '0' as 'No' for those who answered *'no'* in question 3 and women's LFP is coded '1' as 'Yes' for those who responded *'yes'* to all these three questions. The summary of this variable from the BISP data set is reported in table 4.6. Like PSLM, the percentage of women not participating in the labor force is much higher than those part of the labor force in all four survey years.

Table 4.6 Summary: Women's Labor Force Participation from BISP

	2011 Control	2011 Treated	2013 Control	2013 Treated	2014 Control	2014 Treated	2016 Control	2016 Treated
No	0.80	0.76	0.81	0.77	0.78	0.74	0.79	0.76
Yes	0.20	0.24	0.19	0.23	0.22	0.26	0.21	0.24
Observations	8405	8074	8989	8610	7993	7741	8774	12098

Source: BISP Survey, own calculations

4.4.2 Gender role beliefs

There are several indicators to operationalize gender role beliefs from BISP and PSLM: women's self-reported responses on different women empowerment domains. However, as discussed in the first chapter about social desirability bias in using self-reports on empowerment indicators, this thesis utilizes a distinct approach to operationalize women empowerment. However, the original indicators from both surveys are also incorporated in the analysis. The following tables summarize these original indicators as well as the new specification of EGRB.

Indicators of gender role beliefs from PSLM

For the measurement of gender role beliefs, two items are used from PSLM. These items are available only in four provincial rounds of PSLM, which is why these four rounds are mainly selected for this research. There is a specific module in PSLM on 'women in decision making composed of several questions related to the household decision-making process regarding female education, employment, marriage, children, and decision on the purchase of household items. Since this research's central theme is women's LFP, the selected items are based on female employment decisions. One of the items asks respondents, *"who in your household decides whether you can seek or remain paid employment?"* and these respondents are all the females in a household who are between the age of 15 and 49. The variable constructed from this item is named 'employment decision' in further analysis. Females select potential answers from the following options:

1 Woman herself
2 Head/Father of the household decides alone
3 Head/Father in consultation with his/her spouse

4 Head/Father in consultation with the woman concerned
5 Head/Father and spouse of the head in consultation with the woman concerned
6 Head/Father and other male members decide
7 Other combinations of persons decide
8 Too old to study or work
9 The woman concerned has no interest in study/work

Variable of gender role beliefs is dichotomized based on these options mentioned above. Options 1, 4, and 5 are combined and coded as 1 to categorize gender role beliefs, i.e., 'modern/non-traditional.' Women who select option 1 have attributes of non-traditional gender norms as she decides about her participation in the labor market. Similarly, options 4 and 5 also, to some extent, represent the presence of modern social norms because women are somehow involved in the decision-making process about participation in the labor market. On the other hand, options 2, 3, 6, and 7 are combined and coded as 2 to make the second category of social norms, i.e., "traditional." Since these options do not show the involvement of the concerned women in the employment decision, these options show intuition to consider them as 'traditional.' However, options 8 and 9 are not considered for analysis as women themselves do not have work-related orientation.

The second item chosen to construct gender role beliefs asks respondents, *"why are you not actively seeking paid work?"* Again, these respondents are women between 15 and 49, and neither paid employed nor seeking work. This variable is named 'permission to work' in the analysis. Women choose the answer from the following options.

1 Not permitted by husband or father to work outside home
2 Don't want to work outside home
3 Not enough job opportunities in the region
4 Pay too low
5 Too busy doing domestic work
6 Too Old / Retired / Sick / Handicapped
7 Don't know whether there exists an opportunity
8 student
9 Other

In response to this question, women who select options 2, 3, 4, and 5 are combined and coded as 1 to make the first category of this variable, i.e. 'permitted.' The options are categorized as 'permitted' because they portray that woman herself does not want to work, rather than being not permitted by household head. Options 1 and 7 are combined and coded as 2 'not permitted.' Option 1 specify women are not allowed. However, option 7 informs about lack of knowledge or information on woman's part, which in turn can be thought of as 'not permitted' because, in Pakistan's context, a woman with a lack of knowledge about the external environment can be considered as the one who is not permitted to go out of the home. Lastly, options 6, 8, and 9 are combined and coded as 3 'inactive.' Table 4.7 below shows the summary of these two measures of gender role beliefs for the respective four years of the PSLM survey. It can be observed that for the variable 'employment decision', the percentage of women who are reported to have 'traditional' norms regarding their decision to participate in the labor market is relatively higher than women who reported 'modern' norms. However, for the variable 'permission to work', the data suggests that the percentage of 'permitted' to work is relatively higher than those who are 'not permitted' to work.

Table 4.7 Summary: Indicators of Social Norms from PSLM

	2006	2008	2012	2014
Employment Decision:				
Non-Traditional	0.17	0.20	0.22	0.24
Traditional	0.83	0.80	0.78	0.76
Permission to work:				
Permitted	0.55	0.51	0.52	0.53
Not Permitted	0.36	0.38	0.36	0.36
Inactive	0.09	0.11	0.12	0.11
N	21221	18608	20398	24341

Source: BISP Survey, own calculations

The table above shows the summary of responses on gender role beliefs indicators from PSLM data. It is noted from the table that 83 % of women in 2006 did not decide themselves on whether they could seek or remain in paid employment. However, this percentage gradually decreased in subsequent years, i.e., it dropped to 80 %, 78 %, and 76 % in the years 2008, 2012, and 2014, respectively. On the other hand, in 2006, 17 percent of women decided themselves about their employment or are involved in this

decision along with their parents. Nevertheless, this percentage improved in the following years. Related to the question on 'permission to work,' it is observed that for women who are neither employed nor seeking paid work, more than 50% of them have permission to work. However, 35-37 % are not permitted. This summary reveals a striking phenomenon that, on the one hand, there is 80 percent of women who do not have a say in their employment decision. On the other hand, in the second indicator, 53% of women have permission to work on average. It displays the fact that there is inconsistency in these responses regarding who makes women's employment decisions in the household and permission to work, which makes these indicators susceptible to bias.

Therefore, there is a reason to believe that, although these indicators reflect women's responses on their employment decision and whether they are permitted to work. Still, women's stated preferences may be different from what happens. So, using them in the analysis as variables of interest to measure EGRB may lead to biased conclusions. Therefore, this research intends to use these indicators as control variables rather than to identify causal links. Thus, quantitative analysis prefers to use a data-driven indicator on the working-age gender ratio; the rationale for using this measure is discussed in section 4.4.3.

Indicators of gender role beliefs from BISP

BISP survey collects data on different domains of women empowerment. These are women's decision-making in the household, men's and women's gender role beliefs, women's freedom of mobility, women's access to money, and voting in elections. However, these survey questions collect data on women's self-reported responses on each indicator, which means the responses could be potentially biased. Moreover, as discussed in the first chapter, women's emancipation will not make any difference if her husband has traditional views. Therefore, it is considered crucial to construct such a measure that incorporates this critical aspect. The original women empowerment and decision-making indicators from the survey do not take into consideration this approach. So, a measure of husband-wife gender norms coincidence is constructed by using responses to six gender role beliefs questions. The original questions from the survey are:

1. The important decisions in the family should be made only by the men of the family.
2. If the wife is working outside the home, then the husband should help her with the household chores.
3. A married woman should be allowed to work outside the home if she wants to.
4. The wife has a right to express her opinion even when she disagrees with what her husband is saying.
5. A wife should tolerate being beaten by her husband in order to keep the family together.
6. It is better to send a son to a school than it is to send a daughter

Respondents are given four options, *'strongly disagree,' 'disagree,' 'agree,'* and *'strongly agree.'* These questions and their responses are transformed for them to use in further analysis. First, the answers are converted into the binary form of 0/1 by combining *'strongly disagree'* and *'disagree'* (coded as `0'), *'agree,'* and *'strongly agree'* (coded as 1). Second, the statements of 1st, 5^{th,} and 6th questions are rephrased so that respondents who *'disagree'* to the original questions are considered as if they *'agree'* to the rephrased questions. This is done to have consistency in the responses, i.e., the answer *'agree'* to these statements will represent non-traditional gender role beliefs and *'disagree'* to represent traditional gender role beliefs. These original questions are rephrased as follows.

- Both men and women should make important decisions in the family
- A wife should not tolerate being beaten by her husband in order to keep the family together.
- It is better to send a daughter to school rather than a son

All ever-married women in the household answer these questions. To construct the indicator of husband-wife gender norms coincidence, among all the women, only wives' responses are considered and that too considering households with one wife because with multiple wives in a household – which appears in 1.37 % of the households – it would be hard to distinguish which wife's gender norms are changing within a household. Then wife's responses are matched with their husbands' responses. This matching is done so that the variable is coded '1' when husband and wife share their gender role beliefs, i.e., either they both have traditional views – disagree to the gender norms questions – or non-traditional views – agree to the gender norms questions. It would represent gender norms '*coincidence'* between husband and wife.

On the other hand, this variable is coded '0' when there is a *conflict* in the gender norms responses by husband and wife. Conflict in gender norms means that when husband's and wife's gender norms are not the same, i.e., either wife has non-traditional views, and her husband has traditional opinions and vice versa. Constructing a variable through this approach investigates whether changing a wife's labor market participation decision due to being the recipient of cash brings '*coincidence'* or '*conflict'* in gender norms perceptions of a wife and her husband.

The following tables summarize the husband and wife self-reported responses separately on these questions and then compare them with the indicator of *husband-wife gender norms coincidence*. The summaries are presented concerning the average of these binary responses regarding their treatment status in each survey year. As these indicators are binary, the averages can be interpreted as percentages.

Table 4.8 Husbands' gender role beliefs

	2011 Control	2011 Treated	2013 Control	2013 Treated	2014 Control	2014 Treated	2016 Control	2016 Treated
Both men and women should make HH decisions	0.18	0.16	0.19	0.16	0.20	0.18	0.24	0.23
Husband should help in HH	0.70	0.69	0.71	0.70	0.68	0.69	0.75	0.76
Women can work outside home	0.46	0.47	0.46	0.44	0.41	0.41	0.49	0.51
Wife should express opinion	0.68	0.65	0.74	0.71	0.75	0.72	0.84	0.85
Women should not tolerate violence	0.30	0.27	0.29	0.25	0.27	0.22	0.17	0.16
Better to send daughter to school rather than son	0.74	0.66	0.78	0.74	0.76	0.72	0.76	0.74
Observations	3040	2983	2768	2608	2289	2200	2635	3389

Source: BISP Survey, own calculations

The table above shows that husbands have traditional views regarding decision-making in the household and the acceptability of women tolerating violence. Concerning who should make the important family decision, it is further noted that conventional views are found in beneficiary households relative to non-beneficiary. Moreover, the husband's gender role beliefs on who should make family decisions are emerging to be non-traditional over time irrespective of the beneficiary status. On average, husbands have relatively strict traditional norms regarding household decision-making and violence against women. Contrary to the decision-making indicator, the gender norms

on women facing violence are strictly traditional over time, seen from the table as declining percentages of *'agreement'* to the indicator in each successive year in both treated and control households.

Regarding the rest of the four gender norms indicators, there is a noticeable difference in average responses to non-traditional gender norms compared to the previous two indicators. It is seen that on average, approximately 40-50 percent (in case of *'women can work outside home'*) and 70-80 percent respondents (in case of *'husband to help in household chores,*' *'wife to express an opinion'* and, *'sending daughter to school rather than son'*) have non-traditional beliefs on these gender norms indicators. Moreover, husbands from beneficiary households have traditional views on these gender norms than non-beneficiary households; however, their views have slightly improved over time.

The following Table 4.9 summarizes the wife's responses on the same gender norms indicators. Regarding decision-making in the household and violence against women, wives possess similar gender norms as their husbands, i.e., only 12-20 percent agree that the wife should not tolerate violence to keep the family together. Additionally, these traditional beliefs are getting stronger over time as only 12% had non-traditional views on violence against women in 2016, which has dropped from 22% in 2011. A similar pattern is observed concerning wives' beliefs on decision-making in the household where beneficiary women have strong traditional views than non-beneficiaries. Therefore, it can be said that cash transfer has made no impact in changing men and women's gender norms on decision-making in the household and violence against women. However, as to the other four indicators of gender norms, average responses to these indicators show that wives have more non-traditional views specifically regarding *'women can work outside home'* and *'women should express opinion'* than their husbands. Their views are not changed whether they are program beneficiaries or not. In general, this simple comparison of average responses on gender role beliefs by husband and wife reveals that they share the norms on most of these indicators. Table 4.10 summarizes the averages of *'coincidence'* and *'conflict'* in gender role beliefs among wives and husbands.

Table 4.9 Wives' gender role beliefs

	2011 Control	2011 Treated	2013 Control	2013 Treated	2014 Control	2014 Treated	2016 Control	2016 Treated
Both men and women should make HH decisions								
	0.18	0.14	0.22	0.18	0.27	0.22	0.24	0.21
Husband should help in HH								
	0.72	0.72	0.78	0.77	0.75	0.74	0.73	0.73
Women can work outside home								
	0.76	0.75	0.81	0.82	0.75	0.75	0.76	0.77
Wife should express opinion								
	0.82	0.81	0.91	0.90	0.87	0.86	0.87	0.88
Women should not tolerate violence								
	0.22	0.21	0.18	0.14	0.20	0.14	0.12	0.12
Better to send daughter to school rather than son								
	0.85	0.78	0.78	0.76	0.81	0.77	0.84	0.81
Observations	3343	3309	3026	3032	2634	2584	2221	3399

Source: BISP Survey, own calculations

Table 4.10 describes, on average, how much *coincidence* or *conflict* of gender norms exists among husbands and wives. This specification is used in further analysis, as research utilizes this specification rather than wives' and husbands' self-reported responses on gender role beliefs questions from the survey. As this indicator is derived from the binary answers from husband and wives gender norms, there are two categories of this indicator. One represents *'coincidence'* of gender norms coded as '1,' i.e., when husband and wife both have similar responses on a given indicator, either *'agree'* to the statement or *'disagree.'* On the other hand, the second category represents *'conflict'* of gender norms coded as '0,' i.e., when husbands' and wives' responses are not similar. It indicates husband has traditional views on a specific indicator and the wife has non-traditional opinions and vice versa. The average values from Table 4.10 show that in most of the indicators, there is a *'coincidence'* of husband and wife gender norms except *'women can work outside home'* where it is found that husband and wife are less likely to share their beliefs on this indicator as on average 50-57 percent couples have the coincidence of norms which is comparatively lower than coincidence on other indicators. The previous two tables compared the wives' and husbands' norms separately, where it was observed that wives have non-traditional views on *'women can work outside home'* than their husbands. Therefore, there is a reason to believe that this conflict is more likely because wives are more non-traditional and husbands are traditional.

Table 4.10 The coincidence of Husbands' and wives' gender role beliefs

	2011 Control	2011 Treated	2013 Control	2013 Treated	2014 Control	2014 Treated	2016 Control	2016 Treated
Both men and women should make HH decisions								
	0.73	0.75	0.72	0.73	0.70	0.72	0.63	0.67
Husband should help in HH								
	0.64	0.62	0.67	0.63	0.61	0.61	0.66	0.64
Women can work outside home								
	0.55	0.56	0.52	0.49	0.49	0.51	0.55	0.57
Wife should express opinion								
	0.67	0.63	0.72	0.68	0.68	0.68	0.77	0.76
Women should not tolerate violence								
	0.65	0.69	0.68	0.70	0.65	0.71	0.77	0.77
Better to send daughter to school rather than son								
	0.69	0.63	0.68	0.66	0.67	0.64	0.69	0.70
Observations	2570	2570	2177	2144	1839	1817	1695	2512

Source: BISP Survey, own calculations

Although the discussion and analyses in this research revolve around this new indicator, however, it is also pertinent to analyze additional women empowerment indicators from the survey, i.e., women's decision making, women's mobility, and access to money, to compare the results of non-survey and survey indicators of gender norms and women empowerment.

For women's decision-making, the survey asks the question from all ever-married women *"if you wanted to, would you be able to participate in these decisions?"* for which they are given options either *'yes'* or *'no.* Responses are binary where '0' represents *'no'* whereas '1' represents '*yes.'*

- Have another child
- Children's education
- Children's marriage plans
- Use of family planning methods
- Visit friends
- Minor household purchases such as food or other daily items
- What kind of job you will do (or tasks, if you don't work away from home)
- Lending or borrowing
- Small Investment (setting up a small business, buying some livestock)

- If you/ a child has a serious health problem, what to do about it (e.g., consult with someone (doctor, nurse, pharmacist, Lady Health Worker, traditional healer, hakeem, etc
- Whether you can participate in a group outside of your home (list examples)
- Vote in elections

Table 2 in the appendix presents the summary of these indicators. It shows that, on average, women's decision-making in the household has improved over the years, indicating that cash transfer enhances women's say in the household. However, the comparison of these indicators shows that women have relatively less power in making decisions that involve matters outside the home, for example, household purchases, lending or borrowing, what kind of job they can do, small investment, and participation in the group outside the home. It explains that even if women's say has improved, they are still not enough emancipated to make decisions outside the home.

Concerning how much freedom a woman has in terms of their mobility alone, they are asked, *"Are you usually permitted to go to the following places on your own, only if someone accompanies you or not at all?"* for which they answer either *'alone,' 'not alone,'* and *'never.* For the following items on mobility, *'not alone'* and *'never'* are combined and coded as '0', whereas *'alone'* is coded as '1'.

- To the local market to buy things
- To a local health facility or doctor
- To homes of friends in the neighborhood
- To a nearby shrine or mosque

Finally, there is another question asked from the women about their voting rights. Women are asked, *"When there is a local bodies / provincial/ national assembly election do you always vote, sometimes or never?"* for which woman answers either '*always votes,' 'sometimes votes,' 'never vote,' and 'too young to vote.'* For this variable, 'never vote' responses are coded as '0' whereas '*always votes'* and *'sometimes votes'* are combined and coded as '1'. Table 3 in the appendix summarizes these responses concerning the treatment status of women. It is noted that women have gained freedom of mobility over time. On average, beneficiary women have lower permission in contrast to non-beneficiaries except in 2016.

Regarding women's mobility to the local market, it is observed that from 2011 to 2016, on average, non-beneficiary women have gained permission from 30% to 37%. However, for beneficiary women, it has improved from 26% to 40% during this time. For women's mobility to the health facility, friend's home, and religious place, it is also noted that in 2016 beneficiary women were more independent in their mobility alone. However, in 2016 beneficiary women had the lowest autonomy in their mobility to religious places where only 29% of women are reported to have the freedom to go to religious places alone as compared to their mobility to the local market (40%), to health facility (40%), and friend's home (58%). Nevertheless, in comparison to baseline, 2011, these percentages do not show a considerable difference. The only exception is women's participation in casting their votes in elections which rose from 53% in 2011 to 89% in 2016 among beneficiary women and 58% to 75% among non-beneficiary women. However, this substantial change is because women must possess Citizen National Identity Card (CNIC) to receive the BISP cash transfer. CNIC is also required to cast votes in local bodies/ provincial and national assembly elections. Therefore, one cannot certainly interpret a rise in women's voting in elections as progress towards women empowerment because it may be the case that women themselves feel encouraged to vote since they possess CNIC, which they must have to get the cash transfer.

In addition, the BISP survey measures women's empowerment through women's access to cash. These questions are asked from women *"if you needed to, could you personally gain access to PKR 50, PKR 100, PKR 200, PKR 400, PKR 600, PKR 800, and PKR 1000 amounts of money quickly, for example, in an emergency?"* for which women are given options either *'yes'* or *'no.'* The way this question is framed in the survey indicates that it is more of a question related to the household poverty level. This question does not ask women that whether they have their own money for these specific amounts. Still, this question captures whether a household can afford this emergency expense or not. So, these questions are more like indicators of the household poverty level, not essentially indicators of women's empowerment. This observation can be confirmed with the average responses to these questions. A summary of these indicators is given in Appendix Table 4. It shows that as the amount of money gets higher, women are found to have, on average, less access to those amounts of money, which is an indication that a poor household may have access to lower amounts of money, i.e., Rs.

50, Rs. 100 but not the higher amounts. Therefore, it is considered in this research to use these 'access to money' indicators as part of the control variables.

4.4.3 Gender Ratio

To examine gender ratio, we use the four provincial rounds of PSLM 2005-06, 2007-08, 2011-12, and 2013-14. This variable is measured following the method used by Conover, Khamis, and Pearlman (2017). First, 10 categories of age cohorts are formed from 0-99 years of age, each cohort forms 10-years age groups, for example, the first cohort of age includes people from 0-9 years of age, the second cohort of age includes individuals from 10-19 years of age and so on up until the tenth cohort of age includes individuals from 90-99 years of age. Second, these groups of ten age cohorts, district, and year yield 900 combinations. Third, based on these 900 age cohort-district-year combinations, the total number of working-age (10-64 years of age) men and women are calculated. Finally, the following specification is used to calculate the gender ratio.

$$\text{Gender Ratio} = (Total\ Men/Total\ Women)_{cdt}$$

The above ratio is calculated at age cohort 'c,' district 'd' and year 't.' Figure 4.4 shows the average working-age gender ratio in four survey years, and it is seen that it has declined from 1.04 to 0.99 from 2012 to 2014 while there was an increase in gender ratio 1.03 to 1.05 between 2006 and 2008.

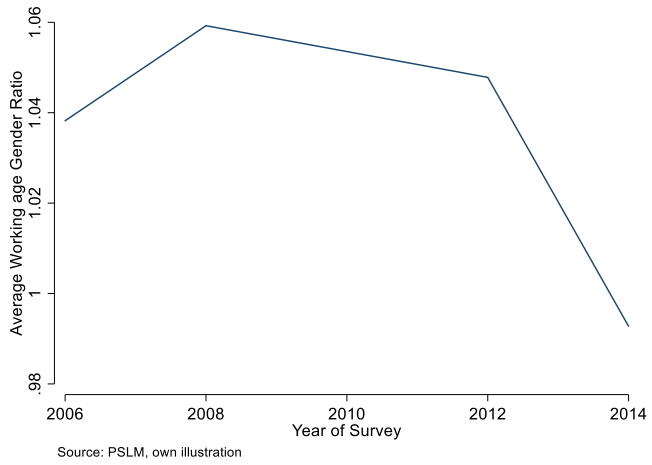


Figure 4.4 Average Working-age Gender Ratio

Figure 4.5 presents the district-level spatial distribution of the average working-age gender ratio in four PSLM survey years. The surveys capture most of the districts on the

eastern side, i.e., from the provinces of Punjab and Sindh. While on the western side of the country, many districts are not surveyed, especially from Balochistan province. The figure shows that most of the districts from Sindh (towards the Indian Ocean in the south) have a consistently high average working-age gender ratio. However, districts from Punjab and KPK have had significant variations in gender ratio over the years.

Moreover, as noted in Figure 4.3, Punjab and Sindh are provinces with a high concentration of extreme weather events such as floods. In the districts where data is available, the maximum working age gender ratio is 1.39, 1.38, 1.35, and 1.30 in 2006, 2008, 2012, and 2014 respectively. On the other hand, the minimum gender ratio is 0.71, 0.58, 0.73, and 0.52 in 2006, 2008, 2012, and 2014, respectively.

