Essays on Economics of Gender and Labor Market

Kanika Mahajan

Thesis submitted to the Indian Statistical Institute

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Essays on Economics of Gender and Labor Market

Kanika Mahajan

December 2015

Thesis Supervisor: Prof. Bharat Ramaswami

Thesis submitted to the Indian Statistical Institute

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Acknowledgements

A doctoral thesis is a long journey. It is a path which cannot be traversed without the support of a number of people, whose encouragement plays an important role in reaching the final goalpost. First and foremost, I would like to thank Indian Statistical Institute for providing me an opportunity to conduct my research in an atmosphere which is intellectually stimulating and challenging. The research fellowship from the institute is gratefully acknowledged.

I would like to express heartfelt thanks to my supervisor Prof. Bharat Ramaswami for giving me the freedom to think and explore, and for being there whenever I faltered. This thesis would not have seen the light of day if it had not been for his constructive criticism, constant encouragement and guidance. I have learnt a lot from him, professionally and personally, and cannot thank him enough for his patience and endurance. I owe this research work to his foresight and academic rigor.

I would also like to thank all the faculty members of the Economic and Planning Unit at the Indian Statistical Institute, Delhi. It was an exceptionally enriching and joyful experience to have been taught by them. I would especially like to thank Prof. Abhiroop Mukhopadhyay for always being available to help me think aloud. His fine sense of judgment helped me refine my ideas and gave me confidence to proceed. I am also grateful to Prof. Farzana Afridi for reading through my work and giving me valuable comments and feedback. I would also like to thank Prof. Tridip Ray and Prof. Prabal Roy Chowdhury for their guidance at various stages of my research work.

I have also benefitted from discussion with several scholars. I would like to extend a special thanks to Prof. Indira Rajaraman, Prof. Vegard Iversen, Prof. Kensuke Kubo, Prof. Mukesh Eswaran, Prof. Ashok Kotwal, Prof. Sonia Bhalotra, Prof. Maria S. Floro and Dr. Yumiko Yamamoto for their invaluable feedback. During my graduate days I had the opportunity to attend several seminars and conferences. I would like to thank the participants of all these meetings for giving me feedback which has helped shape my research.

I am grateful to all my seniors and peers, who made this journey beautiful and joyful. This list is not exhaustive but I have to thank Priyanka Grover, Urvashi Jain, Soham Sahoo, Pawan Gopalakrishnan, Shoumitro Chatterjee, Kumarjit, Dyotona Dasgupta, Akansha Batra, Ridhima Gupta, Soumendu Sarkar and Ashokankur Datta, for having lend me an ear whenever I needed one. Sincere thanks to Mr. Rakesh Sinha for his help with data compilation and to Mr. Ramesh Saini for his technical assistance. I would also like to thank Ms. Ashwini Kulkarni and her organization *Pragati Abhiyan* for having hosted me and given me a ground level understanding of the labor markets in rural India.

My deepest gratitude and thanks to my family who have stood by me in this arduous journey. I am thankful to my parents for giving me a life in which I had the luxury to choose what I wanted to do with it. Thank you for encouraging me to tread a path which was unconventional in our family. I cannot thank my husband, Udbhav, enough for all his words of wisdom whenever I felt overwhelmed by the task at hand. His love and faith in me, has encouraged me to carry on and do what I believe in. Thank you for your patience with me in my times of frustration and hearing me out whenever I wanted a vent. Lastly, thank you for our beautiful daughter, Elakshi, whose presence has made me believe in miracles.

To my Parents and Udbhav, I am truly blessed

to have you in my life

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction and Main Results

For the female of the species is more deadly than the male.

Rudyard Kipling

1.1 Background

The differences in labor market outcomes between males and females have been of interest to the economists for at least past half a century. Gender inequality in the labor market manifests itself in the form of wage and employment gaps between males and females. However, little is understood about why these inequalities emerge. There are taste based theories of discrimination, occupational exclusion and theories of statistical discrimination. In this thesis, we study gender disparities in the labor market of rural India. The main objective of this thesis is to further our understanding about the existing wage and employment disparities in the rural labor market of India and whether these disparities can be explained by the economic forces of labor supply and demand. A deeper understanding about functioning of the labor markets can throw light on the processes which result in emergence of these inequalities.

It is well known that achieving gender equality is important for social and economic development of a country (World Bank 2012). Equality in labor market outcomes has been regarded as an important tool for female empowerment. The importance of female participation in the labor force is well-established in the literature. Participation in the labor market provides access to economic opportunities and a greater bargaining power to women in household decision making.

This dissertation consists of three chapters. Each chapter deals with a particular aspect of gender inequality in the labor market. Chapter 2 and chapter 3 look at the differences in daily wage rate received by males and females in the agricultural labor market. Gender wage gaps are known to be persistent and ubiquitous in nature. On an average, the daily wage rate received by a female laborer is 70 percent of daily wage rate of a male laborer, in the rural agriculture labor market of India. A large proportion of this wage gap is unexplained by the standard analysis of differences in human capital characteristics of males and females (Agrawal and Vanneman 2014). The second chapter is concerned with the spatial variation in the gender wage gap in Indian agriculture and aims to understand the significance of other economic forces in shaping gender wage inequality. Chapter 3 looks at over time responsiveness of gender wage gap in Indian agriculture to aggregate shocks in the labor market like rainfall variability.

The fourth chapter of