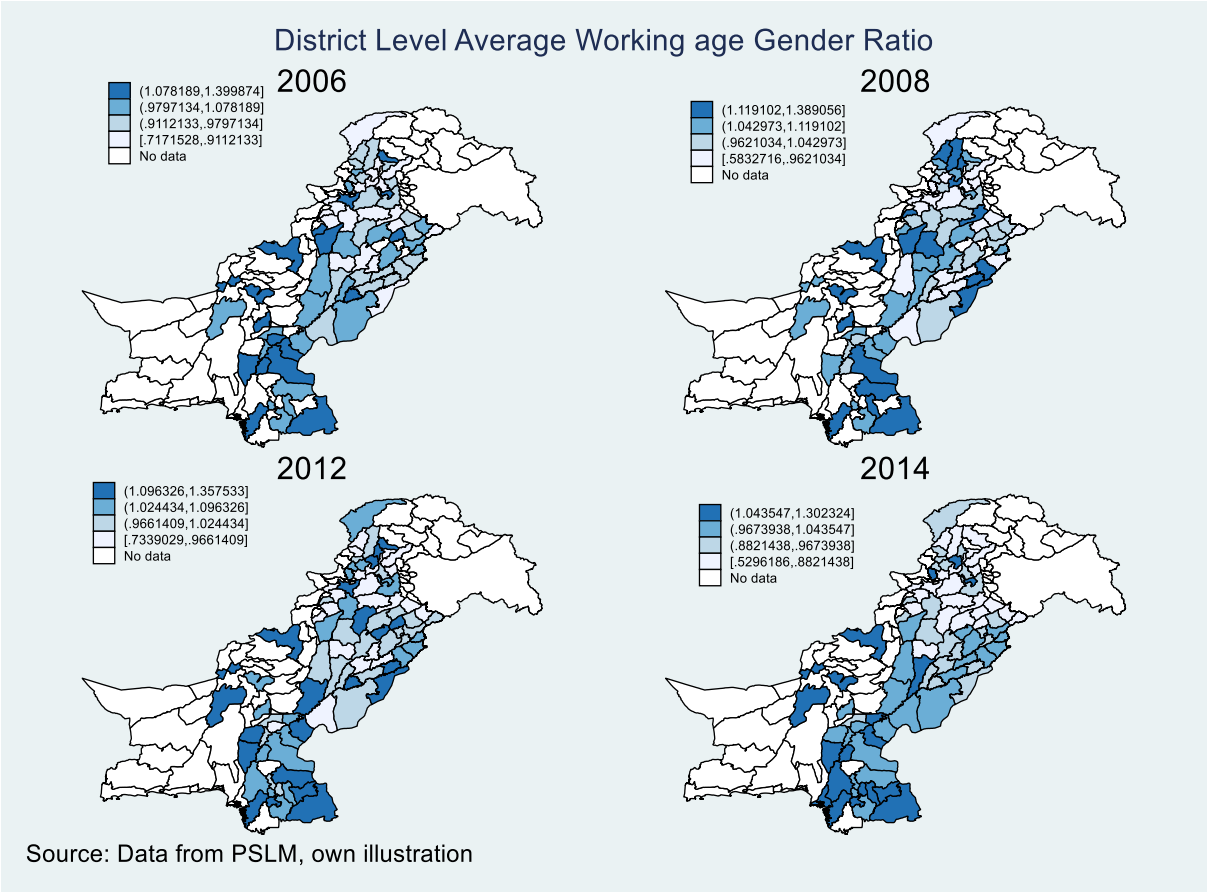


Figure 4.5 District Level Average Working-age Gender Ratio

4.4.4 Exposure to Extreme Weather Events

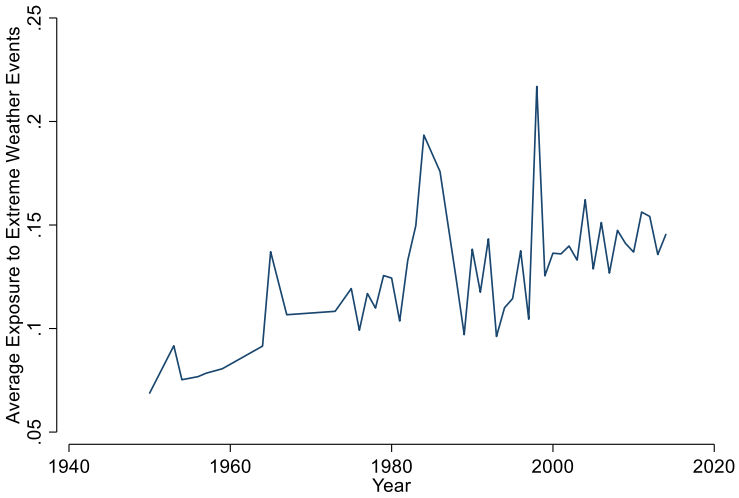
Intergovernmental Panel on Climate Change (IPCC) defines exposure as "presence (location) of people, livelihoods, environmental services and resources, infrastructure, or economic, social, or cultural assets in places that could be adversely affected by physical events and which, thereby, are subject to potential future harm, loss, or damage" (IPCC,

2012 pp. 32). This research concentrates on exposure due mainly to the occurrence of extreme weather events. This variable is constructed from the EM-DAT database and then combined with the PSLM survey data for further analysis. Exposure to extreme weather events is estimated as a birth cohort and district-specific variable to identify how many extreme weather events a person has been exposed to throughout their lifetime. The longer a person lives, the more will be their exposure, hence more likely they will experience out-migration due to weather events exposure. Therefore, this will give an individualized measure of exposure to extreme weather events.

This variable is computed considering a certain age of individuals who can move/migrate to another place. Therefore, firstly, 'moving age' is specified as working age, i.e., 10-64 years. The reason is that since 'exposure' will be used as an instrument to endogenous gender ratio, which is measured by using a total number of working-age men and women, therefore, to make the analysis consistent and comparable, 'exposure' is also calculated by considering 'moving age' the same as 'working age.' Secondly, the total number of extreme weather events and maximum age is calculated based on the year of birth for each district. It identifies how many total events have occurred in a particular district ever since a person was born. Lastly, exposure is calculated by dividing birth-cohort and the district-specific total number of events with individuals' birth-cohort and district-specific maximum age. It is done by using the following expression.

$$\text{Exposure} = (Total\ Extreme\ Weather\ Events / Maximum\ Age)_{ydt}$$

The above expression is measured at 'y' year of birth, 'd' district, and year 't.' The following diagram illustrates the trend in average exposure to extreme weather events from 1950-2014. It clearly illustrates a rising trend in occurrence events. For further analysis about how much exposure has been faced by respondents from the survey datasets of PSLM, the estimated 'exposure' variable is combined with the pooled PSLM data set, which is eventually used to apply TSLS.



Source: Data from EM-DAT, own illustration

Figure 4.6 Average Exposure to Extreme Weather Events 1950-2014

4.4.5 Replicated Poverty Score

The government of Pakistan assigned 85% of households a poverty score based on the Proxy Means Test (PMT) through a census, National Socioeconomic Registry (NSER) in 2011 (Channa, 2012). The eligibility status of each household for BISP cash transfer was determined based on NSER. As discussed earlier in chapter 1, BISP contracted Oxford Policy Management (OPM) to evaluate this program's impact. The fieldwork for the NSER was still underway on a rolling basis (i.e., some districts completed before others) at the time of OPM's baseline survey in 2011. Therefore, to collect "pure baseline data" (i.e., data before beneficiaries started receiving cash transfers), OPM had to identify potential beneficiaries without using data from BISP. To do this, OPM replicated the BISP PMT for purposes of creating a sample of potential beneficiaries. Since the baseline survey, OPM matched the households using CNICs to the BISP data in the NSER to get their PMT scores as collected by BISP.

BSIP poverty score was calculated in 2011 using the PMT model applied on 2005-06 PSLM, and this poverty score has not been updated since then, which makes a downside associated with using PMT for BISP program targeting that also results in the results its non-reproducibility using more recent data on chosen poverty covariates. Jalal (2017), in her study, also claims that BISP's poverty scorecard is an unstable welfare predictor because it determines household eligibility based on point-in-time estimates. For quantitative analysis, this research attempts to replicate the official PMT model of BISP by using the OPM data collected at baseline in 2011. The OPM surveys used in this

research have data on the 23 poverty covariates chosen from PSLM to construct an official poverty score. So, this data availability permits replication of poverty scores using the updated information on these poverty covariates. Hou (2009) reports the official weights used by BISP in its PMT model, and these weights are given in Table 5 in the appendix. In this research, similar weights are used on poverty covariates from more recent rounds of BISP impact evaluation to replicate poverty scores in each year.

Following Figure 4.7 compares the replicated score with official poverty score and replicated score with scorecard generated by OPM. The left panel of figure 4.7 shows that the official score and the replicated score are largely inconsistent. Then the right panel of figure 4.7 compares OPM's generated score with replicated score constructed in this research. The same comparison is presented by using boxplot, Figure 1 in the appendix. The boxplot shows that the distribution of replicated score from this research and OPM's replicated score are precisely the same as shown in the right panel in figure 4.7. These comparisons are in line with findings from Jalal (2017), who also replicated the BISP's PMT model. The author suggests that the difference in official poverty score and the replicated score are due to fieldwork of NSER on a rolling basis, ineffective collection and processing of NSER data, and the chosen poverty covariates are imperfect welfare predictors.

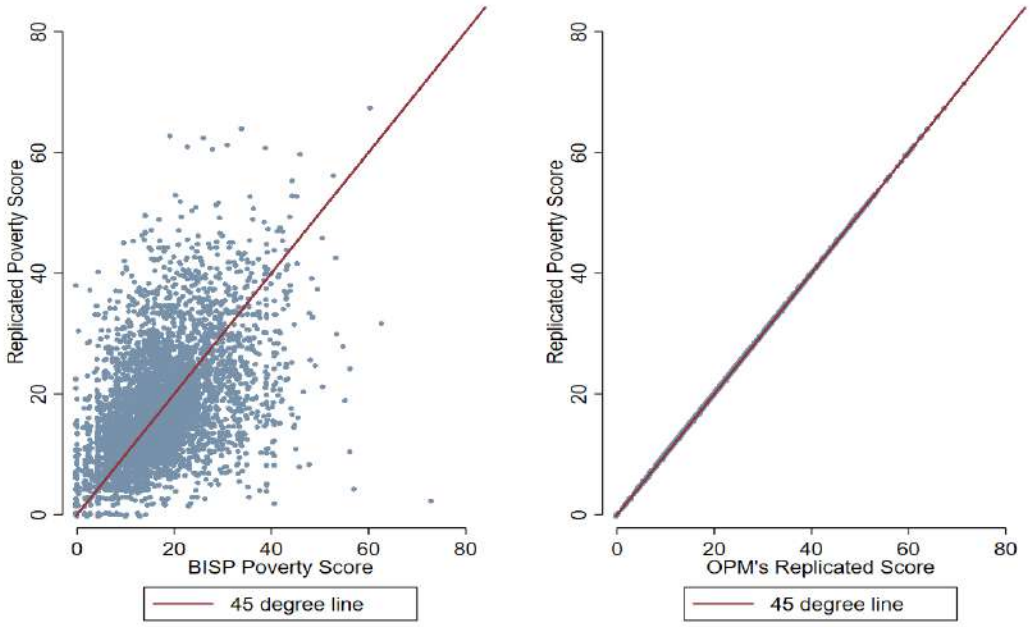


Figure 4.7 Comparison of Official Poverty Score, OPM's Score, and Replicated Poverty Score

As the official BISP PMT model is perfectly replicated using OPM data collected in 2011. The reason for the non-reproducibility of official poverty scores is as described above. Therefore, if the same PMT model is applied to midline data collected by OPM, i.e., 2013, 2014, and 2016, it should calculate the actual poverty score of households using current data on poverty covariates from each year. Hence, rather than using the outdated official poverty score, this research utilizes this replicated score as one of the instruments and official treatment status to predict FLFP, details of which are given in the next chapter.

4.4.6 Placebo Treatment Status

As discussed before, the BISP official poverty scorecard has not been updated since 2011, so the score is replicated in this research to have the actual poverty status of each household. According to the official poverty score, the households eligible for the program in 2011 their eligibility status did not change in 2013, 2014, and 2016. It implies that there are households who are always takers and those who are never takers. Since the score is replicated for each year, it can be easily identified who would have received the cash transfer had they assigned the program eligibility in 2013, 2014, or 2016 rather than 2011. So, we can find those who still get the money in the following years, although they are no longer poor.

Similarly, some households should get the money if the program was not started in 2011 but selected as ineligible in 2011. This way, we can look at those households whose eligibility status has changed over time. These households can be served as a good counterfactual for one another. However, never-takers and always-takers cannot be good counterfactuals because their eligibility status never changed during the program.

With the replicated poverty score and counterfactual households, we construct a placebo treatment status of each household. First, we apply the same cut-off, i.e., 16.17, on replicated poverty score to get the replicated treatment status. Second, we exclude those households officially treated because they are always takers and thus below the threshold in each year. This way, only those households are considered who would have received the cash if they were selected as beneficiaries in the following years, not in 2011. Third, in placebo treatment status, we eliminate all those households who are appeared to be in the control group in other years, i.e., households who would not be treated if the program was decided in a different year. Finally, according to placebo

status, households in the treated group should be poor than those in the control group because the control group would not be under the threshold if it were a different year. Thus, placebo status consists of an officially control group, but they would not be treated based on values of replicated poverty score in other years. This placebo treatment status would eventually be used as a placebo instrumental variable (IV) to lend support to the validity of identification of the official treatment status, which is a true IV. Technically, placebo status should have strictly no effect in explaining FLFP and thus gender norms because it never happened, that is why we call it a placebo treatment status.

4.5 Multiple Correspondence Analysis

There are six qualitative binary indicators of gender role beliefs used in this research. Therefore, we utilize a data reduction technique to summarize these six binary variables into a single composite indicator of gender role belief. A well-known statistical method of Principal Component Analysis (PCA) is widely used in research as a data reduction technique. However, the study says that PCA applies to continuous quantitative variables normally distributed (Kamanou 2005) and assumes a linear relationship between the chosen variables. Therefore, it is said that when dealing with categorical/binary variables, the correct choice is Multiple Correspondence Analysis (MCA). Merola & Baulch (2015) also refer to the unsuitability of PCA to categorical/binary data as the association between categorical variables cannot be measured in terms of covariation and correlation. Moreover, contrary to PCA, which considers the absolute weight of the component, MCA considers their relative frequencies (Tuccio & Wahba, 2018)

The advantage of using MCA instead of PCA is that it considers the binary nature of the variables being used to construct an indicator matrix, an individuals x variables matrix, which is *households x variables* matrix where the rows represent households and columns represent variable categories. Every variable is divided into mutually exclusive dummies, i.e., each category of each variable is divided into a binary indicator of 0/1 (Le Roux and Rouanet, 2004). In other words, each row (household) will have 1 in only one category and otherwise 0. MCA applies standard correspondence analysis on the inner product of this indicator matrix, called Burt Table, to analyze the relationship between several categorical variables (Ferrant, 2014) and extracts dimensions that contain variance of variables used in the indicator matrix. Like PCA components, the

first dimension of MCA captures the maximum variance of all variables followed by other dimensions with successively lower variance.

For the formulation of composite indicator generated with MCA, we consider *k* number of variables (dimensions) with *x = (1, 2, …., X)* where *X=6* is the number of gender norms responses by husbands and wives, *c* the number of categories of each variable with *c=(1, 2, …, Cx).* The gender norms variables are binary; therefore, *c=(0,1)* in our case. As discussed before, each variable category is divided into a dummy variable (I). With 6 binary variables, each with two categories, this gives 12 dummy variables, where the dummy is equal to 1 if the household has a specific category 0 otherwise. Finally, *W* is the weight calculated by MCA which are the coordinates/factor scores of each category of variables normalized by the eigenvalue of the first dimension. Having these descriptions, the functional form of composite indicator or gender norms index for household *i* can be written as

$$(Gender\ Norms\ Index)_i = \frac{1}{X}\sum_{x=1}^{X}\sum_{c_x=1}^{Cx} W_{c_x}^{x}\ I_{c_x\,i}^{x} \qquad 4.1$$

Where $$W_{c_x}^{x} = \frac{Coordinates^{x}}{\sqrt{\lambda_1}} \qquad 4.2$$

Asselin (2009) proposes the monotonicity axiom that must be fulfilled while constructing a composite K categorical/binary variables indicator. Monotonicity axiom requires two conditions 1) First Axis Ordering Consistency (FAOC-I) for an indicator I. It states that variables should have ordinal consistency in their categories and ordering of weights for each category, either decreasing or increasing. In the context of this research, this implies that the first axis (dimension) must have increasing scores showing a movement from traditional to non-traditional gender norms in the household, i.e., if a household has more non-traditional gender norms, then the composite indicator increases. 2) Global First Axis Ordering Consistency (FAOC-G) is about the exact orientation of FAOC-I for all variables, either increasing or decreasing. In the case of binary variables, FAOC-I is always fulfilled (Ezzari and Verne, 2012); however, FAOC-G requires that all indicators have the same orientation for the first dimension. MCA also calculates contribution for each category of each variable equal to the variance of each variable corresponding to each dimension (Franco, 2016), i.e., and it measures the strength with which each category of each variable is associated with dimensions.

Weights generated by MCA can have negative and positive values. Therefore, the variables that do not obey FAOC-I can be dropped from the analysis. There are other criteria to select the variables used to construct a composite indicator: the variance of the first dimension explained by the variable. Moreover, estimated coordinates can be negated by multiplying the coordinates by -1 to provide a more intuitive reading of the scale (Seplaki et al., 2014), i.e., to get the scores increasing with the increase in non-traditional norms in the household. So larger values of scores from the first dimension indicate having more non-traditional gender norms in the household. However, for a new construct of husband-wife gender norms coincidence, the rescaling of score would mean that husband and wife are more likely to share similar gender norms with the increasing score. To simplify the understanding, the resulting weights can be transformed by subtracting the minimum weight from each weight of the household to make the gender norms index greater than equal to zero.

Another decision needs to be made to create an index using MCA, particularly in panel data. The researcher must decide whether to calculate the index using each cross-section sample independently or pooled sample over all the available survey years. There are trade-offs in using pooled and cross-section data to compute MCA. As discussed in (Kohn 2012, p. 230), "there is a need to balance the pros and cons of pooled or cross-sectional computation of index based on specific feature of data and the nature of research question." For this research, we choose to compute the index on pooled data. First, the effect of cash transfer on women's LFP and gender role beliefs may change over time. One may expect that a change in women's LFP in one year because of being a recipient of cash transfer may also influence the women's LFP decision in the following years which may also influence gender norms over time. Second, as MCA maximizes the correlation of responses within the sample, therefore, to avoid the loss of precision in the correlation, the pooled sample is used to compute the index because of the considerably large sample size.

There are two steps involved to carry out the MCA analysis. First, we use all the six variables listed in tables 4.8 and 4.9 for husbands' and wives' gender norms index, respectively, and table 4.10 for an index of husband-wife gender norms coincidence. Initially, MCA is applied to these variables to find the pattern of association of each category of these variables with the extracted dimensions. Second, we remove variables that do not meet the basic MCA requirements discussed above. Table 4.11 presents the

preliminary and final MCA eigenvalues and inertia for husband-wife gender norms coincidence on a panel sample of four years from 2011 to 2016. The preliminary eigenvalues and inertia show that the first dimension explains 22.8% of total inertia, which drops to 17.5% in the second dimension, followed by successively lower inertia for the remaining dimensions. Table 4.12 shows the coordinates and contributions of each category of each variable for the first two dimensions. It is observed that two variables, *'both men and women should make household decisions'* and *'women should not tolerate violence,'* do not follow the FAOC-I in the first dimension because of having negative weights of the higher-order category of *'coincidence'* and they have a meager contribution to the first dimension. Another variable, *'better to send daughter to school rather than son,'* also has low contribution but is retained following FAOC-1. However, the other three variables follow FAOC-I, and they also significantly contribute to the first axis. Therefore, to make the model satisfy MCA's primary conditions, we need to eliminate the variables that are inconsistent with FAOC-I.

Table 4.11 MCA Eigenvalues and Inertia of Husband-wife gender norms coincidence 2011-2016 Panel Sample

	Preliminary Eigenvalues	Preliminary Inertia	Final Eigenvalues	Final Inertia
Dimension 1	0.23	22.83	0.33	33.56
Dimension 2	0.17	17.51	0.25	25.05
Dimension 3	0.17	16.88	0.21	21.18
Dimension 4	0.15	15.30	0.20	20.20
Dimension 5	0.14	14.07	0.00	0.00
Dimension 6	0.13	13.40	0.00	0.00

Table 4.11 also reports the final eigenvalues and inertia. After dropping the two variables mentioned above, the MCA is applied to the remaining four variables, and the explanatory power of the first dimension has increased from 22.8% to 33.5%. However, the final coordinates of these four variables in the first dimension have opposite signs, as shown in Table 4.13. Although they follow FAOC-G, i.e., all indicators have the same direction for the first axis, the coordinates have opposite signs with a higher-order category of *'coincidence.'* This issue can be resolved by flipping the sign of the weights of the first dimension to have a more intuitive explanation of the first dimension as suggested by Seplaki et al. (2014). It concludes that the household composite indicator computed by the first dimension of MCA is a consistent index that can be used for further analysis.

Table 4.12 Preliminary MCA Dimensions and Contributions of Husband-wife gender norms coincidence 2011-2016 Panel Sample

	Coordinates Dimension 1	Contributions Dimension 1	Coordinates Dimension 2	Contributions Dimension 2
Both men and women should make HH decisions				
Conflict	0.93	0.02	2.32	0.11
Coincidence	-0.40	0.01	-1.00	0.05
Husband should help in HH				
Conflict	-1.83	0.01	0.86	0.01
Coincidence	1.02	0.05	-0.4	0.02
Women can work outside home				
Conflict	-1.36	0.07	0.88	0.02
Coincidence	1.20	0.06	-0.77	0.02
Wife should express opinion				
Conflict	-2.04	0.01	-0.05	0.00
Coincidence	0.86	0.04	0.02	0.00
Women should not tolerate violence				
Conflict	0.08	0.00	0.37	0.003
Coincidence	-0.03	0.00	-0.15	0.001
Better to send daughter to school rather than son				
Conflict	-0.83	0.01	-2.25	0.12
Coincidence	0.41	0.01	1.12	0.06

Table 4.13 Final MCA Dimensions and Contributions of Husband-wife gender norms coincidence 2011-2016 Panel Sample

	Coordinates Dimension 1	Contributions Dimension 1	Coordinates Dimension 2	Contributions Dimension 2
Husband should help in HH				
Conflict	1.56	0.12	0.58	0.01
Coincidence	-0.87	0.07	-0.32	0.01
Women can work outside home				
Conflict	1.18	0.09	0.59	0.02
Coincidence	-1.05	0.08	-0.52	0.01
Wife should express opinion				
Conflict	1.70	0.12	-0.50	0.01
Coincidence	-0.72	0.05	0.21	0.004
Better to send daughter to school rather than son				
Conflict	0.53	0.01	-2.61	0.28
Coincidence	-0.29	0.01	1.30	0.14

The same conclusion is drawn when MCA is applied on husband-wife gender norms coincidence from a panel sample of three years from 2011 to 2014. The results of preliminary and final eigenvalues are reported in Appendix Table 6. The table shows that when all the variables are included in MCA analysis, the preliminary results show that the first dimensions explain 23.5% of total inertia, which improves to 34.4% after dropping two variables *'both men and women should make household decisions'* and

'women should not tolerate violence' that do not follow the FAOC-I in the first dimension as shown in Table 7 in the appendix.

MCA allows graphical representation of the contribution of each category of the variable on the first two dimensions. Graphic visualization of categories helps in identifying variables for which the respondents give similar responses. Figures 4.8 and 4.9 display the preliminary and final coordinates of each category of variables for their contribution to both dimensions. Figures also show that variable categories for which respondents show similar behavior are grouped. It is evident from Figure 4.8 that the categories of *'coincidence'* of three variables, namely *'women can work outside home,' 'husband should help in HH,'* and *'wife should express opinion,'* have a positive contribution to the first dimension and the distance between these categories are small which indicate that husbands and wives share the gender norms on these indicators. Whereas for *'sending daughter to school rather than son,'* the coincidence of gender norms is relatively high, it happens to be correlated with the second dimension.

Similarly, for *'both men and women should make HH decision,'* the category *'conflict'* significantly contributes to the second dimension, which shows that husbands and wives have conflicting gender norms on this indicator. This category is well represented on the factor map as the category is positioned relatively far from the origin. For *'women should not tolerate violence,'* both the categories of this indicator are closer to the origin, which shows almost no correlation with the first and second dimensions. Respondents seem to judge this question about domestic violence very differently than the other indicators.

The MCA plot in table 4.9 shows the final coordinates. As discussed previously, the signs of the coordinates are flipped to make them increase with the increase in husband-wife gender norms coincidence. Therefore, in the MCA plot, the horizontal axis (first dimension) is negated. Again, it is observed that categories of *'coincidence'* on the three variables *'women can work outside home,' 'husband should help in HH,'* and *'wife should express opinion'* are closely positioned on the positive quadrant of the first dimension. One can interpret that the first dimension measures gendered divisions of labor in the household.

The point of MCA is to identify latent dimensions underlying the distribution of answers to these questions. If there is a lack of clear drop in inertia between the axes, there is a lack of a straightforward underlying structure. It suggests that variables used in MCA analysis measure slightly different matters, which happens to be in the case of six gender norms questions from the BISP survey on which MCA is utilized in this research. Test items have weak internal consistency because of the low Cronbach's alpha (0.27) reported in Table 8 in the appendix. It also indicates that items do not have high shared covariance and measure a different underlying concept. From Table 4.11, it was seen that inertia of each subsequent dimension does not differ a lot compared to the former one. This pattern is also seen in the MCA coordinates plot in Figures 4.8 and 4.9, categories of three out of six variables appear to be grouped, i.e., *'coincidence'* categories are grouped on positive quadrants of the first dimension whereas *'conflict'* categories are positioned on negative quadrants.

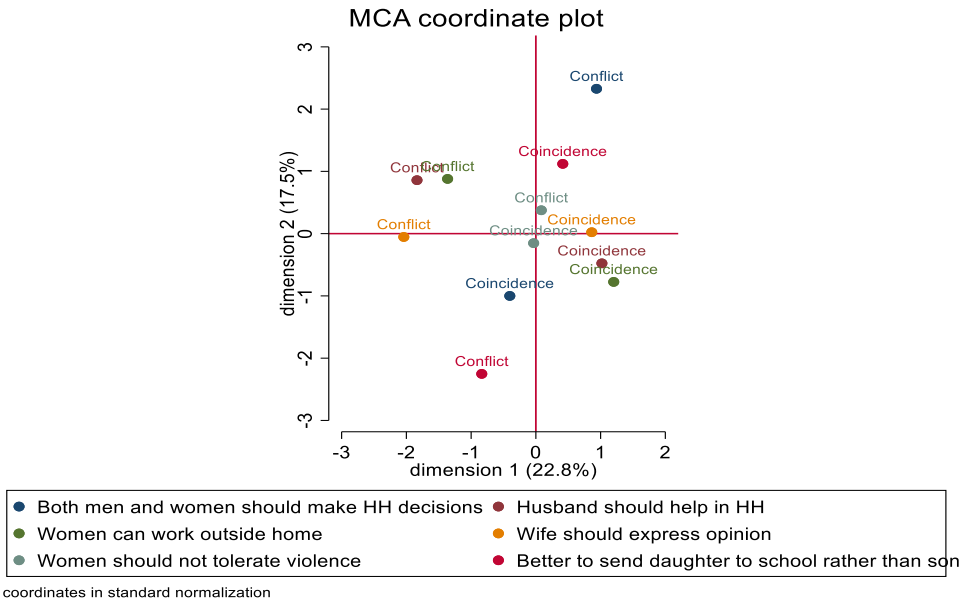


Figure 4.8 Preliminary MCA coordinates plot of Husband-wife gender norms coincidence

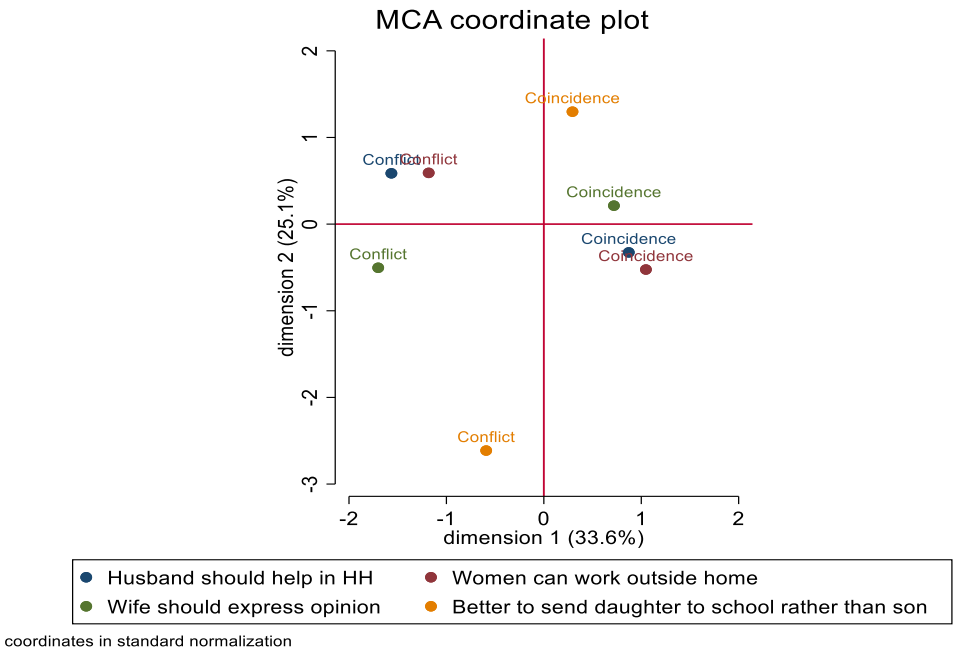


Figure 4.9 Final MCA coordinates plot of Husband-wife gender norms coincidence

One should also realize that six binary variables are a few to construct a scale through data reduction techniques, which one typically uses when many indicators are used to construct an index. However, in the context of this research, the objective is to construct an indicator that illustrates the responses of both husband and wife in terms of coincidence or conflict of gender norms perception among them. The BISP data only possess men and women responses on these six indicators; however, for the rest of the other dimensions of women empowerment, the indicators are only responded by women, which do not meet the criteria to be included in further analysis of this research. Therefore, despite this limitation, analysis is carried out with a limited number of gender norms indicators.

Tables 9 and 10 in the appendix show results of MCA on men's gender norms indicators for the panel sample from 2011-2016. This composite indicator and its interaction with FLFP are used as controls to predict the effect of norms-induced FLFP on husband-wife gender norms coincidence. Table 9 in the appendix shows the preliminary and final inertia with all the variables included, 29% inertia explained by the first dimension, which improved to 54.9 when weakly correlated indicators are dropped from the analysis. As shown in Table 10 in the appendix, the signs of coordinates are contrary to the higher-order category, and they are consistent for each variable category. Therefore, to make the composite indicator increase with the increase in a coincidence of gender

norms among men and women, the signs of the coordinates are flipped to make the explanation more intuitive.

A similar analysis is carried out on a three-year panel sample, and the results of MCA coordinates and contributions are presented in Table 11 in the appendix. MCA is also applied on other women empowerment domains, namely, women's mobility and women's decision making, the results of which are presented in Table 12 and Table 13 in the appendix for women's mobility, and Table 14 and Table 15 for women's decision making in the household for the 4-year panel sample. The MCA results of these domains for the 3-year panel sample are presented in Table 16 and Table 17 for women's mobility and Table 18 and Table 19 for women's decision-making. However, the index of women's access to money is used as one of the control variables in the further analysis because, as mentioned above, these indicators seem to primarily capture the poverty status of the household rather than women's emancipation. The MCA results on women's access to money for the 4- year and 3-year panel samples are presented in the appendix from Table 20 to Table 23. Finally, the summaries of the control variables used in the two studies are presented in Appendix Table 24 from PSLM for study 1 and Table 25 from BISP for study 2.

4.6 Summary

This chapter discussed the datasets used in this research, their sources, and comprehensively stated operationalization of key variables. It starts with the summary of two main data sets, i.e., PSLM and BISP. PSLM is a nationally representative cross-sectional household survey in Pakistan initiated explicitly by the Government of Pakistan to evaluate the progress of United Nations' MDGs in Pakistan. Four provincial rounds of 2005-6, 2007-08, 2011-12, and 2013-14 are used in this research. PSLM and emergency events dataset, EM-DAT, are used to investigate the first research hypothesis, i.e., the causal effect of EGRB on women's LFP. EM-DAT data provided information on the occurrence of extreme weather events at the district level. A wide variety of extreme weather events are considered in this data; however, I utilize data on flood, drought, extreme temperature, landslide, and storm for this research.

Moreover, this data is employed in the first study to operationalize the exposure of extreme weather events individuals face during their lifetime. The variable is operationalized by merging the PSLM survey data with EM-DAT to determine

individual exposure to extreme weather events. This variable serves as an instrument to clean the endogeneity of gender ratio to determine the causality between EGRB and women's LFP.

BISP data set is utilized to examine the second research hypothesis of study 2, i.e., the causal effect of women's LFP on equitable gender norms. BISP is the Government of Pakistan's flagship conditional cash transfer program, initiated in 2011 to assist the poorest of the poor by providing them cash assistance. One of the program's primary objectives is women's empowerment, and the money is particularly disbursed to women in the household. This research uses four rounds of BISP 2011, 2013, 2014, and 2016. The cash assistance is provided based on the poverty score calculated using the Proxy Means Test (PMT) method from the scale of 0 to 100. A poverty score of 16.17 is chosen as a threshold to determine the treatment and control groups. Households with poverty scores less than or equal to 16.17 are considered treated, and households above 16.17 are control households. BISP has not updated the poverty score since baseline 2011. Therefore, the treatment status of the households has not changed since then. It may cause the empirical analysis inconsistent because the poverty status of households may have changed during the time because of being cash recipients. Since this critical aspect is not considered, this research replicates the BISP poverty score by utilizing the same PMT method used by BISP to compute the official poverty score. The replicated poverty score and official treatment status of households will be used as an instrument to determine the causal impact of EGRB on women's LFP.

This research uses the six indicators of gender role beliefs of husband and wife to determine the coincidence and conflict between husband and wife on their perceptions of gender role beliefs. This unique specification will help avoid the problem of social desirability bias arising from self-reported responses to survey-based questions. Moreover, a data reduction technique, multiple correspondence analysis (MCA), is applied to the characteristics of husband-wife gender norms. The chapter provides a methodological explanation of the construction of this composite indicator and highlights the limitations of using MCA on a limited number of binary indicators.

CHAPTER 5 RESULTS AND EMPIRICAL FINDINGS

5.1 Introduction

This chapter presents the empirical findings obtained from analyzing the research hypothesis developed in chapter 1. Sub-section 5.2 of this chapter presents TSLS estimation results on the pooled cross-section of PSLM to examine the causal effect of EGRB – measured through district-level variation in gender ratio – on women's LFP decisions. For analysis, this research considers statistical inference in TSLS regression models where observations are grouped into clusters, with model errors uncorrelated across clusters but correlated within-cluster (Cameron and Miller, 2011). One leading example of "clustered errors" is individual-level cross-section data clustering on geographical regions, such as villages or states. Then model errors for individuals in the same region may be correlated, while model errors for individuals in the different areas are assumed to be uncorrelated. Although the model is estimated at the individual level, the variation in the instrumental variable, i.e., 'exposure,' is at the district level. Therefore, variation in gender ratio is also calculated at the district level. Thus, using robust standard errors clustered at the district level is reasonable because of district-level variations. Standard errors when clustered are much larger than robust standard errors; however, the coefficients are the same. Clustered standard errors are a special kind of robust standard errors that account for heteroskedasticity across "clusters" of observations (Cameron and Miller, 2011).

Sub-section 5.3 describes the empirical results of study 2. This study uses the panel random effect IV estimation method on BISP cash transfer data to investigate the causal effect of EGRB on women's LFP. In addition to employing TSLS, the research plans to apply regression discontinuity design to substantiate the presence of a discontinuity in women's LFP around the poverty score threshold. The RDD results and their validity is presented in the appendix. Finally, section 5.4 summarizes the empirical findings of the two studies.

5.2 Study 1: equitable gender role beliefs and women' LFP

As argued earlier about the endogeneity of gender ratio, the primary variable of interest, Table 5.1 presents the point estimates on this variable with OLS and IV. The complete estimation results with individual and household level controls are shown in Table 26 in

the appendix. It is always expected to get unbiased point estimates, i.e., $\hat{\beta}$ to get the accurate statistical inference. However, another crucial element for correct statistical inference is to get correct standard errors. The presence of heteroskedasticity, i.e., when error variance is not constant, causes standard errors to be biased. OLS assumes that errors are independent and identically distributed; robust standard errors relax either or both assumptions. Therefore, robust standard errors also referred to as a sandwich standard error, an Eicker-White standard error, a Huber standard error, a heteroscedasticity consistent standard error, address the problem of errors that are not independent and identically distributed. The OLS and IV results are analyzed with robust and with standard errors clustered at the district level. Since we are instrumenting gender ratio with exposure to extreme weather events constructed at the district level, endogenous variables and IV are constructed at the district level to have consistency in analysis. Thus, we present the results with standard errors clustered at the district level and robust standard errors.

Table 5.1 OLS and IV regressions on FLFP with Exposure to extreme weather events as IV

	Robust SEs		Clustered SEs	
VARIABLES	OLS	IV-II	OLS	IV-II
Working age Gender Ratio	-0.05***	0.37***	-0.05	0.37
	(0.01)	(0.09)	(0.04)	(0.26)
Constant	0.45***	-0.15	0.45***	-0.15
	(0.03)	(0.13)	(0.09)	(0.36)
Observations	43,879	43,879	43,879	43,879
R-squared	0.12	0.08	0.12	0.08

Regressions include control variables, age, age square, education, marital status, head education, head labor market status, dependency ratio, household size, number of kids, number of infants, family structure, female-headed households, log of monthly income, employment decision and permission to work. *** p<0.01, ** p<0.05, * p<0.1

5.2.1 OLS and Second Stage Results

When interpreting the OLS coefficient of endogenous variable, i.e., working-age gender ratio, in column 1, it is seen that the coefficient is negative and statistically significant at 1% level while controlling for individual and household level characteristics. It shows that with robust standard errors, i.e., assuming independence across individuals, a 1 percent increase in gender ratio will reduce the women's labor force participation by 5 percent and vice versa. This relationship is clearly in line with Becker's marriage market hypothesis. Becker's (1973) intra-household decision-making models argue that marriage market conditions determine the distribution of intra-household utility. Becker

presented the demand and supply model to illustrate the effect of changing sex ratio on the marriage market and, ultimately, women's participation in the labor market. It is argued that when one gender is relatively scarce compared to the other, its bargaining power increases. For instance, when there is a male-biased sex ratio, female bargaining power increases and is less likely to participate in the labor market. The magnitude of the effect of gender ratio on FLFP is the same, i.e., 5%, irrespective of the nature of SEs. However, the standard errors are higher when clustered, making the results statistically insignificant. However, OLS results are inconsistent due to the endogenous gender ratio in the model, thereby, do not the present causal effect of EGRB on women's LFP.

In the second stage of the TSLS IV method, the endogenous variable, part of which is not correlated with the error term, is used to investigate its impact on the outcome variable. The results of the second stage, obtained with robust SEs and robust SEs clustered at the district level, are reported in Table 5.1. The second stage results are strikingly different from OLS. The point estimates and signs of OLS and IV-II are the same with robust and clustered SEs except for the standard errors, which are large when clustered at the district level and make the point estimates statistically insignificant. With robust SEs, the coefficient on gender ratio is positive and statistically significant at 1% level, but insignificant with clustered SEs. The results, with robust SEs, show that a 1% increase in gender ratio leads to a 37% increase in women's LFP. The results are contrary to what OLS predicted, thereby confirming bias in the OLS estimates in the presence of endogeneity. Therefore, once the model is free of endogeneity, the statistical inference with TSLS differs substantially from OLS. The conflict in results also indicates that the policy recommendations based on OLS estimates can be misleading once the endogeneity in the regression model is not corrected.

With SEs clustered at the district level, the SEs are large, which makes the results insignificant. The model is estimated at the individual level, but the instrument 'exposure' is constructed at the district level. Because of this district-level variation in exposure, the research uses clustered SEs calculated too conservatively. Therefore, this excessive clustering of SEs makes the results insignificant. Moreover, variation in gender ratio is also estimated at the district level to operationalize the endogenous variable at the same level as the instrument. However, this is not the proper level of clustering. We are mainly interested in the individual-level analysis of variation in

gender ratio and its impact on women's LFP. However, even if SEs are clustered, the coefficient 0.37 for gender ratio is nearly closer to be significant at 10% level. Since the t-value for the 10% significance level is supposed to be 1.64, with IV result, the t-value associated with the coefficient of gender ratio is 1.42. Hence, the first study's IV results could be considered reliable and consistent even when insignificant.

The interpretation of the IV result suggests that an increase in gender ratio improves the labor market working conditions for women, which is reflected in higher FLFP in response to exogenous variation in gender ratio. Our results are also contrary to the available literature on unbalanced sex ratio and its effect on FLFP. Several studies have looked at the causal relationship between sex ratio and FLFP; among them are (Grosjean and Khattar, 2019; Cardoso and Morin, 2018; Teso, 2014; Fernández et al., 2004; Angrist, 2002). These studies provide evidence of the negative causal impact of an unbalanced sex ratio on women's participation in the labor market. Moreover, our results were also in contrast with Becker's marriage market hypothesis, which states that in the case of male-biased sex ratios, women will have more bargaining power and prefer to enjoy more leisure, which reduces their labor market participation.

Our results can be interpreted so that even if there is a male-biased sex ratio, women still increase their labor market participation. One explanation of this relationship could be that extreme weather events especially, floods and extreme temperatures destroy household livelihoods. Therefore, to mitigate the effect of income loss, women and men will increasingly take the responsibility of providing income to the household. They may be compelled to work in the fields, sell homemade production, or find a small job to meet the households' financial needs. The first stage result in Table 5.2 shows that an increase in 'exposure' leads to a higher gender ratio, i.e., fewer women than men. This result reveals women relocating themselves as compared to men. This exogenous variation in gender ratio leads to higher women's LFP. Therefore, it can be said that more exposure to extreme weather events, which results in a higher gender ratio, will make the prevailing strict social norms undesirable, which is reflected in higher women's LFP.

There is a general understanding that more males out-migration in Pakistan due to weather events (Mueller et al., 2014). Still, our results are contrary to this understanding

and present a different perspective related to weather-induced migration. The results from available literature are also in line with the gender role perceptions in Pakistan. Due to collective decision-making in households, men are dominant in providing the household's essential livelihoods, and women are generally responsible for household chores and child-rearing. However, once valid causal identification methods are taken into account in the analysis, as in the case of our research, the effect of existing gender norms and gender role perception appears to be different from what is proven from the existing literature. Our results show that when the appropriately designed causal identification procedure is applied, and the potential endogeneity is removed, the results are against the available evidence. This research has attempted to resolve the endogeneity of the potentially endogenous variable, gender ratio, and found that a higher gender ratio leads to higher FLFP. These results indicate that existing patriarchal social norms become increasingly costly, and women's LFP increases even if more men than women are. Hence, we can safely interpret our results to present the causal effect of EGRB on women's LFP, i.e., gender role beliefs may evolve or progress when there are circumstances when practicing those norms become unfavorable for society. That is what our results suggest that exogenous variation in gender ratio leads to higher women's LFP. However, the result is statistically insignificant due to the clustering of SEs at the district level. Still, as seen from the first stage results, the first stage produces significant results even if the SEs are clustered at the district level. Therefore, we can safely argue that our second stage results can be taken as reliable statistical inferences.

5.2.2 IV Validity

The point estimate of 'exposure' in the first stage results in Table 5.2 indicates that exposure to extreme weather is positive and statistically significant at a 1% level in both specifications. Although SEs are large due to clustering, the estimated coefficient is statistically significant with clustered SEs. The coefficient of 0.16 indicates that a 1% increase in exposure to extreme weather events leads to a 16% increase in gender ratio in a district, meaning that the higher the exposure to extreme weather events, the more the gender ratio, i.e., male to female ratio. It implies that there are more men relative to women in those districts, and women are found to be less in number than men in response to higher exposure to weather events, which explains the phenomenon of female migration from one district to another rather than male out-migration. This result is different than what was found in the literature on climate-induced migration in

Pakistan. In their longitudinal study, Mueller et al. (2014) found that men are more likely to move out of the villages than women due to extreme temperatures. This study employs data on temperature, flood, and rainfall from five districts, Attock, Badin, Dir, Kalat, and Faisalabad, from 1991 to 2011. From each of these districts, they randomly chose villages based on their distance from the market. This research employs data from 95 districts across the country and looks at the frequency of extreme weather events from 1950 until 2014. Moreover, our analysis considers a variety of weather events like drought, extreme temperature, flood, landslide, and storms. Our results seem to be more representative in terms of precision and generalizability, having a long time series, districts across the whole country, and different extreme weather events.

The study mentioned above found a significant impact of temperature on migration than rainfall and flood and that migration is mainly male-dominated. Our results are different because we found more 'exposure' to increase the gender ratio, which implies that migration is not necessarily male-dominated; the situation is the other way round. If the gender ratio increases in response to more exposure to weather events, the number of women is lower than men, revealing that women are more likely to migrate than men. Moreover, the study also found that rainfall, flood, or moisture is associated with retention (as opposed to migration), which means people do not migrate if there is a flood or heavy rainfall. However, the EM-DATA database found that the frequency of flood occurrence is the highest during 1950 to 2014 among all other extreme weather events. It shows that most of the variation in gender ratio comes from floods in a district. Another reason for conflict in our results and Mueller et al.'s (2014) study is that gender-based migration patterns may respond differently to different weather events.

Table 5.2 also presents the diagnostic tests that confirm the validity and relevance of the chosen instrument. As discussed earlier in the previous section, the gender ratio is endogenous; therefore, to undertake causal analysis, we need to exploit exogenous variation in gender ratio to determine its causal impact on women's LFP. Table 5.2 presents the results of the first stage of TSLS, which shows point estimates of exposure to extreme weather events with robust SEs and SEs clustered at the district level and several test statistics to determine the validity and relevance of IV. The models include a set of individual and household level control variables and two gender norms indicators from PSLM. As discussed earlier, the IV method is used to remove the endogeneity of

the gender ratio. In the first stage, we instrument gender ratio by exposure to extreme weather events. Causal identification requires the instrument to be random and exogenous, and it should affect the outcome variable only through its effect on the endogenous variable.

Therefore, an IV must satisfy the two conditions: (1) Relevance, this condition requires IV to be correlated with an endogenous variable in the model, and (2) Exogeneity, i.e., IV should not be correlated with the error process. To check the first condition it requires testing of under-identification and weak identification, for which there are several tests available in the literature. However, for the second condition, there is a need to test for the orthogonality condition. For under-identification, Kleibergen-Paap rk Lagrange Multiplier is the widely used test in case of non-i.i.d errors, therefore, with robust or clustered SEs, this test statistic will be heteroscedastic-robust (Kleibergen & Paap, 2006), which happens to be the case in our research since we intend to cluster the SEs at the district level. The results reported in Table 5.2 below strongly reject the null hypothesis of under-identification in both versions of SEs. For identifying weak instruments, the tests are available such as first-stage F statistic on excluded instruments (Stock et al., 2002) which measures whether the coefficient on the excluded instrument is significantly different from zero. Staiger and Stock (1997) report that the first stage F-statistic must be large, which as a rule of thumb must be greater than 10. However, Lee, McCrary, Moreira, & Porter (2020) suggest this threshold to be greater than 104.7 rather than 10 to avoid distorted Type 1 error or coverage rates with the weak instrument. The authors also recommend having an alternative approach to rely on the F statistics threshold of 10. Still, in that case, critical values have to be adjusted from 1.96 to 3.43 to have the correct size or coverage.

Table 5.2 First Stage Results

VARIABLES	(1) Working-age gender ratio - Robust SEs	(2) Working-age gender ratio - Clustered SEs
Exposure to Extreme Weather	0.16***	0.16***
	(0.007)	(0.04)
Constant	1.37***	1.37***
	(0.01)	(0.07)
Observations	43,879	43,879
Kleibergen-Paap rk LM statistic	419.6	9.81
Kleibergen-Paap rk LM P-value	0.00	0.001

First stage F-Statistic	430.3	13.37
Kleibergen-Paap rk Wald F statistic	430.3	13.37
Endogeneity Test Statistic	21.66	3.87
Endogeneity Test P-value	0.00	0.05

Regressions include control variables, age, age square, education, marital status, head education, head labor market status, dependency ratio, household size, number of kids, number of infants, family structure, female-headed households, log of monthly income, employment decision and permission to work. Stock–Yogo critical values for weak identification tests (used for Kleibergen-Paap rk Wald F statistics) are 16.38 for 10%, 8.95 for 15% maximal IV size. Critical values are for one instrument and one endogenous regressor, *** p<0.01, ** p<0.05, * p<0.1.

The weak instrument problem is also tested with Kleibergen–Paap Wald rk F statistic, with the null hypothesis of the estimator being weakly identified. Stock and Yogo (2005) critical values are used for this test statistics. The critical values are determined by the number of endogenous variables in the model, the number of excluded instruments, and the size of the acceptable estimator bias. The critical values are the same for Cragg–Donald Wald statistic – a weak instrument test with homoscedasticity assumption – and Kleibergen–Paap Wald rk F statistic – a weak instrument test with non-i.i.d. errors. The first stage F statistics and Kleibergen–Paap Wald rk F statistic is reported for robust SEs and clustered SEs. With robust SEs, the test results show the instrument is strong, i.e., 'exposure' is strongly correlated with gender ratio with first-stage F statistics to be 430.3, greater than the threshold of 104.7 suggested by Lee et al. (2020).

Similarly, Kleibergen–Paap Wald rk F statistic is 430.3, greater than 10% critical value. However, when we assume dependence of observations across clusters, i.e., non-i.i.d error, the first stage F statistics is 13.37, and Kleibergen–Paap Wald rk F statistic is above 15% critical value. The reason for the low values of these statistics is because of excessive clustering of SEs at the district level because of which SEs of 'exposure' happen to be larger than the robust version of SEs. Overall, we may infer the instrument 'exposure' to be strong enough to explain exogenous variation in gender ratio. The reason is that extreme weather events are expected to drive peoples' relocation from the place of occurrence of weather events, which would be a source of changing gender ratio in a disaster-hit area.

Having discussed the instrument relevance, we now discuss the instrument exogeneity, i.e., whether the chosen instrument is sufficiently uncorrelated with the errors. Therefore

since we have one excluded instrument, we cannot test this condition through some formal tests. But we have a solid reason to believe that the chosen instrument is in no way associated with the error process. We use extreme weather events data from EM-DAT, which is then matched with PSLM survey data to measure the exposure to these extreme weather events. The instrument is constructed from this matched data by calculating how many total extreme weather events in a particular district a person has been exposed to ever since they are born. This instrument is perfectly exogenous because weather events are beyond one's control and happen exogenously and randomly due to climate change.

However, there is an argument in the literature on the exogeneity of weather events. Schultz & Mankin (2019) reveal that weather events are affected by civil conflicts and found that weather stations that record temperature is lost due to civil conflict, thereby leading to a negative correlation between conflicts and the number of weather stations. Moreover, Pretis (2021) is also of a similar view that the earth's climate affects human activity and vice versa, identifying reverse causality between climatic events and economic activity. These mechanisms are more likely to hold when country-level weather events are considered. The frequency of weather events may change in the long term due to environment-friendly government policies like the carbon tax. However, within the context of this research, we are using several weather events to construct 'exposure' to these events by individuals during their lifetime, starting from their year of birth. The year of birth of a person and how much exposure to weather events they would have during their whole life is by no chance going to have any feedback effect on the frequency of weather events. Therefore, we have a strong reason to believe that our chosen instrument is exogenous and thereby not correlated with the error process.

Next, we look at the test of endogeneity of gender ratio, which confirms with the null hypothesis that the specified endogenous regressor can be treated as exogenous. If the problem of endogeneity does not exist, then there is no need to use the IV method. In that case, OLS would produce reliable estimate than IV. The endogeneity test implemented is defined as the difference of two Sargan-Hansen statistics: one for the equation with the smaller set of instruments, where the suspect regressor(s) are treated as endogenous, and one for the equation with the more extensive collection of instruments, where the suspect regressors are treated as exogenous, and the test is also

robust to violations of conditional homoscedasticity. The endogeneity test statistics and associated probability values are reported for both versions of SEs. The null hypothesis that gender ratio can be treated as exogenous is strongly rejected, and therefore, the gender ratio is confirmed to be treated as endogenous as expected.

Finally, we apply a robustness check for the validity of our IV results. The results presented in Table 5.1 are obtained considering gender ratio measures with working age of men and women from 10 to 64 years of age. The lower age limit is chosen because PSLM data capture labor force statistics from 10 years of age and older. The upper limit is selected because it is the official limit in the definition of the working-age population in Pakistan (Pakistan Economic Survey, 2020). We also check the robustness of these results by considering working ages from 15-49. The lower limit is according to the official ILO working-age definition, and the upper limit is chosen because it is the realistic age for men/women who migrate or relocate themselves. The OLS and IV results with clustered and robust standard errors are robust to having gender ratio constructed with different working ages. The comparison of OLS and IV results with this specification is presented in Table 27, and the first stage results are shown in Table 28 in the appendix.

5.3 Study 2: Women's LFP and equitable gender role beliefs

As argued earlier, gender role beliefs and women's LFP have bidirectional causal relationships. Study 1 analyzed the causality from gender role beliefs to women's LFP. In this section, study 2 analyzes whether a causal direction exists from women's LFP to EGRB. For this purpose, we employ the random effect instrumental variable (REIV) TSLS method by utilizing panel BISP cash transfer data. As discussed in the previous chapter, in the 2016 round of BISP, there was a significant drop in panel households from 7563 surveyed from 2011 to 2014 to 3497 in round 2016. Because of this substantial difference in the number of panel households, we intend to conduct a separate analysis for a 3-year panel sample, i.e., from 2011, 2013, and 2014, and a 4-year panel sample, i.e., from 2011, 2013, 2014, and 2016. In doing so, the study aims to determine whether access to cash helps women change their labor market participation – the notion presented by modernization theory – and whether women's LFP, in turn, leads to equitable gender norms or not – the viewpoint of gender stratification theory.

As discussed in the previous chapter, different dimensions of women empowerment indicators are available in BISP data. However, these subjective measures of women empowerment/gender role beliefs are susceptible to self-reporting bias. Therefore, this research intends to construct a specific measure of gender role beliefs from both husbands' and wives' responses to questions on gender norms. This construct measures the 'coincidence' and 'conflict' of gender role perceptions between husband and wife in a household. Furthermore, the index of gender role belief is constructed with MCA, which is in a standardized form with 0 mean and approximately one standard deviation. To present the analysis, Table 5.3 compares OLS and IV results of this specification with two different models from a 3-year panel sample followed by first stage results of TSLS REIV in Table 5.4 to check the validity of IV used. The complete estimation results are presented in Table 29 in the appendix. I present the results of two models wherein Model 1, I regress husband-wife gender norms coincidence on women's LFP – in case of OLS – and instrumented women's LFP – in case of REIV along with a set of control variables. In model 2, the interaction term of husband gender norms index and women's LFP is included to explore further how the relationship between women' LFP and coincidence of gender role beliefs depends on husbands' gender role beliefs

5.3.1 OLS and Second Stage Results

Looking at OLS results in column 1 in Table 5.3, it is observed that compared to women who do not participate in the labor force, we expect that husband-wife gender norms coincidence to increase by 0.07 points on average when women participate in LF, and the result is statistically significant at 1% level. However, when the endogeneity of women'sLFP is removed, the relationship does not change but turns out to be statistically insignificant. As discussed in the previous chapter, if the woman is emancipated, but her husband is of traditional views, even if women are provided with financial resources, they are more likely to have similar gender norms as their husbands'. Therefore, given this insight, there is a need to control for husbands' gender role beliefs in the model to assess how the relationship between women's LFP and gender norms coincidence changes when controlling for husbands' gender norms perceptions. It is done by using the interaction effect of women's LFP and husbands' gender norms. Therefore, OLS and REIV results from columns 1 and 2 do not seem to provide a meaningful understanding of the causal effect of FLFP on EGRB.

Columns 3 and 4 in Table 5.3 present the OLS and REIV results after incorporating the interaction effect of FLFP and husbands' gender norms to analyze how the impact of women's LFP on gender norms coincidence changes depending on husbands' attitudes gender role beliefs. In the case of REIV, the result in column 4 suggests that the average value of husband-wife gender norms coincidence is 0.70 units lower for women who participate in the labor market than women who do not, and the result is statistically significant at 1 percent level. After controlling for this interaction effect, the main impact of women's LFP is contrary to the results from REIV in column 2, leading us to infer that when access to financial resources make women more emancipated, and they tend to participate in the labor force, their attitudes towards gender norms become different from their husbands'. Regarding OLS results in column 3, it is observed that the relationship between women's LFP and gender norms coincidence does not change. However, the magnitude of the effect of women's LFP is much smaller as compared to REIV. Since OLS does not capture the endogeneity of FLFP; therefore, the results are biased and inconsistent. A sharp change can be seen in the point estimate of women's LFP from 0.04 to 0.70. It shows that after cleaning the women's LFP from its endogeneity, 1 point increase in LFP leads to gender role belief coincidence to decreases by 0.70 points as compared to 0.04 points decrease in the case of OLS.

Study 2 found that the relationship between FLFP and gender norms coincidence is flipped due to adding the interaction effect in the model. The inverse relationship between FLFP and the coincidence of gender role beliefs implies that women who participate in the labor market will cause a change in their and husbands' gender role beliefs, thereby transforming gender role beliefs within the household. In other words, in response to higher women's LFP, the gender role perceptions among husbands and wives draw apart, and the differences in opinion arise among them on gender role beliefs. It shows that either husband or wife changed his/her beliefs on gender roles in the household in response to women's LFP. Since women's LFP rises due to being cash recipients of BISP, it makes sense to believe that its' wives' gender role beliefs change rather than husbands' and that too in favor of equitable gender norms. It confirms the gender stratification theory that when women participate in the labor force, they cause gender norms to be more equitable. So, the study approves that gender role beliefs change in response to women's LFP. In other words, the conflict of gender norms arises

when women become more emancipated in terms of their labor market participation because of being beneficiaries of cash transfers.

Regarding the main effect of husbands' gender norms, measured as an index with MCA, which is 0.53 means that with one standard deviation increase in husbands' gender role beliefs, i.e., husbands having more non-traditional attitudes, leads to 0.53 units increase in the likelihood that husband and wife have similar gender role beliefs. Moreover, the point estimate of the interaction term is 0.25, which shows that the presence of a participating woman in the household is associated with 0.25 units increase in the effect of husbands' gender norms on the coincidence of gender role beliefs between husbands and wives which is the difference between husbands' gender norms effect with participating and non-participating wife. It proposes that when a husband has non-traditional gender role beliefs, higher FLFP leads to a higher coincidence of gender norms. In other words, husband and wife are more likely to share gender role beliefs when husbands have modern gender norms. The coincidence of gender role beliefs increases with an increase in husbands' gender norms for women who are participating in the labor force than women who do not, thereby making the effect of FLFP on the coincidence of norms an increasing function of husbands' gender norms. This interaction term effect suggests that when a woman participates in the labor market, she has modern gender role beliefs. If a husband is also modern, the couple is more likely to have similar gender role beliefs. As husband and wife both are modern, there is a reason to believe that they would have this similarity in favor of EGRB.

Table 5.3 OLS and IV on MCA Index of Husband-wife gender norms coincidence with Official Treatment Status, Replicated Poverty Score, and their interaction as IVs- Panel Sample 2011-2014

	(1)	(2)	(3)	(4)
VARIABLES	OLS	REIV	OLS	REIV
FLFP	0.07***	0.32	-0.04**	-0.70***
	(0.02)	(0.22)	(0.02)	(0.20)
Husband gender norms index			0.52***	0.53***
			(0.01)	(0.01)
Husband gender norms index x FLFP			0.20***	0.25***
			(0.02)	(0.03)

Constant	-0.16***	-0.20***	-0.03	0.06
	(0.06)	(0.06)	(0.05)	(0.06)
Observations	9,527	9,527	9,508	9,508
Number of Groups	5,290	5,290	5,285	5,285
Overall R-sq	0.001	0.006	0.32	0.27

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Controls include Age, Completed Years of Education, Household Head Employment Status, Remittances received, Ownership of agriculture, non-agriculture, commercial and residential land, Access to Money Index, Number of working-age people in Household.

5.3.2 IV Validity

As discussed earlier endogeneity of women's LFP, we need to exploit exogenous variation in women's LFP to understand its causal impact on EGRB. For this purpose, we employ REIV, wherein we use instruments: the official BISP treatment status of households and the replicated poverty score and their interaction. In the BISP program, the households are categorized into treatment and control based on poverty scores. The poverty score is computed by employing the PMT method on a set of predetermined 23 poverty covariates. These poverty covariates are primarily taken from several dimensions: the number of dependents, household head education, children education, the ratio of rooms per person in the household, type of toilet, ownership of assets, ownership of livestock, and ownership of agriculture land. As these poverty indicators are hard to manipulate by the members of a household, therefore, poverty score based on these covariates is by no means influenced by the households. Hence, the treatment status determined based on poverty score cutoff can be an exogenous instrument in explaining women's LFP. However, since the poverty score has not been updated since 2011, households' treatment status has not changed since then, so this indicator lacks variation.

We utilize another instrument which is replicated poverty score of the household. This variable is computed using the OPM data collected at baseline in 2011 and then three midline data from 2013, 2014, and 2016. The official weights used in the BISP PMT model for each poverty covariate are taken from Hou (2009) to replicate poverty scores in each year. As the official poverty score is not updated since 2011, therefore, we use replicated score as an additional instrument to consider the actual poverty status of households based on recent data on poverty covariates. The replicated poverty score is based on the similar covariates used for computing the official poverty score; therefore, our replicated score can also serve as an instrument in explaining variations in FLFP.

So, we intend to analyze how an unchanged official BISP treatment status and varying replicated poverty scores affect women's LFP and look at its causal impact on EGRB.

It is also checked if the effect of these two instruments on women's LFP is somehow dependent on each other, and for that, we include the interaction term of BISP status and replicated poverty score. However, the inclusion of interaction terms comes at the cost of multicollinearity among predictors. Multicollinearity can produce misleading results; therefore, we standardize replicated score and then interact it with binary BISP status to avoid this problem. Table 5.4 presents the first stage REIV results for the models with interaction effect of FLFP and husband gender norms index and without interaction. However, the results do not change in both models. The point estimate of BISP status is positive but relatively low, and it is also statistically insignificant. It suggests that being a beneficiary of BISP cash transfer merely influences women to become part of the labor force in the panel sample from 2011 to 2014. In statistical language, there is no effect of being treated in this program on FLFP. However, the main impact of standardized replicated poverty score shows that for one standard deviation increase in replicated poverty score, the probability for women to be part of the labor force decreases by 0.05 units in model without interaction. It shows that women from lower poverty score households prefer to participate in the labor market, and the relationship is statistically significant at a 1% level. As shown in column 2, this relationship does not change even after controlling the interaction of husbands' gender norms with FLFP. The interaction effect of treatment status and the replicated score is insignificant, showing that although the main impact of replicated poverty score is statistically significant, the effect of treatment status on FLFP does not depend on replicated poverty score and can be due to inadequate variation in official treatment status.

Now we present some diagnostic statistics to check the relevance and exclusion restriction of our chosen instruments. We use under-identification and weak identification tests to check relevance conditions, which are reported in Table 5.3. The p-value of Kleibergen-Paap rk LM statistics shows that the null hypothesis of under-identification is rejected at a 1% significance level. First-stage F statistics and Kleibergen-Paap rk Wald F statistics are used to check the weak instrument, with the null hypothesis of the estimator being weakly identified. Stock and Yogo (2005) critical values are used for Kleibergen-Paap rk Wald F statistics. For both the models, the calculated value for this test is greater than 10% critical values, leading to the rejection

of the null hypothesis of the weak instrument. Hence, we can confirm our chosen instrument to be strong enough in explaining exogenous variation in FLFP.

For detecting the exogeneity of instruments, the Sargan-Hansen test of overidentifying restriction is used, which tests the null hypothesis that instruments are valid, i.e., the instruments are uncorrelated with the error terms. The Hansen-J statistics p-value fails to reject this hypothesis for the model with interaction in column 2. However, for the model without interaction in model 1, the test rejects the hypothesis of validity of the instrument. The rationale for choosing these instruments has already been discussed, based on which we can safely argue that instruments are exogenous and uncorrelated with the error term. Next, we look at the endogeneity of FLFP. We use the endogeneity test implemented as the difference of two Sargan-Hansen statistics with the null hypothesis that the specified endogenous regressor can be treated as exogenous. The p-value for this test in model with interaction rejects the null hypothesis and concludes that FLFP must be treated as an endogenous variable. However, the test fails to reject the null hypothesis for the model without interaction. From the diagnostic tests of chosen instruments, it is observed that instruments performed well in the presence of interaction effect of FLFP and husband gender role beliefs.

Table 5.4 First Stage Results - Panel Sample 2011-2014

VARIABLES	(1) REIV-I	(2) REIV-I
BISP Status	0.009	0.009
	(0.01)	(0.01)
Replicated Poverty Score	-0.05***	-0.05***
	(0.006)	(0.006)
BISP Status x Replicated Poverty Score	0.01	0.01
	(0.01)	(0.01)
Constant	0.16***	0.16***
	(0.02)	(0.02)
Observations	9,527	9,508
R-squared	0.02	0.03
Kleibergen-Paap rk LM statistic	113	106.9
Kleibergen-Paap rk LM P-value	0	0
First stage F-Statistic	43.51	40.08
Kleibergen-Paap rk Wald F statistic	43.51	40.08
Hansen J Statistic	12.52	1.886
Hansen J P-value	0.001	0.39
Endogeneity Test P-Value	0.35	0.001

Robust standard errors clustered at the household level in parentheses. Controls include Age, Completed Years of Education, Household Head Employment Status, Remittances received, Ownership of agri, non-

agri, commercial, and residential land, Access to Money Index, Number of working-age people in Household. Stock–Yogo critical values for weak identification tests (used for Kleibergen-Paap rk Wald F statistics) are 16.38 for 10%, 8.95 for 15% maximal IV size. Critical values are for one instrument and one endogenous regressor, *** p<0.01, ** p<0.05, * p<0.1.

The results of REIV and OLS for a 4-year panel sample are given in Table 30 in the appendix. The first stage results are shown in Table 31. Overall, the results do not differ, but the effect of women's LFP on gender norms coincidence in both OLS and REIV are no more statistically significant. However, husbands' gender norms significantly and positively affect the coincidence of norms, which is also true for the interaction effect. It implies that, although we cannot infer statistically significant results concerning the main effect of women's LFP, its point estimates do suggest some correlation between the outcome variables and the FLFP. Therefore, in a small number of panel households from 2011 to 2016, we do not find a causal effect of FLFP on shifting gender norms in a household. However, the presence of a participating woman in the household is associated with a 0.19 units increase in the effect of husbands' attitudes on the coincidence of norms.

For the first stage results, the point estimate of BISP status is positive and insignificant, indicating no effect of being treated in this program on FLFP. However, the main impact of standardized replicated poverty score shows that for one standard deviation increase in replicated poverty score, the probability for women to be part of the labor force decreases by 0.027, and 0.026 in models without interaction and with the interaction of women's LFP with gender norms coincidence, respectively. It shows that women from lower poverty score households prefer to participate in the labor market, and the relationship is statistically significant at a 1% level. The interaction effect of treatment status and replicated score happens to be insignificant, showing that although the main effect of replicated poverty score is statistically significant, its interaction effect with official treatment status is insignificant.

To validate the results of REIV, we provide the robustness check by using placebo treatment status as IV. As discussed before, as the placebo treatment status was never assigned to the households, this IV should not affect FLFP or the outcome variable. The REIV results with placebo status are reported in Table 32 in the appendix for the panel sample and two estimation models. The results align with what we expected to be, confirming the statistically insignificant effect of FLFP on husband-wife gender norms coincidence. Moreover, to further validate the results of this study, we apply RDD,

which verifies the presence of a discontinuity in FLFP just above and below the poverty score cutoff. The results of RDD, their interpretation, graphical representation, and validity of RDD results are presented in appendix 2

5.4 Summary

Study 1 deals with the hypothesis that EGRB have a causal impact on women's LFP. To investigate this hypothesis, I utilized the pooled cross-section household survey data from the Pakistan Social and Living Standard Measurement (PSLM) survey. I constructed the instrumental variable of exposure to extreme weather events by combining the PSLM survey data and emergency data events data that capture extreme weather events. In examining the causal impact of gender role beliefs on FLFP, we do not explicitly consider the social norms indicators measured through the PSLM survey because of potential self-reporting bias. To prevent our analysis from this bias, we look at how variation in gender ratio in a region influences FLFP in those regions, a link based on Becker's marriage market hypothesis. The underlying sense is that when there is variation in gender ratio, for example, more men relative to women, the labor market condition for women become favorable through implicitly affecting existing gender norms and making traditional attitude in society challenging to persist. Since gender ratio is itself endogenous in the model, therefore, to examine the effect of gender ratio, we instrument it with people's exposure to extreme weather events. I check the relevance and exogeneity of instrument through formal tests, and we provide a valid explanation as to why our chosen instrument is exogenous to the model.

I instrumented gender ratio with district-level variation in exposure to extreme weather events, because of which we used standard errors clustered at the district level. Clustering the standard error was used for both OLS and IV models because of instrument measurement at the district level. However, we presented the comparison of results with robust standard errors and cluster-robust standard errors. The first stage results showed that increased 'exposure' leads to a higher gender ratio, i.e., more men than women. This effect could be because of more inter-district migration of women than men after exposure to weather events. Therefore, it can be said that extreme weather events will make the prevailing strict social norms undesirable, which is why women migrate despite the prevalent traditional norms. The findings from study 1 revealed contrasting OLS and IV results which confirms that biasedness of OLS results in the presence of endogeneity. IV results showed that even in the presence of a male-

biased gender ratio, women are still participating in the labor market reflected through higher women's LFP in response to higher gender ratio in the second stage of IV. One explanation of this relationship could be that extreme weather events especially, floods and extreme temperatures destroy household livelihoods. Therefore, to mitigate the effect of income loss, women and men will increasingly take responsibility for providing income to the household. They may be compelled to work in the fields and fix the damages caused by flood or heatwave, sell homemade production, or find a small job to meet the households' financial needs. Given these results, the predictions of Becker's marriage market hypotheses are less likely to hold in the case of Pakistan. It may be due to the nature of the instrument we are looking at, which affects the gender ratio.

Study 2 examined causality between EGRB and women's LFP by using nationwide cash transfer data from Pakistan, Benazir Income Support Program (BISP). Random effect instrumental variables estimation strategy is used to investigate the hypothesis of study 2. The choice of random effect is based on the nature of the variable being instrumented, women's LFP, and the chosen instrument, BISP official treatment status. Women's LFP is found to have less variation over time. BISP official treatment status has not changed since the program was initiated, making the fixed effect model inappropriate to utilize. The BISP treatment status of households had not changed because the poverty score had not been updated since 2011. This research attempts to replicate the poverty score using the BISP proxy means test (PMT) method and official weights by using the recent rounds of BISP to obtain the actual poverty score of the household based on current data on poverty covariates. I use this replicated poverty score and official treatment status as instrumental variables to explain exogenous variation in women's LFP. These instruments are found to fulfill the relevance and exclusion restriction.

The study uses an indicator of husband-wife gender norms coincidence by matching the responses given by husbands and wives on six gender norms questions. We then further construct an index of these matched responses through multiple correspondence analysis (MCA). The investigation is conducted on a 3-year panel sample, i.e., panel households from 2011, 2013, and 2014, and 4-year panel households from 2011, 2013, 2014, and 2016. The analysis was conducted on two different models 1) I examine the effect of women's LFP on the coincidence of husband-wife gender norms index 2) I include the interaction of husbands' gender norms with women's LFP to examine how the effect of

women's LFP on the coincidence of norms depends on whether a husband is modern or traditional. The first stage results show that official treatment status was found to have a negligible impact on women's LFP. However, there is a statistically significant negative impact of replicated poverty score on women's LFP. It suggested that women from lower poverty score households tend to prefer to participate in the labor market. This relationship confirms the prediction of modernization theory that when women belong to poor households and are given access to financial resources, they intend to increase their participation in the labor force, indicating them to be more modern.

We compared the OLS and second stage REIV results for the models discussed above. Results of the model without interaction suggested that the more women participate in the labor force, the more they will comply with their husbands' gender norms. However, after incorporating the interaction effect, the results differ for both OLS and REIV. These results show that the main impact of women's LFP on a coincidence of norms is negative and statistically significant. It proposes that women who participate in the labor market will cause their gender norms to be different from their husbands, thereby bringing in the transformation of gender norms within the household leading us to confirm the gender stratification theory that when women participate in the labor force, they cause gender norms to be more equitable.

Moreover, the main effect of husbands' gender norms and the interaction is positive and statistically significant. It indicates that when the husband has non-traditional gender norms, higher women's LFP leads to a higher coincidence of gender norms. In other words, husband-wife gender norms coincidence tends to be higher in households where husbands have modern gender norms. One may expect that with non-traditional husband and wife participation in LF, both husband and wife would have similar non-traditional gender role beliefs.

The structure of BISP panel data permitted us to conduct the regression discontinuity analysis to examine how women behaved, in terms of their labor market participation, in response to being treated, those who are just above and just below the poverty score cutoff. Regression discontinuity design helped us examine this local average treatment effect. The results of RDD presented in appendix 2 suggest that closer to the poverty score cutoff, the beneficiary women do not intend to participate in the labor force as compared to women above the cutoff. However, this inference is not statistically

significant considering the optimal bandwidth. RDD is employed on a narrow subsample from all the available panel households, making the analysis weak in its external validity. When the beneficiary households outside the chosen bandwidth are systematically different from those within the bandwidth, then a lack of external validity is likely to occur.

From the first stage of REIV, it is found that being a beneficiary of BISP cash transfer has a very small influence on women to become part of the labor force and is too statistically insignificant. It was derived by considering all the panel households from the data, understanding the phenomenon from a broader perspective. If we link the first stage REIV results with RDD estimates, it turned out that although there is a small positive insignificant impact of cash transfer on FLFP in a broader range of households, if a narrower range of households is considered who are just above and below the threshold, the effect of cash transfer appears to be negative but still statistically insignificant with optimal bandwidth. Therefore, the RDD results can be taken as additional evidence of first stage REIV that either we look at all the households in the panel sample of households closer to the cutoff. The BISP cash transfer does not contribute to increasing the women's labor force participation. However, let's look at the impact of replicated poverty scores. In households with lower poverty scores, women tend to participate in the labor force, and this relationship is statistically significant.

After presenting the main findings, the results are also checked for their robustness. In examining the first hypothesis, the robustness of the results is checked by using different working ages. In the main analysis, I considered working age from 10-64, while in robustness check, the age limit of 15-49 years is considered. The results are consistent in both the specification. Concerning the second hypothesis, I constructed a placebo treatment status. I used it as instrumental variable to verify that it does not produce a statistically significant effect of FLFP on gender norms coincidence. The results with this placebo treatment status are in line with the expectations. Overall, based on the results of this research, it can be inferred that in Pakistan, there exists bidirectional causality between social norms and women's labor force participation

CHAPTER 6 CONCLUSIONS AND POLICY IMPLICATIONS

6.1 Conclusions

The central idea of this research was to investigate the bidirectional causal relationship between EGRB and women's LFP focusing comprehensively on one specific society: Pakistan, where favorable conditions and data availability allows verifying the validity of bidirectional causality at the same time. There is strong theoretical evidence that confirms the correlation between social norms and women's labor force participation. These correlations are nested in theories, but the existing theories propose opposing explanations for the observed correlation. The neoclassical and modernization theories point that the causality runs from social norms to labor force participation of women: LFP decisions among women are conditioned by EGRB. The neoclassical theory relies on a single utility function and assumes that household members have the same preferences and pool their resources to maximize the single utility function.

Given that Pakistan is a patriarchal society and the male household heads make essential decisions, the neoclassical theory predictions were expected to hold. Along with this theory, the research also tested Becker's marriage market hypothesis, which suggests that variation in gender ratio affects female labor force participation. On the other hand, gender stratification theory provides another perspective. It proposes that women's participation in economic production and control over material resources can increase bargaining power, motivating a shift to more equitable gender norms. The theory argues that gender stratification is inversely linked to women's access to material resources and the level of women's economic power. An increase in women's access to resources and their ability to participate in the labor market can reduce their ideological and physical oppression in societies.

The study 1 findings are based on the instrumental variables method on pooled cross-section household survey of PSLM, and for study 2, BISP cash transfer data is utilized. PSLM data is used to examine the first hypothesis of the causal impact of EGRB on FLFP. I used the instrument of exposure to extreme weather events to explore this causal relationship. Gender ratio, the endogenous variable in the model, is instrumented with exogenous variation in exposure to extreme weather events. The first stage of IV results indicates the statistically significant positive impact of exposure on gender ratio.

In the second stage of IV, it is found that a higher gender ratio in a specific district leads to higher women's LFP, thereby making Becker's marriage market hypothesis fail to hold in Pakistan given the nature of the instrument constructed through the number of extreme weather events. Overall, from the results of study 1, it is inferred that EGRB, implicitly considered through the effect of changing gender ratio, causally impact women' LFP, i.e., higher the exposure to weather events, higher will be the gender ratio which in turn leads to more women's LFP.

Contrary to the empirical findings on climate-induced migration in Pakistan, such as (Mueller et al., 2014), women are active agents for climate-induced migration. However, the result is in line with the findings of the study from Bangladesh, which concludes that "women should not be perceived as passive, but rather active agents in migration processes, who at the same time face greater struggles than men in achieving their goals" (Evertsen, & van der Geest, 2020, p. 20). Thus, study 1 of this thesis confirms that the equitability of gender role beliefs caused women to increase their labor market participation in response to exogenous variation in gender ratio.

Study 2 of this research was investigated through the panel BISP cash transfer data. I utilized a random effect instrumental variable approach to examine the direction of causality from FLFP to EGRB. In this study, exogenous variation in women's LFP is considered by using the official treatment status of household and replicated poverty score and their interaction effect. These results show that the impact of FLFP on the coincidence of gender role beliefs is negative and statistically significant. It proposes that women who participate in the labor market will cause their gender norms to be different from her husband thereby bringing in the transformation of gender norms within the household leading us to confirm the gender stratification theory that when women participate in the labor force, they cause gender norms to be more equitable. Thus, higher FLFP leads to conflict in gender role perceptions between husband and wife.

6.2 Policy Implications

The findings from the studies have some meaningful policy implications. From the results of study 1, women are found to be relocating themselves due to exposure to extreme weather events, and these women are mainly poor, less educated, and less

skilled. Therefore, there should be such a mechanism at the district level to observe the migration patterns after any extreme weather event. Since women are mainly the migrants, either alone or with their whole families, there should be the availability of some employment opportunities of these families and specifically women, so that they can safely mitigate the dual adverse effect of climate change, one in terms of their livelihood loss and second leaving their place of residence. Secondly, since women tend to work more in the labor force after the adverse climate shock, which makes them more vulnerable, there should be a strategy that reduces their vulnerability. It can be done by providing them with financial resources.

Moreover, the donor agencies give the funds after extreme weather events; specifically, floods should be disbursed fairly to the vulnerable groups. Thirdly, since Pakistan is among the countries with high vulnerabilities to climate change, the frequency of extreme weather events is most likely to increase in the future. Given this fact, there are prospects for more climate-induced migration. Therefore, there is a need to have an official database that captures the population mobility, specifically including the gender dimension of migration. There is also a need to establish a weather events database at narrow geographical levels such as Tehsils, an administrative sub-division under a district.

Regarding study 2 from BISP cash transfer data, significant policies are derived. The study encountered challenges concerning data and survey questions specifically related to gender role beliefs. The indicators of gender role beliefs from the survey have self-reporting bias. Therefore, there is a need to change the way gender norms questions in surveys are framed. Three out of six questions on gender role beliefs that are asked from both men and women are stated in negative connotations, making the analysis, especially constructing an index out of those indicators, challenging to implement. If these questions were asked consistently and in a similar tone, the investigation would have been more interesting. Second, the women empowerment questions in the survey should be asked from both women and their husbands or other male head of the household. It can help researchers evaluate and assess dynamics in gender norms holistically at the household level rather than the individual level. Third, survey questions on gender role beliefs and women empowerment should be framed to extract respondents' perceptions on whether their responses are according to what exists or whether they want or like to have those norms prevail. If this is done, the social

desirability bias can be overcome. For REIV results, it was found that women beneficiaries of cash transfer increased their labor force participation. Regarding how women's emancipation affects gender role beliefs in households, it was found that after becoming part of the labor market, wives' gender role beliefs are different from their husbands'. Thus there is a need to create awareness among male household members of the potential benefits the household could have if women work and that working women should not be considered as social stigma.

6.3 Contribution of this thesis

This thesis undertakes two separate studies on the bidirectional causality between social norms – measured through equitable gender role beliefs – and women's labor force participation by adopting the causal identification method of Instrumental Variables. The existing empirical evidence from separate settings shows that gender norms or gender role beliefs significantly impact women's LFP. More women in the labor force transform the current gender norms. This thesis relies on neoclassical labor supply theory, modernization theory, and gender stratification theory and contributes to this literature by verifying both causal directions within a single country's context. The findings from both studies confirm the two hypotheses 1) more EGRB cause an increase in women's LFP, and 2) women's LFP causes gender role beliefs to be more equitable. Thus, the results prove the presence of reverse causality between equitable gender role beliefs and women's LFP in Pakistan, thereby providing evidence that these variables amplify each other. This feedback loop between gender role beliefs and women's LFP also presents an essential insight into whether a virtuous or vicious circle exists. Based on the results, it is suggested that EGRB cause higher women's LFP. Therefore, there is a virtuous cycle effect rather than a vicious cycle because women's economic empowerment improves their skills and resources and positively contributes to transforming traditional gender role beliefs, further providing a better social and economic environment for future generations. So, empowered women are the source of creating gender role beliefs to be equitable in their households, i.e., at the micro-level in the short run, and transform gender norms in the whole society at the macro level in the long run. For countries like Pakistan with a GDP of $5000 per capita per annum, this transformation is invaluable as almost half of the country's population (women) may use their potential to bring economic progress.

Specific to study 1, it contributes to the literature on the gender aspect of climate-induced migration. It is observed in the research that there is a dearth of literature on environmental impact on female migration (Chindarkar, 2012). The relevant literature is missing specifically from South Asia, a region highly vulnerable to environmental changes that may contribute to significant human migration in the future. The findings from study 1 approve of the female outmigration due to climatic events, thereby presenting an interesting and unique insight that climate-induced migration is not always male-dominated; instead, women also tend to migrate and become part of the labor market in response to weather events. It ultimately exhibits the equitability of gender role beliefs in a household that causes women to become part of the labor force.

Moreover, the thesis has a unique specification regarding the measurement of equitable gender role beliefs in study 2. The underlying principle is that available gender role beliefs from survey data are susceptible to self-reporting bias. Therefore, to overcome this issue, data-driven measures of gender role beliefs are utilized. For study 2, the thesis uses 'coincidence of gender role beliefs' between husband and wife to determine the causality. With this specification, the husband's gender role beliefs are considered along with the wife's gender role beliefs to analyze how changes in FLFP influence the 'coincidence' of beliefs between them. Therefore, this thesis presents another significant contribution to the literature by showing that it is not only women's emancipation that makes them empowered. Instead, the husband's gender norms perception is also essential. So, in a patriarchal society such as Pakistan, to make women's empowerment more sustainable and to have a virtuous cycle impact, husbands should be enablers of women's empowerment and not the barriers.

6.4 Limitations of the study and directions for future research

Existing literature provides evidence on correlational links between social norms and women's LFP without addressing the problem of endogeneity. This thesis addresses this literature gap and comprehensively investigates the bidirectional causal impact by utilizing instrumental variable strategy. However, there are certain limitations of this thesis. First, regarding study 1, the main limitation concerns the operationalization of gender ratio. Since the study chooses the exposure to extreme weather events as an instrument calculated by considering total extreme weather events that occurred in a district from 1950 to 2014, this 'exposure' is used to clean endogeneity of gender ratio

in the model. Therefore, variation in gender ratio is also considered at the district level, whereas the analysis is carried out at the individual level. It has led to large standard errors and insignificant IV results. The results could have been significant if the data had permitted calculating the occurrence of extreme weather events at a narrower geographical level, such as the Tehsil level. This way, standard errors would not have been too conservative as in IV results of study 1.

In the second study, the 4-year panel sample was small, i.e., 3497 households from 2011 to 2016, compared to a 3-years panel sample size with the 7563-panel household sample. The small 4-year panel sample limited the power to analyze the variation in infrequent variables such as women's LFP. The consequences of this limitation can be observed in the form of insignificant results in the case of a 4-year panel sample. If BISP had a bigger panel sample in recent years, the results would have been significant. An additional limitation concerns the poverty score calculation by BISP. As discussed before, the BISP has not updated the households' poverty score since baseline because of which treatment status has not changed too. Suppose there is little or no within-subject variation in this variable over time. In that case, the fixed-effect model becomes inappropriate to use because of time-differencing the variables, resulting in large standard errors (Wooldridge, 2013). There must be time-varying variables for them to be estimated using the FE model.

In contrast, the random effect model estimates the time-constant variables but at the cost of resulting in biased estimates because of not controlling for omitted variables. Moreover, women's LFP, the key variable of interest, also exhibited sluggish variation within the panel sample. Thus, the nature of this variable and the chosen instrument led to rely on random effect instrumental variable, which has serious drawbacks. Hence, the results from study 2 should be interpreted with caution. However, the study has tried to mitigate this bias by replicating the poverty score of each household. Regarding future research, it is recommended that researchers ensure sufficient variation in the instrument and the key variable of interest.

Another limitation of the second study of the thesis is that the analyses have been restricted to labor force participation of 'wives' in a household, and 'coincidence' of gender role beliefs have been considered between husbands and wives only. Future studies can undertake a broader analysis of variation in labor force participation of all

the women in the households in response to being beneficiaries of BISP. Also, the coincidence of gender role beliefs can be observed between fathers and daughters and between mothers and sons. It may also help understand the intergenerational transfer of norms in a household due to program recipients.

REFERENCES

Abramitzky, R., Delavande, A., & Vasconcelos, L. (2011). Marrying up: the role of sex ratio in assortative matching. *American Economic Journal: Applied Economics*, *3*(3), 124-57.

Acemoglu, D., Autor, D. H., & Lyle, D. (2004). Women, war, and wages: The effect of female labor supply on the wage structure at midcentury. *Journal of political Economy*, *112*(3), 497-551.

Adamo, S. B. (2010). Environmental migration and cities in the context of global environmental change. *Current Opinion in Environmental Sustainability*, *2*(3), 161-165.

Aguirre, D., L. Hoteit, C. Rupp, and K. Sabbagh (2012). *Empowering the Third Billion. Women and the World of Work in 2012*, Booz and Company.

Ahmad, E., & Hafeez, A. (2007). Labour supply and earning functions of educated married women: A case study of Northern Punjab. *The Pakistan Development Review*, 45-62.

Akerlof, G. A., & Kranton, R. E. (2000). Economics and identity. *The quarterly journal of economics*, *115*(3), 715-753.

Alesina, A., Giuliano, P., & Nunn, N. (2013). On the origins of gender roles: Women and the plough. *The quarterly journal of economics*, *128*(2), 469-530.

Algan, Y., & Cahuc, P. (2005) "The roots of low European employment: Family culture?" *CEPR Discussion Paper No. 5169*: Centre for Economic Policy Research.

Ambler, K., & De Brauw, A. (2016). *The impacts of cash transfers on women's empowerment: learning from Pakistan's BISP program*. World Bank.

Ambler, K., & De Brauw, A. (2019). *Household labor supply and social protection: Evidence from Pakistan's BISP cash transfer program* (Vol. 1815). Intl Food Policy Res Inst.

Amuedo-Dorantes, C., & Pozo, S. (2006). Migration, remittances, and male and female employment patterns. *American Economic Review*, *96*(2), 222-226.

Angrist, J. (2002), How Do Sex Ratios Affect Marriage and Labour Markets? Evidence from America's Second Generation, *The Quarterly Journal of Economics*, 117, 997–1038.

Angrist, J. D., & Krueger, A. B. (1999). Empirical strategies in labor economics. In *Handbook of labor economics* (Vol. 3, pp. 1277-1366). Elsevier.

Angrist, J. D., & Lavy, V. (1999). Using Maimonides' rule to estimate the effect of class size on scholastic achievement. *The Quarterly journal of economics*, *114*(2), 533-575.

Angrist, J. D., & Pischke, J. S. (2010). The credibility revolution in empirical economics: How better research design is taking the con out of econometrics. *Journal of economic perspectives*, *24*(2), 3-30.

Antecol, H. (2003). *Why is there cross-country variation in female labor force participation rates? The role of male attitudes toward family and sex roles* (No. 2003-03). Claremont Colleges Working Papers.

Antonakis, J., Bendahan, S., Jacquart, P., & Lalive, R. (2010). On making causal claims: A review and recommendations. *The leadership quarterly*, *21*(6), 1086-1120.

Arif, G. M. (2005). Internal migration and household well-being: Myth or reality. *Internal Labour Migration in Pakistan. Institute of Developing Economies, Japan External Trade Organisation, Chiba, Japan.*

Asian Development Bank (2012). Addressing Climate Change and Migration in Asia and the Pacific', Hugo, G., Bardsley, D., Sharma, V., Tan Y., Williams, M., and Bedford, R. (eds), Asian Development Bank, Mandaluyong City, Philippines. ISBN: 978-92-9092-611-5

Asian Development Bank (2016). Policy brief on female labor force participation in Pakistan‖ No. 70

Asselin, L. M. (2009). *Analysis of multidimensional poverty: Theory and case studies* (Vol. 7). Springer Science & Business Media.

Azid, T., Khan, R. E. A., and Alamasi, A.M.S. (2010). Labor force participation of married women in Punjab (Pakistan), *International Journal of Social Economics*, Vol. 37 No. 8, pp. 592-612

Baird, S., McKenzie, D., & Özler, B. (2018). The effects of cash transfers on adult labor market outcomes. *IZA Journal of Development and Migration*, *8*(1), 1-20.

Banerjee, A. V., Hanna, R., Kreindler, G. E., & Olken, B. A. (2017). Debunking the stereotype of the lazy welfare recipient: Evidence from cash transfer programs. *The World Bank Research Observer*, *32*(2), 155-184.

Banerjee, S., Gerlitz, J. Y., & Hoermann, B. (2011). *Labour migration as a response strategy to water hazards in the Hindu Kush-Himalayas*. International centre for integrated mountain development (ICIMOD).

Barbara, B., & Ratzel, J. (2005). The Effects of Social Norms and Perceptions on the Relationship Between Labor Market Opportunities and Women's Status in India. Paper presented at annual IAFFE meetings, Washington, DC.

Barnett, J., & Adger, W. N. (2007). Climate change, human security and violent conflict. *Political geography*, *26*(6), 639-655.

Becker, G. S. (1973), "A Theory of Marriage: Part I", *Journal of Political Economy*, 81, 813–46.

Becker, G. S. (1981). Altruism in the Family and Selfishness in the Market Place. *Economica*, *48*(189), 1-15.

Becker, Gary S. (1965). A theory of the allocation of time, *The Economic Journal* 75(299), 493-517.

Bertrand, M., Kamenica, E., & Pan, J. (2015). Gender identity and relative income within households. *The Quarterly Journal of Economics*, *130*(2), 571-614.

Bertrand, M., Mullainathan, S., & Miller, D. (2003). Public policy and extended families: Evidence from pensions in South Africa. *The World Bank Eeconomic Review*, *17*(1), 27-50.

Binzel, C., & Assaad, R. (2011). Egyptian men working abroad: Labour supply responses by the women left behind. *Labour Economics*, *18*, S98-S114.

Black, R., Adger, W. N., Arnell, N. W., Dercon, S., Geddes, A., & Thomas, D. (2011). The effect of environmental change on human migration. *Global environmental change*, *21*, S3-S11.

Black, S. E. (1999). Do better schools matter? Parental valuation of elementary education. *The quarterly journal of economics*, *114*(2), 577-599.

Blau, F. D., Ferber, M. A., & Winkler, A. E. (1992). *The Economics of Women. Men and Work, Englewood Cliffs.*

Blumberg, R. L. (1988). Income under female versus male control: Hypotheses from a theory of gender stratification and data from the Third World. *Journal of Family Issues*, *9*(1), 51-84.

Brainerd, E. (2017). The lasting effect of sex ratio imbalance on marriage and family: Evidence from World War II in Russia. *Review of Economics and Statistics*, *99*(2), 229-242.

Bütikofer, A. (2013). Revisiting 'mothers and sons' preference formation and the female labor force in Switzerland. *Labour Economics*, *20*, 82-91.

Calonico, S., Cattaneo, M. D., & Titiunik, R. (2014a). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica*, *82*(6), 2295-2326.

Calonico, S., Cattaneo, M. D., & Titiunik, R. (2014b). Robust data-driven inference in the regression-discontinuity design. *The Stata Journal*, *14*(4), 909-946.

Cameron, A. C., & Trivedi, P. K. (2005). *Microeconometrics: methods and applications*. Cambridge university press.

Cardoso, A. R., & Morin, L. P. (2018). Can economic pressure overcome social norms? The case of female labor force participation.

Cattaneo, M. D., Idrobo, N., & Titiunik, R. (2019). A practical introduction to regression discontinuity designs: Foundations. Cambridge University Press.

Chafetz, J. S. (1988). The gender division of labor and the reproduction of female disadvantage: Toward an integrated theory. *Journal of Family Issues*, *9*(1), 108-131.

Channa, A. (2012). Social protection in Pakistan: a profile of existing programmes and an assessment of data available for analysis. *Background paper, Asia Research Centre, London School of Economics, London*.

Chant, S. (1991). Gender, migration and urban development in Costa Rica: the case of Guanacaste. *Geoforum*, *22*(3), 237-253.

Charles, M., & Bradley, K. (2009). Indulging our gendered selves? Sex segregation by field of study in 44 countries. *American journal of sociology*, 114(4), 924-976.

Chaudhry, I. S., & Nosheen, F. (2009). The determinants of women empowerment in Southern Punjab (Pakistan): An empirical analysis. *European Journal of Social Sciences*, *10*(2), 216-229.

Cheema, I., Farhat, M., Hunt, S., Javeed, S., Pellerano, L., & O'Leary, S. (2014). Benazir income support programme.

Cheema, I., Hunt, S., Javeed, S., Lone, T., & O'Leary, S. (2016). Benazir income support programme: Final impact evaluation report.

Chiappori, P. A. (1992). Collective labor supply and welfare. *Journal of political Economy*, *100*(3), 437-467.

Chiappori, P.-A., Fortin, B. and Lacroix, G. (2002), "Marriage Market, Divorce Legislation, and Household Labour Supply", *Journal of Political Economy*, 110, 37–72.

Chintagunta, P., Erdem, T., Rossi, P. E., & Wedel, M. (2006). Structural modeling in marketing: review and assessment. *Marketing Science*, *25*(6), 604-616.

Chishti, S., A. Lodhi and S. Rashid (1989), Female Labor Participation Behavior: A Case Study of Karachi. *Pakistan Journal of Applied Economics*, 8(2): 157-165.

Cislaghi, B., & Heise, L. (2018). Theory and practice of social norms interventions: eight common pitfalls. *Globalization and health*, *14*(1), 1-10.

Cislaghi, B., & Heise, L. (2020). Gender norms and social norms: differences, similarities and why they matter in prevention science. *Sociology of health & illness*, *42*(2), 407-422.

Codazzi, K., Pero, V., & Albuquerque Sant'Anna, A. (2018). Social norms and female labor participation in Brazil. *Review of Development Economics*, *22*(4), 1513-1535.

Contreras, D., & Plaza, G. (2010). Cultural factors in women's labor force participation in Chile. *Feminist Economics*, 16(2), 27-46.

Cook, T. D. (2008). Waiting for life to arrive: a history of the regression-discontinuity design in psychology, statistics and economics. *Journal of Econometrics*, *142*(2), 636-654.

Corrigall, E. A., & Konrad, A. M. (2007). Gender role attitudes and careers: A longitudinal study. *Sex roles*, *56*(11-12), 847-855.

Cunningham, M. (2008). Influences of gender ideology and housework allocation on women's employment over the life course. *Social science research*, *37*(1), 254-267.

De Brauw, A., Gilligan, D. O., Hoddinott, J., & Roy, S. (2015). The impact of Bolsa Família on schooling. *World Development*, *70*, 303-316.

De Haas, H. (2008). The myth of invasion: the inconvenient realities of African migration to Europe. *Third world quarterly*, *29*(7), 1305-1322.

De Haas, H., & Van Rooij, A. (2010). Migration as emancipation? The impact of internal and international migration on the position of women left behind in rural Morocco. *Oxford development studies*, *38*(1), 43-62.

Desai, S., & Banerji, M. (2008). Negotiated identities: male migration and left-behind wives in India. *Journal of Population Research*, *25*(3), 337-355.

Dildar, Y. (2015). Patriarchal norms, religion, and female labor supply: Evidence from Turkey. *World Development*, *76*, 40-61.

Easterlin, R. A. (1980). *Birth and fortune: The effects of generation size on personal welfare*. New York: Basic Books.

Eckstein, D., Kunzel, V., Schafer, L., & Winges, M. (2020). *Global Climate Risk Index*, 2020. Germanwatch e.V, Bonn, Germany

Ejaz, M. (2007). Determinants of Female labor force participation in Pakistan: An empirical analysis of PSLM (2004-05) Micro data. *The Lahore Journal of Economics*, *12*(1), 203-235.

Ejaz. M., (2011). Labour force participation in Pakistan: An instrumental variable approach. Centre for Research in Economics and Business working paper No. 01-11.

EM-DAT: The Emergency Events Database - Université catholique de Louvain (UCLouvain) - CRED, D. Guha-Sapir - www.emdat.be, Brussels, Belgium

Ervin, P., Elisenda, E., Silvio, D., Benjamin, D., Paul, V. U., Stanfeld, M., ... & Gelson, T. (2017). Learning about labour impacts of cash transfers in Zambia. *Journal of African Economies*, *26*(4), 433-442.

Evertsen, K. F., & van der Geest, K. (2020). Gender, environment and migration in Bangladesh. *Climate and Development*, 12(1), 12-22.

Ezzrari, A., & Verme, P. (2012). *A multiple correspondence analysis approach to the measurement of multidimensional poverty in Morocco, 2001–2007*. The World Bank.

Faridi, M. Z., Chaudhry, I. S., & Anwar, M. (2009). The socio-economic and demographic determinants of women work participation in Pakistan: evidence from Bahawalpur District. *South Asian Studies*. Vol. 24, No. 2, pp. 353-369

Farooq, S. (2014). Effectiveness of cash transfer programmes for household welfare in Pakistan: The case of the Benazir Income Support Programme. *The Pakistan Development Review*, *53*(2), 145-174.

Fernández, R. (2013). Cultural change as learning: The evolution of female labor force participation over a century. *American Economic Review*, *103*(1), 472-500.

Fernandez, R., & Fogli, A. (2009). Culture: An empirical investigation of beliefs, work, and fertility. *American economic journal: Macroeconomics*, *1*(1), 146-77.

Fernández, R., Fogli, A., & Olivetti, C. (2004). Mothers and sons: Preference formation and female labor force dynamics. *The Quarterly Journal of Economics*, *119*(4), 1249-1299.

Ferrant, G. (2014). The Multidimensional Gender Inequalities Index (MGII): A descriptive analysis of gender inequalities using MCA. *Social indicators research*, *115*(2), 653-690.

Fortin, N. M. (2005). Gender role attitudes and the labour-market outcomes of women across OECD countries. *Oxford Review of Economic Policy*, 21(3), 416-438.

Franco, D. G. (2016). Multiple correspondence analysis: one only or several techniques?. *Quality & Quantity*, *50*(3), 1299-1315.

Fukuda, K. (2006). A cohort analysis of female labor participation rates in the US and Japan. *Review of Economics of the Household*, *4*(4), 379-393.

Gazdar, H. (2011). Social protection in Pakistan: in the midst of a paradigm shift? *Economic and Political Weekly*, 59-66.

Giavazzi, F., Schiantarelli, F., & Serafinelli, M. (2009). *Culture, policies and labor market outcomes* (No. w15417). National Bureau of Economic Research.

Gisbert, M. E., Painter, M., & Quiton, M. (1994). Gender issues associated with labor migration and dependence on off-farm income in rural Bolivia. *Human Organization*, 110-122.

Galarneau, D. (1994). *Female baby boomers: A generation at work* (Vol. 96, No. 315). Scarborough, Ont.: Prentice-Hall Canada.

Glennerster, R., Walsh, C., & Diaz-Martin, L. (2018). A practical guide to measuring women's and girls' empowerment in impact evaluations. *Gender Sector, Abdul Latif Jameel Poverty Action Lab*.

Goldin, C. (1990). *Understanding the gender gap. An economic history of American women*. Oxford: Oxford University Press.

Goldin, C., & Katz, L. F. (2000). Career and Marriage in the Age of the Pill. *American Economic Review*, 90(2), 461-465.

Government of Pakistan (2019). Pakistan Economic Survey. Finance Division, Islamabad, Pakistan

Grant, I., Kesternich, I., Steckenleiter, C., & Winter, J. (2018). Historic sex-ratio imbalances predict female participation in the market for politicians. *Journal of Economic Behavior & Organization*, *156*, 144-165.

Greene, W. H. (2008). The econometric approach to efficiency analysis. *The measurement of productive efficiency and productivity growth*, *1*(1), 92-250.

Grosjean, P., & Khattar, R. (2019). It's raining men! Hallelujah? The long-run consequences of male-biased sex ratios. *The Review of Economic Studies*, *86*(2), 723-754.

Grossbard-Shechtman, A. (1984). A theory of allocation of time in markets for labour and marriage. *The Economic Journal*, *94*(376), 863-882.

Guner, D., & Uysal, G. (2014). Culture, religiosity and female labor supply. IZA DP No. 8132

Guttentag, M., & Secord, P. F. (1983). *Too many women?: The sex ratio question*. SAGE Publications, Incorporated.

Hafeez, A., & Ahmad, E. (2002). Factors determining the labour force participation decision of educated married women in a district of Punjab. *Pakistan Economic and Social Review*, 75-88.

Hakim, A., & Aziz, A. (1998). Socio-cultural, Religious, and Political Aspects of the Status of Women in Pakistan. *The Pakistan Development Review*, Vol. 37, No. 4

Halstead, P. (1989). The economy has a normal surplus: economic stability and social change among early farming communities in Thessaly, Greece'. In: Halstead, P.,

O'Shea, J. (eds), Bad year economics: Cultural responses to risk and uncertainty, Cambridge University Press, Cambridge, pp.68-80.

Hampshire, K. (2002). Fulani on the move: seasonal economic migration in the Sahel as a social process. *Journal of Development Studies*, *38*(5), 15-36.

Hartmann, W., Nair, H. S., & Narayanan, S. (2011). Identifying causal marketing mix effects using a regression discontinuity design. *Marketing Science*, *30*(6), 1079-1097.

Haseeb, M., & Vyborny, K. (2016). *Imposing institutions: Evidence from cash transfer reform in Pakistan* (No. 2016-36). Centre for the Study of African Economies, University of Oxford.

Himmelweit, S., & Sigala, M. (2004). Choice and the relationship between identities and behaviour for mothers with pre-school children: Some implications for policy from a UK study. *Journal of Social Policy*, *33*(3), 455-478.

Hou, X. (2009). Poverty Scorecard for Pakistan-an update using the 2007-2008 Data. Technical Report, The World Bank.

Hussain, I. (2014). Urbanization in Pakistan. Keynote address delivered at South Asia Cities Conference and Pakistan Urban Forum, 9 Jan 2014.

Ibraz, T. S., Fatima, A., & Aziz, N. (1993). Uneducated and Unhealthy: The Plight of Women in Pakistan. *The Pakistan development review*, *32*(4), 905-915.

Imbens, G. W., & Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of econometrics*, *142*(2), 615-635.

Imbens, G., & Angrist, J. (1994). Identification and Estimation of Local Average Treatment Effects. *Journal of Econometrics* 61(2): 467-476.

Inglehart, Ronald, and Pippa Norris (2003) *Rising tide. Gender equality and cultural change around the world*. Cambridge: Cambridge University Press.

International Labor Organization (2016) *Key indicators of the labour market. Ninth edition*. Geneva: International Labour Office.

Iqbal, T., Farooq, S., & Padda, I. U. H. (2020). Can Empowerment be Enhanced by Putting Cash in the Hands of Poor Women? Learning from Pakistan's BISP Program. *The European Journal of Development Research*, 1-33.

Iqbal, T., Padda, I. U. H., & Farooq, S. (2020). Unconditional cash transfers and women empowerment: the case of Benazir Income Support Programme (BISP) in Pakistan. *Journal of Business and Social Review in Emerging Economies*, *6*(2), 401-418.

Irfan, M. (1986). Migration and development in Pakistan: Some selected issues. *The Pakistan Development Review*, *25*(4), 743-755.

Ishfaq, S., M. (2019). *Rural-urban migration and climate change adaptations: policy implications for Pakistan*. Sustainable Development Policy Institute, Islamabad, Pakistan.

Jacob, R., Zhu, P., Somers, M. A., & Bloom, H. (2012). *A practical guide to regression discontinuity*. MDRC.

Jalal, A. (2017). The Targeting Performance and Short-Term Welfare Effects of Female Income Support Programs: Evidence from Pakistan. Unpublished Thesis, Department of Economics, Yale University, available at https://economics.yale.edu/sites/default/files/files/Undergraduate/Nominated% 20Senior% 20Essays/2016-17/Amen_Jalal Senior% 20Essay. pdf.

Kalin, W. (2010). Conceptualizing climate-induced displacement, in J.McAdam (ed.), Climate Change and Displacement: Multidisciplinary Perspectives, Oxford and Portland, OR: Hart, pp. 81-104.

Kamanou, G. (2005). Challenges of poverty statistics in the UN Statistical Commission framework and in the present context of follow up of the UN Millennium Development Goals. *Washington DC: United Nations*.

Kandiyoti, D. (1988). Bargaining with patriarchy. *Gender & society*, *2*(3), 274-290.

Kanji, S. (2011). What keeps mothers in full-time employment? *European Sociological Review*, 27(4), 509-525.

Kawaguchi, D., & Miyazaki, J. (2009). Working mothers and sons' preferences regarding female labor supply: direct evidence from stated preferences. *Journal of Population Economics*, *22*(1), 115-130.

Kazi, S., Raza, B., & Aziz, N. (1988). Households headed by women: Income, employment and household organization. *The Pakistan Development Review*, *27*(4), 781-790.

Kesternich, I., Siflinger, B., Smith, J. P., & Steckenleiter, C. (2020). Unbalanced sex ratios in Germany caused by World War II and their effect on fertility: A life cycle perspective. *European Economic Review*, *130*, 103581.

Khan, A. H., Shehnaz, L., & Ahmed, A. M. (2000). Determinants of Internal Migration in Pakistan: Evidence from the Labour Force Survey, 1996-97. *The Pakistan Development Review*, 695-712.

Klasen, S. (2019). What explains uneven female labor force participation levels and trends in developing countries? *The World Bank Research Observer*, *34*(2), 161-197.

Kleibergen, F., & Paap, R. (2006). Generalized reduced rank tests using the singular value decomposition. Journal of econometrics, 133(1), 97-126.

Kohn, J. L. (2012). What is health? A multiple correspondence health index. *Eastern Economic Journal*, *38*(2), 223-250.

Kotkin, J., & Cox, W. (2013). The World's Fastest-Growing Megacities. Forbes (2013, April 8).

Kugelman, M. (2014). Understanding Pakistan's unstoppable urbanization. In M., Kugelman (Ed). *Pakistan's Runaway Urbanization: What Can Be Done?* (pp.1-20). Woodrow Wilson International Center for Scholars, Pennsylvania.

Le Roux, B., & Rouanet, H. (2004). *Geometric data analysis: from correspondence analysis to structured data analysis*. Springer Science & Business Media.

Lee, D. L., McCrary, J., Moreira, M. J., & Porter, J. (2020). Valid t-ratio Inference for IV. *arXiv preprint arXiv:2010.05058*.

Lee, D. S., & Lemieux, T. (2010). Regression discontinuity designs in economics. *Journal of economic literature*, *48*(2), 281-355.

Lokshin, M., & Glinskaya, E. (2009). The effect of male migration on employment patterns of women in Nepal. *The World Bank Economic Review*, *23*(3), 481-507.

Loayza, N. V., Olaberria, E., Rigolini, J., & Christiaensen, L. (2012). Natural disasters and growth: Going beyond the averages. *World Development*, *40*(7), 1317-1336.

Luke, N., & Munshi, K. (2011). Women as agents of change: Female income and mobility in India. *Journal of development economics*, *94*(1), 1-17.

Lundberg, S., & Pollak, R. A. (1996). Bargaining and distribution in marriage. *Journal of economic perspectives*, *10*(4), 139-158.

Mahmud, M., Musaddiq, T., Said, F., & Sabir, M. (2010). Internal Migration Patterns in Pakistan—The Case for Fiscal Decentralisation. *The Pakistan Development Review*, 593-607. -607.

Majid, H., & Zahir, H. (2014). Climate Change and the impact on farmer productivity: The role of socioeconomic vulnerability in rural Pakistan. *Climate Change in Pakistan: Working Paper Series*, *1*.

Massey, D. S., Axinn, W. G., & Ghimire, D. J. (2010). Environmental change and out-migration: Evidence from Nepal. *Population and environment*, *32*(2-3), 109-136.

McLeman, R. A., & Hunter, L. M. (2010). Migration in the context of vulnerability and adaptation to climate change: insights from analogues. *Wiley Interdisciplinary Reviews: Climate Change*, *1*(3), 450-461.

McLeman, R., & Smit, B. (2006). Migration as an adaptation to climate change. *Climatic change*, *76*(1), 31-53.

Merola, G. M., & Baulch, B. (2019). Using sparse categorical principal components to estimate asset indices: new methods with an application to rural Southeast Asia. *Review of Development Economics*, *23*(2), 640-662.

Mincer, Jacob (1962). Labor force participation of married women. 63-97 in *Aspects of labor economics*, edited by H.G. Lewis. Princeton, N.J. Princeton University Press.

Morgan, C. (2009). Climate change, uncertainty and prehistoric hunter-gatherer mobility. *Journal of Anthropological Archaeology*, 28(4): 382-396.

Morrill, M. S., & Morrill, T. (2013). Intergenerational links in female labor force participation. *Labour Economics*, 20, 38-47.

Mueller, V., Gray, C., & Kosec, K. (2014). Heat stress increases long-term human migration in rural Pakistan. *Nature climate change*, *4*(3), 182-185.

Naqvi, Z. F., Shahnaz, L., & Arif, G. M. (2002). How do women decide to work in Pakistan? [with comments]. *The Pakistan development review*, 495-513.

Nordenmark, M. (2004). Does gender ideology explain differences between countries regarding the involvement of women and of men in paid and unpaid work? *International Journal of Social Welfare*, 13(3), 233-243.

Pakistan Bureau of Statistics (2015), Pakistan Labor Force Survey, 2014-15

Pampel, F. C., & Tanaka, K. (1986). Economic development and female labor force participation: A reconsideration. *Social forces*, *64*(3), 599-619.

Parker, S. W., & Todd, P. E. (2017). Conditional cash transfers: The case of Progresa/Oportunidades. *Journal of Economic Literature*, *55*(3), 866-915.

Parry, M., Parry, M. L., Canziani, O., Palutikof, J., Van der Linden, P., & Hanson, C. (Eds.). (2007). *Climate change 2007-impacts, adaptation and vulnerability: Working group II contribution to the fourth assessment report of the IPCC* (Vol. 4). Cambridge University Press.

Peterman, A., Schwab, B., Roy, S., Hidrobo, M., & Gilligan, D. O. (2021). Measuring women's decisionmaking: Indicator choice and survey design experiments from cash and food transfer evaluations in Ecuador, Uganda and Yemen. *World Development*, *141*, 105387.

Pretis, F. (2021). Exogeneity in climate econometrics. *Energy Economics*, *96*, 105122.

Raddatz, C. (2009). *The wrath of God: macroeconomic costs of natural disasters*. The World Bank.

Richerson, P. J., Boyd, R., & Bettinger, R. L. (2001). Was agriculture impossible during the Pleistocene but mandatory during the Holocene? A climate change hypothesis. *American Antiquity*, *66*(3), 387-411.

Ridgeway, C. L., & Correll, S. J. (2004). Unpacking the gender system: A theoretical perspective on gender beliefs and social relations. *Gender & society*, *18*(4), 510-531.

Safana, S., Masood, S., Muhammad, W., & Amir, A. (2012). *Economic Analysis of Earnings in Pakistan: A Case of Sargodha District*. MPRA Paper No. 41923.

Schultz, K. A., & Mankin, J. S. (2019). Is temperature exogenous? The impact of civil conflict on the instrumental climate record in Sub-Saharan Africa. *American Journal of Political Science*, *63*(4), 723-739.

Seguino, S. (2007). PlusÇa Change? Evidence on global trends in gender norms and stereotypes. *Feminist Economics*, *13*(2), 1-28.

Seplaki, C. L., Agree, E. M., Weiss, C. O., Szanton, S. L., Bandeen-Roche, K., & Fried, L. P. (2014). Assistive devices in context: cross-sectional association between challenges in the home environment and use of assistive devices for mobility. *The Gerontologist*, 54(4), 651-660.

Shah, N. M. (1986), Changes in female role in Pakistan: Are the volume and pace adequate? *The Pakistan Development Review*, Volume 25(3).

Sørensen, N. N. (2004). Migrant transfers as a development tool. *DIIS (Dansk Institut For Internationale Studier). Copenhagen*.

South, S. J., & Trent, K. (1988). Sex ratios and women's roles: A cross-national analysis. *American Journal of sociology*, *93*(5), 1096-1115.

Stähli, E. M., Le Goff, J. M., Levy, R., & Widmer, E. (2009). Wishes or constraints? Mothers' labour force participation and its motivation in Switzerland. *European Sociological Review*, 25(3), 333-348.

Staiger, D., & Stock, J. H. (1997). Instrumental variables regression with weak instruments. *Econometrica: journal of the Econometric Society*, 557-586.

Stam, K., Verbakel, E., & de Graaf, P. M. (2014). Do values matter? The impact of work ethic and traditional gender role values on female labour market supply. *Social Indicators Research*, 116(2), 593-610.

Stark, O. (1991). Migration in LDCs: risk, remittances, and the family. *Finance and Development*, *28*(4), 39.

Steinmann, S. H. (1993) Effects of International Migration on Women's Work in Agriculture. *Revue de Géographie du Maroc*, 15(1-2): 105-124.

Stock, J. H., & Yogo, M. (2005). Testing for weak instruments in linear IV regression. Chapter 5 In J. Rothenberg, J. H. Stock, & D. W. K. Andrews (Eds.), Identification and Inference for Econometric Models: Essays in Honor of Thomas (pp. 80–108). NY: Cambridge University Press.

Sultana, H., Ali, N., Iqbal, M. M., & Khan, A. M. (2009). Vulnerability and adaptability of wheat production in different climatic zones of Pakistan under climate change scenarios. *Climatic Change*, *94*(1), 123-142.

Tacoli, C. (1999). International migration and the restructuring of gender asymmetries: Continuity and change among Filipino labor migrants in Rome. *International migration review*, *33*(3), 658-682.

Tacoli, C. (2009). Crisis or adaptation? Migration and climate change in a context of high mobility. *Environment and urbanization*, *21*(2), 513-525.

Teso, E. (2019). The long-term effect of demographic shocks on the evolution of gender roles: Evidence from the transatlantic slave trade. *Journal of the European Economic Association*, *17*(2), 497-534.

Thistlethwaite, D. L., & Campbell, D. T. (1960). Regression-discontinuity analysis: An alternative to the ex post facto experiment. *Journal of Educational psychology*, *51*(6), 309.

Tuccio, Michele, and Jackline Wahba. "Return migration and the transfer of gender norms: evidence from the Middle East." *Journal of Comparative Economics* 46, no. 4 (2018): 1006-1029.

United Nations (2015). Human Development Report 2015, United Nations Development Program, New York, USA.

Ward, K. B., & Pampel, F. C. (1985a). Structural determinants of female labor force participation in developed nations, 1955–75. *Social Science Quarterly*, 66(3), 654–667.

Ward, K. B., & Pampel, F. C. (1985b). More on the meaning of the effect of the sex ratio on female labor force participation. *Social Science Quarterly*, 66(3), 675–679.

Emery, J. H., & Ferrer, A. (2009). Marriage market imbalances and labor force participation of Canadian women. *Review of Economics of the Household*, *7*(1), 43-57.

Warner, K. (2010). Global environmental change and migration: Governance challenges. *Global environmental change*, *20*(3), 402-413.

Wold Bank (2013). *Pakistan: Towards an integrated national safety net system*. Report No. 66421-PK

Wooldridge, J. M. (2002). Introductory Econometrics: A Modern Approach, 2003. *New York: South-Western College Publishing*.

Wooldridge, J. M. (2013). Introductory Econometrics: A Modern Approach. *South-Western Cengage Learning*

Zuneira, S. Amjad., Usman M., & Farooq S. (2018) Targeting and Effectiveness of Social Safety Net Programmes: The Case of Zakat and BISP in Pakistan. *Nust Journal of Social Sciences and Humanities*. Volume 4, Number 2.

Appendix 1: Figures and Tables

Figure 1 Comparison of OPM Poverty Score and Replicated Poverty Score

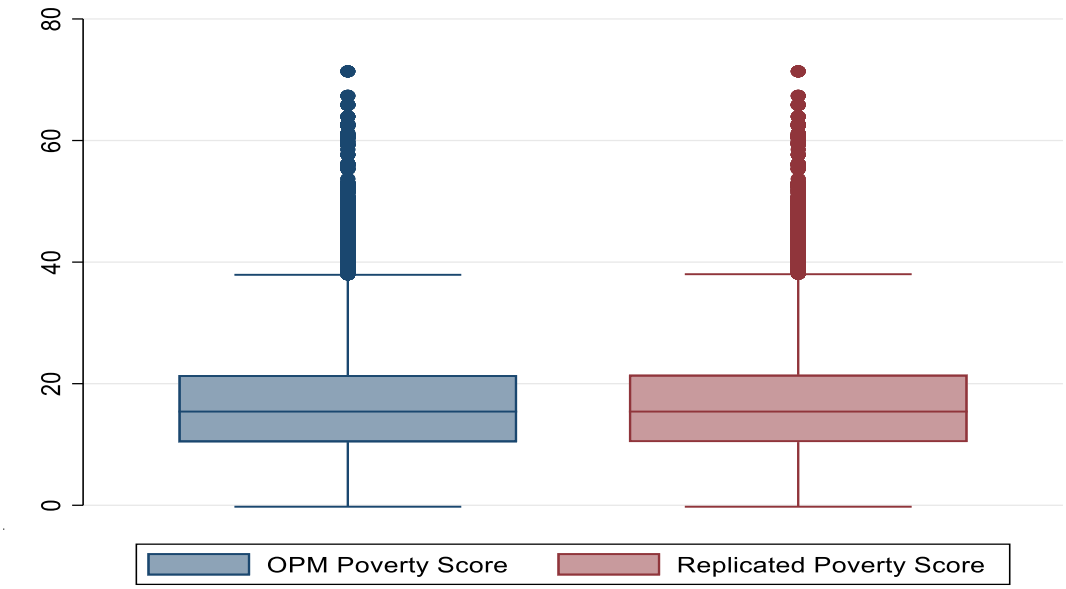


Table 1 Classification of Disasters by EM-DAT

Disaster Group	Disaster Subgroup	Definition	Disaster Main Type
Natural	Geophysical	A hazard originating from solid earth. This term is used interchangeably with the term geological hazard.	Earthquake
			Mass Movement
			Volcanic Activity
	Meteorological	A hazard caused by short-lived, micro- to meso-scale extreme weather and atmospheric conditions that last from minutes to days.	Extreme Temperature
			Fog
			Storm
	Hydrological	A hazard caused by the occurrence, movement, and distribution of surface and subsurface freshwater	Flood
			Landslide

		and saltwater.	Wave Action
	Climatological	A hazard caused by long-lived, meso-to macro-scale atmospheric processes ranging from intra-seasonal to multi-decadal climate variability.	Drought
			Glacial Lake Outburst
			Wildfire
	Biological	A hazard caused by the exposure to living organisms and their toxic substances	Epidemic
			Insect Infestation
			Animal Accident
	Extraterrestrial	A hazard caused by asteroids, meteoroids, and comets as they pass near-earth, enter the Earth's atmosphere, and/or strike the Earth, and by changes in interplanetary conditions that effect the Earth's magnetosphere, ionosphere, and thermosphere.	Impact
			Space Weather
Technological	Industrial Accidents		Chemical Spill
			Collapse
			Explosion
			Fire
			Gas Leak
			Poisoning
			Radiation
			Oil Spill
			Other
	Transport Accidents		Air
			Road
			Rail
			Water
	Miscellaneous Accidents		Collapse
			Explosion
			Fire
			Other

Source: https://www.emdat.be/classification

Table 2 Women's decision making in household

	2011 Control	2011 Treated	2013 Control	2013 Treated	2014 Control	2014 Treated	2016 Control	2016 Treated
Have another child								
	0.68	0.66	0.83	0.83	0.85	0.82	0.81	0.81
Children's education								

	0.68	0.65	0.85	0.84	0.85	0.84	0.83	0.83
Children's marriage plans								
	0.67	0.64	0.83	0.82	0.85	0.83	0.83	0.86
Use of family planning methods								
	0.61	0.61	0.83	0.84	0.84	0.82	0.78	0.79
Visit to friends								
	0.63	0.61	0.84	0.83	0.83	0.81	0.79	0.81
Minor household purchases								
	0.59	0.57	0.79	0.76	0.79	0.76	0.73	0.76
What kind of job she can do								
	0.53	0.52	0.77	0.79	0.75	0.72	0.66	0.70
Lending or borrowing								
	0.54	0.53	0.75	0.73	0.73	0.74	0.68	0.74
Small Investment								
	0.50	0.50	0.73	0.71	0.66	0.67	0.55	0.62
Own and child's health treatment								
	0.60	0.58	0.84	0.83	0.81	0.81	0.83	0.86
Participation in a group outside home								
	0.48	0.47	0.65	0.61	0.61	0.60	0.56	0.59
Vote in elections								
	0.49	0.47	0.68	0.66	0.67	0.64	0.60	0.65
Observations	3627	3758	4589	4642	4291	4150	3207	4240

Source: BISP Survey, own calculations

Table 3 Women's freedom of mobility

	2011 Control	2011 Treated	2013 Control	2013 Treated	2014 Control	2014 Treated	2016 Control	2016 Treated
Permitted to go to Local Market Alone								
	0.30	0.26	0.35	0.29	0.35	0.34	0.37	0.40
Permitted to go to Health Facility Alone								
	0.32	0.29	0.40	0.33	0.38	0.37	0.36	0.40
Permitted to go to Friend's Home Alone								
	0.43	0.40	0.51	0.46	0.49	0.50	0.52	0.58
Permitted to go to Religious Place Alone								
	0.28	0.25	0.31	0.25	0.31	0.30	0.26	0.29
Vote in Elections								
	0.58	0.53	0.67	0.67	0.73	0.73	0.75	0.89
Observations	5164	5158	5154	5044	4663	4448	3856	5175

Source: BISP Survey, own calculations

Table 4 Women's Access to Money

	Control	Treated	Control	Treated	Control	Treated	Control	Treated
	2011		2013		2014		2016	

Access to Rs. 50	0.80	0.75	0.90	0.88	0.93	0.91	0.83	0.81
Access to Rs. 100	0.68	0.60	0.82	0.79	0.88	0.85	0.75	0.73
Access to Rs. 200	0.50	0.42	0.67	0.61	0.74	0.69	0.60	0.60
Access to Rs. 400	0.37	0.30	0.50	0.45	0.59	0.53	0.46	0.47
Access to Rs. 600	0.28	0.20	0.40	0.34	0.48	0.42	0.36	0.37
Access to Rs. 800	0.23	0.16	0.35	0.30	0.43	0.37	0.31	0.31
Access to Rs. 1000	0.22	0.15	0.34	0.29	0.41	0.36	0.29	0.29
Observations	5162	5156	5124	5016	4625	4416	3859	5178

Source: BISP Survey, own calculations

Table 5 Weights of Poverty Indicators from PSLM

Poverty Indicators	**Weights**
Less than or equal to 2 Dependents	13.34
3 or 4 Dependents	8.19
5 or 6 Dependents	4.3
Head completed less than class 1 to 5	1.63
Head completed class 6 to 10	2.38
Head completed class 11, or beyond	10
All children 5-16 years old are attending school	5.62
Some children 5-16 years old are attending school	2.65
Ratio of room per person greater than 0.2 and less than equal to 0.3	2.23
Ratio of room per person greater than 0.3 and less than equal to 0.4	4.94
Ratio of room per person greater than 0.4	14.18
Flush connected to public sewerage, to a pit or to an open drain system	1.60
Dry raised latrine or dry pit latrine	-0.24
At least one refrigerator, freezer or washing machine	2.46
At least one AC, air cooler, geyser or heater	7.04
At least one cooking stove, cooking range, or microwave oven	5.86
At least one TV	1.21
One car or one car with one motorcycle	22.71

One motorcycle but no car	6.04
At least one buffalo or bullock	4.42
No buffalo or bullock but at least one cow or goat or sheep	0.26
Agriculture land less than equal to 12.5 acres	2.02
Agriculture land greater than 12.5 acres	6.73

Table 6 MCA Eigenvalues and Inertia of Husband-wife gender norms coincidence 2011-2014 Panel Sample

	Preliminary Eigenvalues	Preliminary Inertia	Final Eigenvalues	Final Inertia
Dimension 1	0.23	23.48	0.34	34.41
Dimension 2	0.17	17.26	0.25	24.90
Dimension 3	0.17	16.95	0.20	20.69
Dimension 4	0.15	15.34	0.20	19.99
Dimension 5	0.13	13.65	0.00	0.00
Dimension 6	0.13	13.29	0.00	0.00

Table 7 MCA Dimensions and Contributions of Husband-wife gender norms coincidence 2011-2014 Panel Sample

	Coordinates Dimension 1	Contributions Dimension 1	Coordinates Dimension 2	Contributions Dimension 2
Both men and women should make HH decisions				
Conflict	1.02	0.02	2.21	0.09
Coincidence	-0.40	0.01	-0.86	0.03
Husband should help in HH				
Conflict	-1.76	0.09	0.83	0.01
Coincidence	1.01	0.05	-0.47	0.02
Women can work outside home				
Conflict	-1.35	0.07	0.89	0.02
Coincidence	1.22	0.06	-0.81	0.02
Wife should express opinion				
Conflict	-1.93	0.09	-0.006	0.00
Coincidence	0.92	0.04	0.003	0.00
Women should not tolerate violence				
Conflict	0.07	0.00	-0.47	0.005
Coincidence	-0.03	0.00	0.21	0.002
Better to send daughter to school rather than son				
Conflict	-0.87	0.02	-2.36	0.13
Coincidence	0.44	0.01	1.21	0.07

Table 8 Cronbach's Alpha: Husband-wife gender norms coincidence 2011-2016 Panel Sample

	Item-test Correlation	Item-rest Correlation	Average Interitem Covariance	Cronbach Alpha
Both men and women should make HH decisions				

0.43	0.10	0.01	0.25
Husband should help in HH			
0.51	0.17	0.01	0.19
Women can work outside home			
0.51	0.15	0.01	0.20
Wife should express opinion			
0.52	0.19	0.01	0.17
Women should not tolerate violence			
0.355	0.014	0.019	0.313
Better to send daughter to school rather than son			
0.425	0.072	0.015	0.270
Test Scale		0.013	0.272

Table 9 MCA Eigenvalues and Inertia of Men gender norms 2011-2016 Panel Sample

	Preliminary Eigenvalues	Preliminary Inertia	Final Eigenvalues	Final Inertia
Dimension 1	0.29	29.27	0.55	54.94
Dimension 2	0.19	18.80	0.25	24.96
Dimension 3	0.15	15.64	0.20	20.09
Dimension 4	0.14	14.18	0.00	0.00
Dimension 5	0.12	12.12	0.00	0.00
Dimension 6	0.10	9.98	0.00	0.00

Table 10 MCA Dimensions and Contributions of Men gender norms 2011-2016 Panel Sample

	Coordinates Dimension 1	Contributions Dimension 1	Coordinates Dimension 2	Contributions Dimension 2
Both men and women should make HH decisions				
Disagree	0.36	0.01	0.56	0.02
Agree	-1.53	0.04	-2.52	0.09
Husband should help in HH				
Disagree	2.16	0.12	-0.75	0.01
Agree	-0.86	0.045	0.3	0.01
Women can work outside home				
Disagree	1.19	0.07	-0.46	0.01
Agree	-1.40	0.08	0.55	0.01
Wife should express opinion				
Disagree	2.11	0.10	-0.75	0.01
Agree	-0.70	0.03	0.25	0.003
Women should not tolerate violence				
Disagree	0.05	0.00	0.84	0.04
Agree	-0.17	0.001	-2.82	0.13
Better to send daughter to school rather than son				
Disagree	1.05	0.03	2.00	0.08
Agree	-0.38	0.01	-0.73	0.03

Table 11 MCA Dimensions and Contributions of Men gender norms 2011-2014 Panel Sample

	Coordinates Dimension 1	Contributions Dimension 1	Coordinates Dimension 2	Contributions Dimension 2
Both men and women should make HH decisions				
Disagree	0.32	0.008	0.59	0.02
Agree	-1.47	0.03	-2.73	0.09
Husband should help in HH				
Disagree	2.08	0.12	-0.60	0.00
Agree	-0.90	0.05	0.26	0.001
Women can work outside home				
Disagree	1.16	0.07	-0.40	0.01
Agree	-1.46	0.08	0.51	0.01
Wife should express opinion				
Disagree	1.94	0.01	-0.52	0.006
Agree	-0.78	0.04	0.21	0.002
Women should not tolerate violence				
Disagree	-0.02	0.00	0.95	0.05
Agree	0.07	0.00	-2.75	0.14
Better to send daughter to school rather than son				
Disagree	1.081	0.03	1.88	0.07
Agree	-0.394	0.01	-0.68	0.02

Table 12 MCA Eigenvalues and Inertia Women's Mobility 2011-2016 Panel Sample

	Eigenvalues	Inertia
Dimension 1	0.65	64.69
Dimension 2	0.19	18.89
Dimension 3	0.08	8.28
Dimension 4	0.05	5.22
Dimension 5	0.03	2.91

Table 13 MCA Dimensions and Contributions of Women's Mobility 2011-2016 Panel Sample

	Coordinates Dimension 1	Contributions Dimension 1	Coordinates Dimension 2	Contributions Dimension 2
Permitted to go to Local Market Alone				
No	0.80	0.07	0.13	0.001
Yes	-1.64	0.14	-0.27	0.002
Permitted to go to Health Facility Alone				
No	0.85	0.07	0.16	0.001
Yes	-1.56	0.14	-0.29	0.000
Permitted to go to Friend's Home Alone				
No	1.00	0.08	0.06	0.000
Yes	-1.05	0.08	-0.06	0.000
Permitted to go to Religious Place Alone				
No	0.67	0.05	0.14	0.001

Yes	-1.77	0.14	-0.36	0.003
Vote in Elections				
No	0.53	0.01	-3.33	0.29
Yes	-0.23	0.006	1.46	0.13

Table 14 MCA Eigenvalues and Inertia Women's Decision Making in Household 2011-2016 Panel Sample

	Eigenvalues	Inertia
Dimension 1	0.81	81.11
Dimension 2	0.07	6.90
Dimension 3	0.03	2.67
Dimension 4	0.02	1.85
Dimension 5	0.01	1.62
Dimension 6	0.01	1.39
Dimension 7	0.01	1.15
Dimension 8	0.01	0.83
Dimension 9	0.01	0.76
Dimension 10	0.01	0.65
Dimension 11	0.01	0.56
Dimension 12	0.005	0.49

Table 15 MCA Dimensions and Contributions of Women's Decision Making in Household 2011-2016 Panel Sample

	Coordinates Dimension 1	Contributions Dimension 1	Coordinates Dimension 2	Contributions Dimension 2
Have another child				
No	1.30	0.04	1.89	0.03
Yes	-0.74	0.02	-1.08	0.01
Children's education				
No	1.33	0.05	1.74	0.02
Yes	-0.74	0.02	-0.97	0.02
Children's marriage plans				
No	1.31	0.05	1.69	0.02
Yes	-0.76	0.03	-0.98	0.01
Use of family planning methods				
No	1.23	0.04	1.68	0.02
Yes	-0.74	0.02	-1.02	0.01
Visit to friends				
No	1.29	0.05	0.32	0.001
Yes	-0.80	0.03	-0.20	0.001
Minor household purchases				
No	1.247	0.05	-0.39	0.001
Yes	-0.84	0.03	0.26	0.001
What kind of job she can do				
No	1.21	0.04	-0.90	0.007
Yes	-0.79	0.03	0.59	0.005
Lending or borrowing				

No	1.14	0.04	-1.08	0.01
Yes	-0.90	0.03	0.85	0.01
Small Investment				
No	1.09	0.04	-1.05	0.01
Yes	-0.95	0.03	0.91	0.01
Own and child's health treatment				
No	1.25	0.04	-0.62	0.003
Yes	-0.79	0.03	0.39	0.002
Participation in a group outside home				
No	1.04	0.04	-1.06	0.01
Yes	-0.96	0.03	0.97	0.01
Vote in elections				
No	1.04	0.04	-1.06	0.01
Yes	-0.92	0.03	0.93	0.01

Table 16 MCA Eigenvalues and Inertia Women's Mobility 2011-2014 Panel Sample

	Eigenvalues	Inertia
Dimension 1	0.66	66.08
Dimension 2	0.19	18.90
Dimension 3	0.07	7.63
Dimension 4	0.04	4.56
Dimension 5	0.03	2.82

Table 17 MCA Dimensions and Contributions of Women's Mobility 2011-2014 Panel Sample

	Coordinates Dimension 1	Contributions Dimension 1	Coordinates Dimension 2	Contributions Dimension 2
Permitted to go to Local Market Alone				
No	0.77	0.06	0.12	0.001
Yes	-1.68	0.14	-0.26	0.002
Permitted to go to Health Facility Alone				
No	0.84	0.07	0.13	0.001
Yes	-1.56	0.14	-0.25	0.001
Permitted to go to Friend's Home Alone				
No	0.96	0.08	0.14	0.001
Yes	-1.10	0.09	-0.16	0.001
Permitted to go to Religious Place Alone				
No	0.69	0.05	0.10	0.001
Yes	-1.77	0.14	-0.27	0.002
Vote in Elections				
No	0.46	0.01	-2.00	0.27
Yes	-0.25	0.007	1.63	0.15

Table 18 MCA Eigenvalues and Inertia Women's Decision Making in Household 2011-2014 Panel Sample

	Eigenvalues	Inertia

Dimension 1	0.83	82.90
Dimension 2	0.07	7.11
Dimension 3	0.02	2.48
Dimension 4	0.01	1.73
Dimension 5	0.01	1.39
Dimension 6	0.01	1.18
Dimension 7	0.01	1.007
Dimension 8	0.006	0.63
Dimension 9	0.005	0.54
Dimension 10	0.004	0.38
Dimension 11	0.003	0.33
Dimension 12	0.003	0.30

Table 19 MCA Dimensions and Contributions of Women's Decision Making in Household 2011-2014 Panel Sample

	Coordinates Dimension 1	Contributions Dimension 1	Coordinates Dimension 2	Contributions Dimension 2
Have another child				
No	1.27	0.04	1.85	0.03
Yes	-0.76	0.02	-1.10	0.01
Children's education				
No	1.29	0.04	1.70	0.02
Yes	-0.78	0.03	-1.01	0.02
Children's marriage plans				
No	1.28	0.04	1.68	0.02
Yes	-0.78	0.03	-1.02	0.01
Use of family planning methods				
No	1.20	0.04	1.69	0.02
Yes	-0.75	0.02	-1.05	0.01
Visit to friends				
No	1.25	0.04	0.29	0.001
Yes	-0.83	0.03	-0.19	0.001
Minor household purchases				
No	1.21	0.04	-0.45	0.002
Yes	-0.85	0.03	0.31	0.001
What kind of job she can do				
No	1.19	0.04	-0.96	0.008
Yes	-0.80	0.03	0.65	0.006
Lending or borrowing				
No	1.11	0.04	-1.02	0.01
Yes	-0.94	0.03	0.87	0.01
Small Investment				
No	1.08	0.04	-1.04	0.01
Yes	-0.96	0.03	0.92	0.010
Own and child's health treatment				
No	1.20	0.04	-0.60	0.003
Yes	-0.83	0.03	0.41	0.002
Participation in a group outside home				
No	1.02	0.04	-1.02	0.01
Yes	-0.98	0.03	0.97	0.01

Vote in elections				
No	1.03	0.04	-0.93	0.01
Yes	-0.98	0.04	0.88	0.01

Table 20 MCA Eigenvalues and Inertia Women's Access to money 2011-2016 Panel Sample

	Eigenvalues	Inertia
Dimension 1	0.65	65.50
Dimension 2	0.12	19.70
Dimension 3	0.06	6.86
Dimension 4	0.03	3.42
Dimension 5	0.02	2.58
Dimension 6	0.01	1.53
Dimension 7	0.004	0.42

Table 21 MCA Dimensions and Contributions of Women's Access to money 2011-2016 Panel Sample

	Coordinates Dimension 1	Contributions Dimension 1	Coordinates Dimension 2	Contributions Dimension 2
Access to Rs 50				
No	1.70	0.05	3.76	0.13
Yes	-0.29	0.01	-0.64	0.02
Access to Rs 100				
No	1.53	0.06	2.55	0.09
Yes	-0.47	0.02	-0.78	0.01
Access to Rs 200				
No	1.25	0.07	0.78	0.01
Yes	-0.81	0.04	-0.51	0.01
Access to Rs 400				
No	1.02	0.06	-0.16	0.001
Yes	-1.17	0.07	0.18	0.001
Access to Rs 600				
No	0.84	0.05	-0.54	0.01
Yes	-1.47	0.09	0.96	0.02
Access to Rs 800				
No	0.74	0.04	-0.60	0.01
Yes	-1.60	0.09	1.31	0.03
Access to Rs 1000				
No	0.70	0.04	-0.60	0.01
Yes	-1.62	0.09	1.37	0.03

Table 22 MCA Eigenvalues and Inertia Women's Access to money 2011-2014 Panel Sample

	Eigenvalues	Inertia
Dimension 1	0.65	64.97
Dimension 2	0.20	19.89
Dimension 3	0.07	7.03
Dimension 4	0.03	3.54
Dimension 5	0.02	2.67

Dimension 6	0.01	1.50
Dimension 7	0.004	0.38

Table 23 MCA Dimensions and Contributions of Women's Access to money 2011-2014 Panel Sample

	Coordinates Dimension 1	Contributions Dimension 1	Coordinates Dimension 2	Contributions Dimension 2
Access to Rs 50				
No	1.70	0.04	3.90	0.13
Yes	-0.27	0.00	-0.63	0.02
Access to Rs 100				
No	1.52	0.06	2.60	0.10
Yes	-0.46	0.02	-0.78	0.01
Access to Rs 200				
No	1.23	0.07	0.77	0.01
Yes	-0.81	0.04	-0.51	0.01
Access to Rs 400				
No	1.003	0.06	-0.16	0.001
Yes	-1.12	0.07	0.19	0.001
Access to Rs 600				
No	0.82	0.05	-0.52	0.01
Yes	-1.51	0.09	0.96	0.02
Access to Rs 800				
No	0.73	0.04	-0.58	0.01
Yes	-1.64	0.09	1.30	0.03
Access to Rs 1000				
No	0.70	0.04	-0.57	0.01
Yes	-1.66	0.09	1.36	0.03

Table 24 Summary of all Variables: PSLM and EM-DAT

	2005-06	2007-08	2011-12	2013-14
Female Labor Force Participation	0.19	0.17	0.17	0.20
Age	23.21	23.48	23.90	23.39
Employment Decision	1.83	1.80	1.78	1.76
Permission to Work	1.54	1.60	1.60	1.57
Highest Class Completed	1.11	1.17	1.30	1.19
Marital Status	1.42	1.43	1.44	1.44
HH Head Education	1.43	1.51	1.60	1.43
HH Head Labor Market	1.88	1.87	1.86	1.83

Status				
HH Head Employment Status	3.31	3.31	3.22	3.34
Female Head of Household	0.03	0.04	0.04	0.07
Dependency Ratio	0.63	0.61	0.60	0.65
Household Size	9.16	8.68	8.47	8.23
Kids: Age<10	2.75	2.53	2.42	2.46
Infants: Age<=2	0.72	0.65	0.64	0.67
Family Structure	1.57	1.54	1.54	1.52
Log Monthly Income	8.68	8.98	9.60	9.68
Working age Gender Ratio	1.04	1.06	1.05	0.99
Exposure to Extreme Weather	0.15	0.15	0.15	0.15
Observations	106101	102016	100024	119109

Source: PSLM Surveys and EM-DAT, own calculations

Table 25 Summary of all Variables: BISP

	2011	2013	2014	2016
FLFP	0.22	0.21	0.24	0.23
Husband-wife gender norms coincidence Index	2.88	2.91	2.84	3.05
Men gender norms index	2.15	2.22	2.16	2.41
Age	20.93	22.08	22.41	22.70
Completed Years of Education	2.22	2.36	2.49	2.78
Household Head Employment Status	4.11	4.01	3.93	4.02
Head Years of Education	2.33	2.42	2.68	2.71

Number of Kids	4.37	4.11	3.90	3.88
Number of working age people in Household	4.23	4.64	4.78	4.70
Dependency Ratio	1.42	1.23	1.13	1.19
Remittances received	0.14	0.18	0.21	0.16
Ownership of agri, non-agri, commercial and residential land	0.79	0.84	0.82	0.62
Access to Money Index	1.05	1.43	1.62	1.41
Observations	64120	65435	62202	87386

Source: BISP Survey, own calculations

Table 26 Full Estimation Results - PSLM

	Robust SEs		Clustered SEs	
VARIABLES	OLS	IV	OLS	IV
Working age Gender Ratio	-0.051***	0.37***	-0.051	0.37
	(0.01)	(0.095)	(0.044)	(0.26)
Age	-0.002	0.012***	-0.002	0.012
	(0.001)	(0.003)	(0.002)	(0.01)
Age Square	0.00005***	-0.0001***	0.00005	-0.0001
	(0.000022)	(0.00005)	(0.00004)	(0.000160)
Education = 2, Primary and Less	-0.06***	-0.05***	-0.06***	-0.05***
	(0.004)	(0.005)	(0.01)	(0.01)
Education = 3, Higher Secondary School	-0.08***	-0.07***	-0.08***	-0.07***
	(0.003)	(0.004)	(0.01)	(0.01)
Education = 4, Bachelor and Above	-0.04***	-0.04***	-0.04***	-0.04**
	(0.007)	(0.007)	(0.01)	(0.01)
Head Education = 2, Primary and Less	-0.006	-0.008*	-0.006	-0.008
	(0.004)	(0.004)	(0.008)	(0.01)
Head Education = 3, Higher Secondary School	-0.031***	-0.030***	-0.031***	-0.030***
	(0.003)	(0.003)	(0.007)	(0.007)
Head Education = 4, Bachelor and Above	-0.01**	-0.02***	-0.01	-0.02***

	(0.004)	(0.005)	(0.008)	(0.008)
HH Head Employment Status = 2, Self Employed non Agri	-0.0007	0.0004	-0.0007	0.0004
	(0.006)	(0.006)	(0.009)	(0.009)
HH Head Employment Status = 3, Paid Employee	0.002	-0.007	0.002	-0.007
	(0.006)	(0.006)	(0.008)	(0.01)
HH Head Employment Status = 4, Contributing Family Worker	0.08***	0.1***	0.08*	0.1**
	(0.03)	(0.03)	(0.04)	(0.04)
HH Head Employment Status = 5, Own Account Worker	0.17***	0.18***	0.17***	0.18***
	(0.007)	(0.007)	(0.02)	(0.02)
Marital Status = 2, Currently Married	-0.02***	-0.03***	-0.02***	-0.03***
	(0.005)	(0.005)	(0.007)	(0.008)
Marital Status = 3, Widow	0.0006	-0.01	0.0006	-0.01
	(0.02)	(0.02)	(0.02)	(0.02)
Marital Status = 4, Divorced	-0.02	-0.01	-0.02	-0.01
	(0.02)	(0.02)	(0.03)	(0.03)
Dependency Ratio	0.009*	0.01***	0.009	0.01*
	(0.005)	(0.005)	(0.007)	(0.007)
Household Size	-0.004***	-0.004***	-0.004***	-0.004***
	(0.0007)	(0.0007)	(0.001)	(0.001)
Kids: Age<10	0.01***	0.01***	0.01**	0.01**
	(0.002)	(0.002)	(0.004)	(0.004)
Infants: Age<=2	-0.01***	-0.01***	-0.01***	-0.01*
	(0.002)	(0.002)	(0.003)	(0.004)
Family Structure = 2, Extended	0.005	0.01***	0.005	0.01
	(0.003)	(0.003)	(0.005)	(0.006)
Female Head of Household	0.21***	0.23***	0.21***	0.23***
	(0.01)	(0.02)	(0.02)	(0.03)
Log Monthly Income	-0.02***	-0.02***	-0.02***	-0.02***
	(0.002)	(0.002)	(0.006)	(0.01)
Employment Decision = 2, Traditional	-0.02***	-0.02***	-0.02**	-0.02***
	(0.004)	(0.004)	(0.01)	(0.01)
Permission to Work = 2, Not Permitted	-0.04***	-0.03***	-0.04***	-0.03***
	(0.003)	(0.003)	(0.01)	(0.01)
Permission to Work = 3, Inactive	-0.03***	-0.02***	-0.03***	-0.02**
	(0.005)	(0.005)	(0.01)	(0.01)
Constant	0.45***	-0.15	0.45***	-0.15
	(0.03)	(0.13)	(0.01)	(0.36)
Observations	43,879	43,879	43,879	43,879
R-squared	0.12	0.08	0.12	0.08

Table 27 Robustness Check: OLS and IV regressions on FLFP with Exposure to extreme weather events as IV

	Robust SEs		Clustered SEs	
VARIABLES	OLS	IV-II	OLS	IV-II
Working age Gender Ratio	-0.05***	0.54***	-0.04	0.54
	(0.01)	(0.18)	(0.038)	(0.45)
Constant	0.44***	-0.38	0.44***	-0.38
	(0.03)	(0.25)	(0.09)	(0.63)
Observations	43,879	43,879	43,879	43,879
R-squared	0.12	0.02	0.125	0.0

Regressions include control variables, age, age square, education, marital status, head education, head labor market status, dependency ratio, household size, number of kids, number of infants, family structure, female headed households, log of monthly income, employment decision and permission to work. *** p<0.01, ** p<0.05, * p<0.1

Table 28 Robustness Check: First Stage Results

	(1)	(2)
VARIABLES	Working age gender ratio - Rboust SEs	Working age gender ratio - Clustered SEs
Exposure to Extreme Weather	0.11***	0.11*
	(0.01)	(0.06)
Constant	1.4***	1.4***
	(0.01)	(0.07)
Observations	43,879	43,879
Kleibergen-Paap rk LM statistic	110.9	3.63
Kleibergen-Paap rk LM P-value	0.00	0.05
First stage F-Statistic	111.8	3.65
Kleibergen-Paap rk Wald F statistic	111.8	3.65
Endogeneity Test Statistic	12.22	2.59
Endogeneity Test P-value	0.0004	0.10

Regressions include control variables, age, age square, education, marital status, head education, head labor market status, dependency ratio, household size, number of kids, number of infants, family structure, female headed households, log of monthly income, employment decision and permission to work. *** p<0.01, ** p<0.05, * p<0.1

Table 29 Full Estimation Results - BISP Data

	3-Year Panel Sample		4-Year Panel Sample	
VARIABLES	OLS	REIV	OLS	REIV
FLFP	-0.04**	-0.70***	-0.02	-0.23
	(0.02)	(0.2)	(0.02)	(0.35)
Husband gender norms index	0.52***	0.53***	0.50***	0.50***
	(0.01)	(0.01)	(0.01)	(0.02)

Husband gender norms index x FLFP	0.20***	0.25***	0.20***	0.20***
	(0.02)	(0.03)	(0.02)	(0.03)
Age	0.001	0.001	0.002*	0.002*
	(0.001)	(0.001)	(0.001)	(0.001)
Completed Years of Education	0.01***	0.01***	0.01**	0.01***
	(0.003)	(0.004)	(0.005)	(0.006)
Household Head Employment Status	0.007	0.02***	0.001	0.006
	(0.006)	(0.01)	(0.01)	(0.01)
Remittances received	0.02	0.01	0.04	0.03
	(0.03)	(0.03)	(0.03)	(0.03)
Access to Money Index	0.006	0.02**	0.03***	0.03**
	(0.01)	(0.01)	(0.01)	(0.01)
Number of working age people in Household	-0.01**	-0.01***	-0.01**	-0.01**
	(0.004)	(0.005)	(0.006)	(0.006)
Constant	-0.03	0.06	-0.07	-0.02
	(0.04)	(0.05)	(0.06)	(0.10)
Observations	9,508	9,508	5,962	5,962
Number of Groups	5,285	5,285	3,780	3,780
Overall R-sq	0.32	0.26	0.30	0.30

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 30 OLS and IV on MCA Index of Husband-wife gender norms coincidence with Official Treatment Status, Replicated Poverty Score, and their interaction as IVs- Panel Sample 2011-2016

VARIABLES	(1) OLS	(2) REIV	(3) OLS	(4) REIV
FLFP	0.03	0.27	-0.03	-0.23
	(0.03)	(0.41)	(0.02)	(0.35)
Husband gender norms index			0.50***	0.50***
			(0.01)	(0.01)
Husband gender norms index x FLFP			0.18***	0.20***
			(0.02)	(0.03)
Constant	-0.13*	-0.18	-0.07	-0.02
	(0.08)	(0.12)	(0.06)	(0.10)
Observations	5,967	5,967	5,962	5,962
Number of Groups	3,782	3,782	3,780	3,780
Overall R-sq	0.01	0.005	0.30	0.30

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Controls include Age, Completed Years of Education, Household Head Employment Status, Remittances received, Ownership of agriculture, non-agriculture, commercial and residential land, Access to Money Index, Number of working age people in Household.

Table 31 First Stage Results - Panel Sample 2011-2016

	(1)	(2)

VARIABLES	REIV-I FLFP	REIV-I FLFP
BISP Status	0.006	0.006
	(0.01)	(0.01)
Replicated Poverty Score	-0.02***	-0.02***
	(0.01)	(0.01)
BISP Status x Replicated Poverty Score	-0.007	-0.008
	(0.01)	(0.01)
Constant	0.2***	0.2***
	(0.04)	(0.04)
Observations	5,967	5,962
R-squared	0.01	0.02
Kleibergen-Paap rk LM statistic	25.96	24.95
Kleibergen-Paap rk LM P-value	0.00	0.00
First stage F-Statistic	9.12	8.73
Kleibergen-Paap rk Wald F statistic	9.12	8.72
Hansen J Statistic	0.75	3.47
Hansen J P-value	0.68	0.17
Endogeneity Test P-Value	0.58	0.55

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Controls include Age, Completed Years of Education, Household Head Employment Status, Remittances received, Ownership of agri, non-agri, commercial and residential land, Access to Money Index, Number of working age people in Household.

Table 29 Robustness Check with Placebo Treatment Status as IV

	3-Year Panel Sample		4-Year Panel Sample	
VARIABLES	REIV-II FLFP	REIV-II FLFP	REIV-II FLFP	REIV-II FLFP
FLFP	0.40	2.60	14.45	22.51
	(2.1)	(6.54)	(67.22)	(160.0)
Husband gender norms index		0.52***		0.27
		(0.04)		(1.63)
Husband gender norms index x FLFP		0.036		-0.90
		(0.38)		(7.76)
Constant	-0.21	-0.55	-3.21	-4.90
	(0.31)	(1.31)	(14.38)	(34.34)
Observations	9,527	9,508	5,967	5,962
Number of Groups	5,290	5,285	3,782	3,780
Kleibergen-Paap rk LM statistic	1.18	0.27	0.04	0.02
Kleibergen-Paap rk LM P-value	0.27	0.60	0.82	0.88
First stage F-Statistic	1.19	0.27	0.05	0.02
Kleibergen-Paap rk Wald F statistic	1.14	0.27	0.05	0.02

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 33 Hausman Test Gender Norms Index

VARIABLES	(1) FEIV	(2) REIV
FLFP	-8.66	-0.30
	(20.44)	(0.29)
Constant	6.62	2.69***
	(9.92)	(0.12)
Observations	5,836	5,836
Number of Groups	3,735	3,735
Chi Square		2.35
P-value		0.99

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Appendix 2: Regression Discontinuity Design

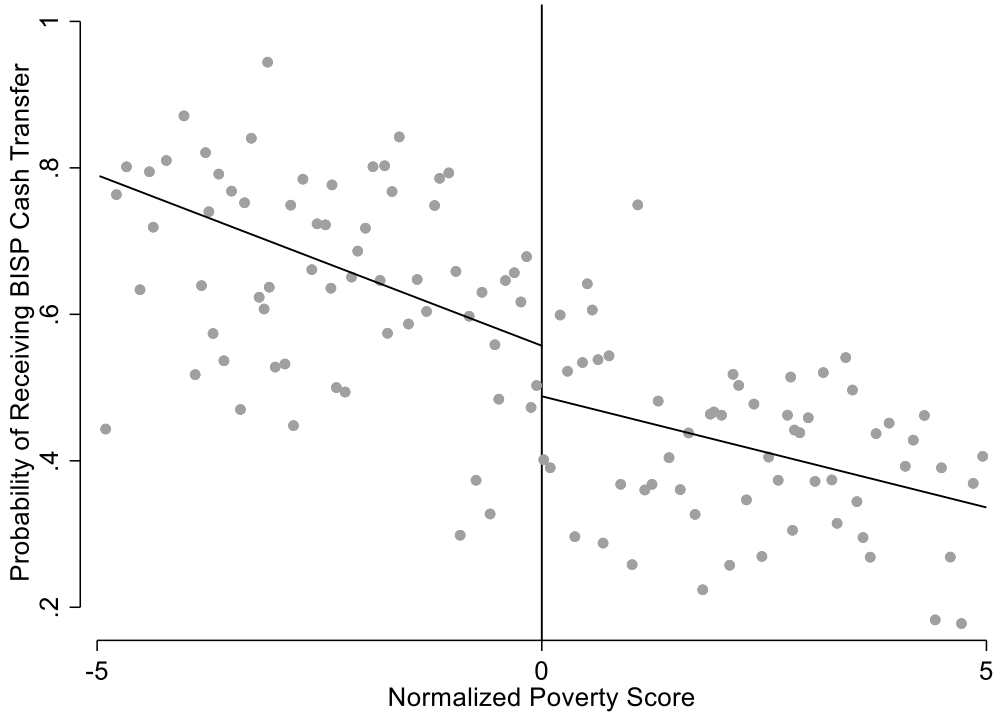
In this research, we also intend to study the impact of BISP cash transfer on female labor force participation. This program is unconditional, and the cash is distributed to the women in a household whose eligibility is determined through a poverty score cutoff. Households below the threshold receive the transfer, making a treatment group; households above the threshold are ineligible to receive the cash, and therefore, they are in the control group. Since there is a discontinuity in the probability of obtaining the cash transfer at the cutoff score, the impact of this program on the outcome variable can be measured through regression discontinuity design. However, in the BISP program, the discontinuity is not sharp, meaning that households above the above threshold may receive the transfer and those below the threshold do not. Therefore, to estimate the treatment effect of BISP cash transfer on FLFP, we use fuzzy RDD that uses eligibility as an instrument for beneficiary status (Ambler and Brauw, 2019).

In RDD analysis, the results are unbiased if many observations are located closer to the cutoff. However, with a small number of observations, the variation in the estimator may be quite high. So, there is a "bias-variance trade-off" (Cattaneo, Idrobo, Titiunik, 2019). Therefore, before doing RDD analysis, there is a need to select the range of values of the running variable, which is the poverty score in our case. This range of values is called bandwidth, and RDD inference is often sensitive to its value. Therefore, the selection of appropriate bandwidth for estimation should be a data-driven and

automatic way. There are many procedures available to determine the optimal bandwidth for estimation of treatment effect; I use the method by Calonico, Cattaneo, and Titunik (2014a, 2014b).

In Figure 2, we show a discontinuity in the probability of receiving the cash transfer around the poverty cutoff score of 16.17. We estimate this result on the panel sample of households from 2011 to 2016. The figure shows that there is a discontinuity in receiving the cash transfer in the panel sample. The graph shows the scatter plot of the probability of receipt by binned poverty score, normalized to have 0 as a cutoff. The linear regression line is estimated on both sides of the cutoff, and the graph is limited to the plus or minus 5 points range of poverty score. At the cutoff, there is a higher probability of receiving the cash transfer for the eligible households on the left of the cutoff. There is a decline in probability on the right of the cutoff.

Figure 2 Effect of poverty score on probability of receiving BISP cash transfer – Panel Sample 2011-2016



Now we present the result of the impact of cash transfer on FLFP. Table 34 shows the results for fixed bandwidth of ±5 and ±3 along with optimal bandwidth. The table presents the conventional, bias-corrected, and robust estimates and standard errors with

control variables included in the estimation. However, in the analysis, we offer statistical inference using the robust p-values. The results show that with a fixed bandwidth of ±5 and optimal bandwidth around the threshold, there is no significant impact of BISP cash transfer on women's labor force participation. The coefficient of the RDD robust estimate appears to be negative. However, with a fixed bandwidth of ±3 around the threshold, this negative effect is statistically significant, leading to failure to reject the null hypothesis of no treatment effect. It suggests that the local average treatment effect of BISP on FLFP is negative, i.e., BISP cash recipient women do not participate in the labor force. The same results are shown through the graphical representation in Figure 3.

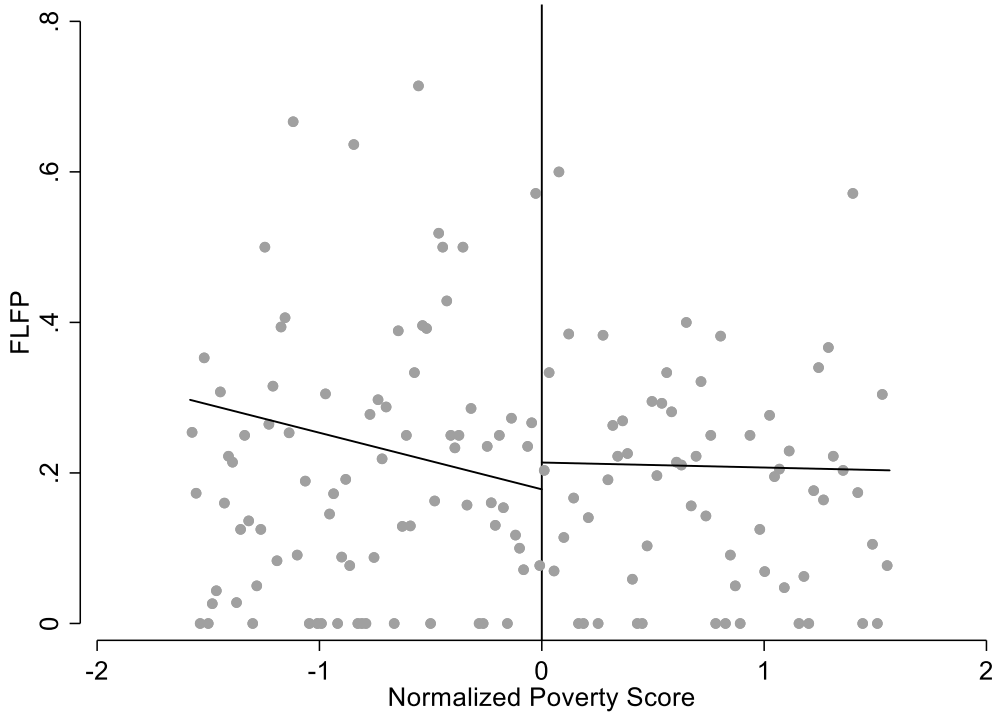
Nevertheless, this result holds for the women in close vicinity of the poverty score threshold of 16.17, i.e., it is the comparison of LFP of women who are just below the cutoff to those who are just above the cutoff. Therefore, the RDD estimation results are consistent in terms of internal validity. However, the same results may not hold when the bandwidth increases and more observations are included in the RDD analysis.

Table 34 RD Estimates of FLFP (without controls) - Panel Sample 2011-2016

VARIABLES	(1) Fixed Bandwidth=±5	(2) Fixed Bandwidth=±3	(3) Optimal Bandwidth
Conventional	0.017	-0.86	-0.60
	(0.28)	(0.98)	(0.58)
Bias-corrected	-0.57**	-2.49**	-0.61
	(0.28)	(0.98)	(0.58)
Robust	-0.57	-2.48*	-0.61
	(0.39)	(1.44)	(0.69)
Conventional p-value	0.95	0.37	0.3
Robust p-value	0.14	0.08	0.37
Bandwidth Left	5	3	4.70
Bandwidth Right	5	3	1.80
Effective Sample Size Left	3850	2485	3747
Effective Sample Size Right	2905	1882	1087

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Figure 3 Effect of BISP cash transfer on FLFP - Panel Sample 2011-2016



Validity of RDD Estimates

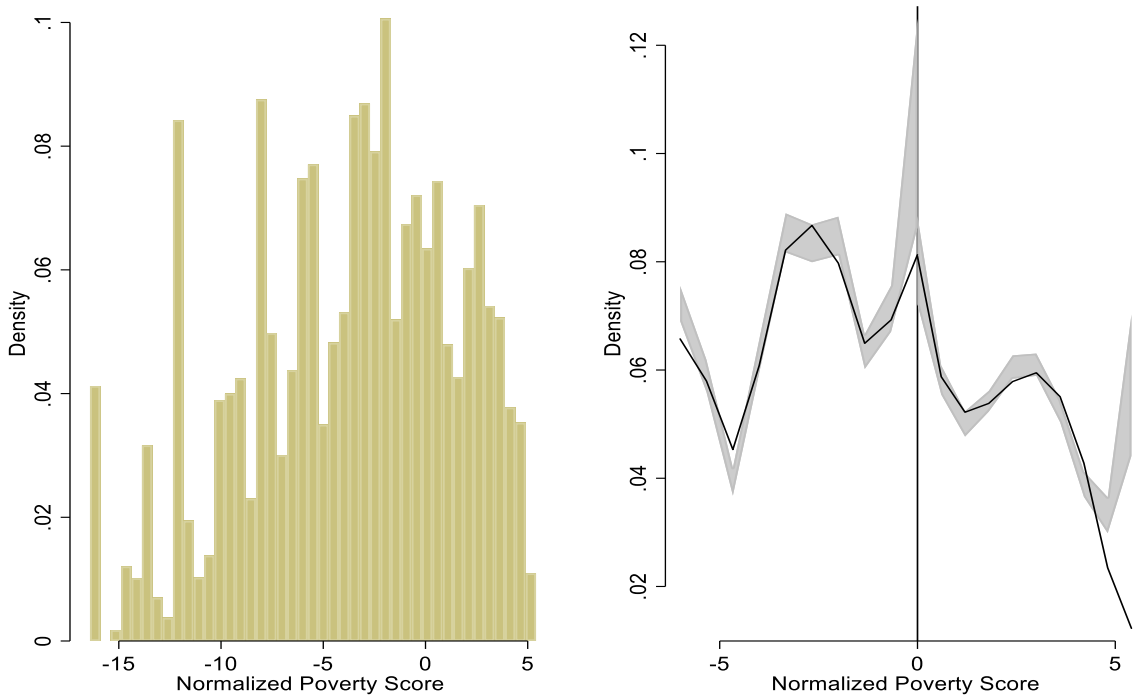
After interpreting the RDD estimation results, it is essential to test the validity of RDD estimates. There are different validations tests offered by RD design that support the plausibility of its assumptions. First, as McCrary (2008) suggested, RD design will be invalid if subjects can manipulate the score or forcing variable. In the case of BISP cash transfer data, poverty score is computed with 23 poverty covariates and households neither know their poverty score nor the cutoff that determines their eligibility. Moreover, the possibility of misreporting the treatment status can be ruled out because households may become ineligible if they are already receiving the treatment and vice versa. However, households may appeal if their poverty score is greater than 16.17 and less than 21.17 and there is a disabled person, 4 or more children, and at least one old-age member in the household. We can formally test for any manipulation by using the method of validation and falsification of RD estimates.

Density of Poverty Score

The first test that we use to check the validity of RD estimates is the density of the running variable, i.e., poverty score. This test examines whether observations near the cutoff are remarkably different on either side. The basic assumption is that if households

cannot manipulate the poverty score, then the number of households above and below the threshold should be similar, giving the continuous probability density function when the score is continuously distributed (Cattaneo, Idrobo, Titiunik, 2019). When the number of observations above and below the cutoff change abruptly, the RD estimates become invalid. This assumption can be verified by visualizing the histogram of the poverty score. However, (Cattaneo, Idrobo, Titiunik, 2019) provide another approach that tests the null hypothesis that the density of running variable around threshold is continuous. This test is implemented by estimating the density of the observations above and below the threshold. Figure 4 provides the histogram and density plot of the poverty score. The histogram shows that there is no abrupt change in the number of observations around the cutoff. The density plot in the right panel of figure 4 also confirms the continuity in the density test approach with a shaded 95% confidence interval. The density plot shows that at the cutoff, the density estimates of the control and treated groups are very close to each other.

Figure 4 Density Plot of Poverty Score



RD Estimates of Predetermined Covariates

An important RD validity test is based on whether treated, and control units are similar in terms of observable characteristics around the threshold. The idea is that if units cannot manipulate their score, they should be identical for all the variables that could

not have been affected by the treatment. These variables are predetermined covariates which are the variables that were determined before the treatment was assigned (Cattaneo, Idrobo, Titiunik, 2019). To validate the RD estimates, we use few baseline characteristics of households and investigate whether there is a continuity in the RD estimates for these predetermined covariates. We test the treatment effect on baseline values of years of education, household head employment status, household head education, whether the household has a female head, working-age people, and household size at baseline. The results are presented in Table 35, and the analysis is conducted similarly to choose the optimal bandwidth. However, the selected bandwidth will differ from what is chosen for the original outcome variable, i.e., FLFP. The results show that we fail to reject the null hypothesis of no treatments effect and thereby infer that the validity of RDD holds.

Table 35 RD Estimates of Predetermined Covariates

VARIABLES	(1) Years of Education	(2) Head Employment Status	(3) Head Education	(4) Female Head	(5) Working-age in HH	(6) Household Size
Conventional	0.69	-19.14	-89.65	0.0323	34.68	108.2
	(2.76)	(31.11)	(205.3)	(0.313)	(152.4)	(532.1)
Bias-corrected	0.715	-24.70	-90.94	-0.0697	-32.00	-11.35
	(2.769)	(31.11)	(205.3)	(0.313)	(152.4)	(532.1)
Robust	0.715	-24.70	-90.94	-0.0697	-32.00	-11.35
	(3.062)	(33.23)	(211.2)	(0.379)	(167.1)	(565.9)
Conventional p-value	0.80	0.53	0.66	0.92	0.82	0.84
Robust p-value	0.81	0.45	0.67	0.85	0.85	0.98
Bandwidth Left	2.57	1.64	1.60	2.01	1.92	1.98
Bandwidth Right	1.22	1.07	0.52	0.80	0.63	0.65
Effective Sample Size Left	3890	2168	2658	3515	3477	3483
Effective Sample Size Right	1487	1291	757	1435	1031	1088

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Placebo Cutoffs

Another test for the validity of RD estimates is to analyze the treatment effect by using placebo cutoffs. The underlying motivation is to check whether the regression functions for the treatment and control group are continuous at a threshold other than the eligibility threshold. If there is evidence of discontinuity on placebo thresholds, then the validity of RD is questionable. This test uses another cutoff value at which there is no change in treatment status and provides inference based on these placebo cutoffs with the expectation that there will be no treatment effect at these placebo cutoffs. Figure 5 shows RD plots of different cutoff values other than 0 in which the true treatment effect occurs. The RD plots are drawn for cutoffs from -3 to +3 except zero. The graphs show that there is no discontinuity at these placebo cutoffs. The results are also confirmed with RD estimates in Table 36.

Figure 5 Placebo Cutoffs on FLFP

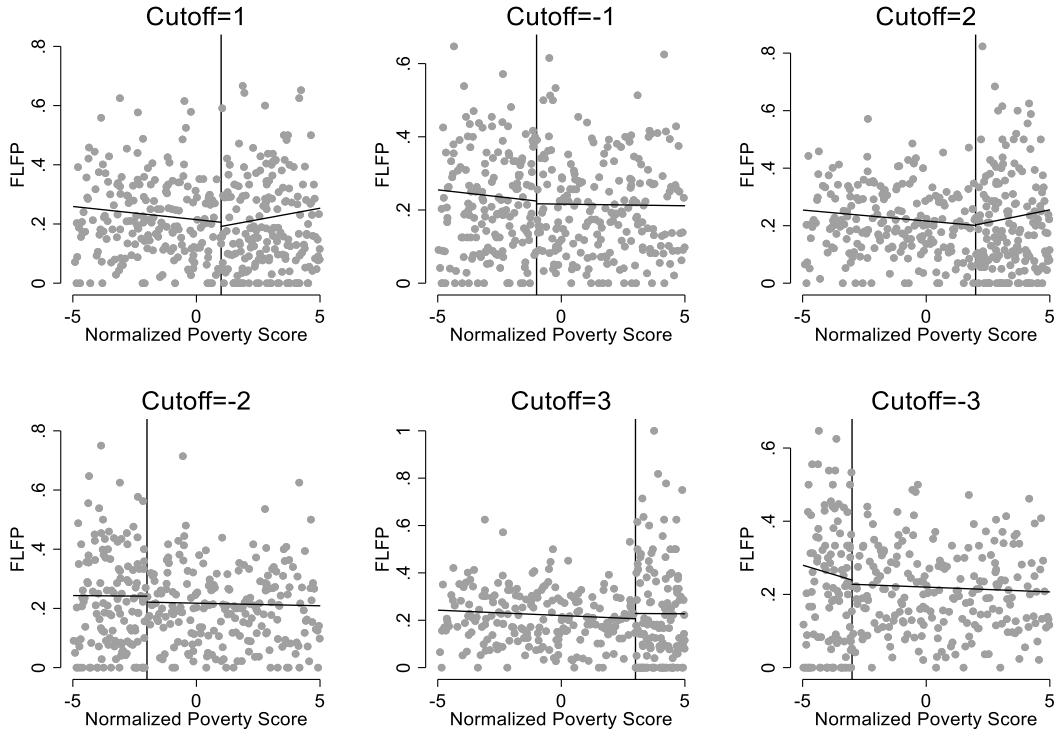


Table 36 Placebo Cutoffs on FLFP

VARIABLES	(1) Cutoff= -3	(2) Cutoff=-2	(3) Cutoff=-1	(4) Cutoff=1	(5) Cutoff=2	(6) Cutoff=3
Conventional	4.67	0.29	0.15	-0.38	12.76	10.28
	(20.30)	(0.340)	(0.149)	(0.518)	(126.5)	(66.08)
Bias-corrected	4.155	0.206	0.134	-0.250	29.75	-23.99
	(20.30)	(0.340)	(0.149)	(0.518)	(126.5)	(66.08)
Robust	4.155	0.206	0.134	-0.250	29.75	-23.99
	(25.40)	(0.396)	(0.173)	(0.594)	(153.7)	(75.02)
Conventional p-value	0.81	0.39	0.29	0.45	0.92	0.87
Robust p-value	0.87	0.60	0.44	0.67	0.84	0.74
Bandwidth Left	0.7	0.98	0.88	1.87	2.16	1.69
Bandwidth Right	1.92	2.41	1.13	1.16	1.02	0.54
Effective Sample Size Left	902	951	616	1326	1375	1050
Effective Sample Size Right	1617	1757	926	727	668	340

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

LIST OF ABBREVIATIONS

ADB	Asian Development Bank
BISP	Benazir Income Support Program
EGRB	Equitable Gender Role Beliefs
EM-DAT	Emergency Events Database
FE	Fixed Effect
FLFP	Female Labor Force Participation
IV	Instrumental Variable
LFP	Labor Force Participation
MCA	Multiple Correspondence Analysis
NADRA	National Database and Registration Authority
NSER	National Socioeconomic Registry
OLS	Ordinary Least Square
OPM	Oxford Policy Management
PKR	Pakistani Rupee
PMT	Proxy Means Test
PSLM	Pakistan Social and Living Standard Measurement Survey
RCT	Randomized Control Trials
RDD	Regression Discontinuity Design
RE	Random Effect
REIV	Random Effect Instrumental Variable
TSLS	Two Stage Least Square
UCT	Unconditional Cash Transfer

LIST OF FIGURES

LIST OF TABLES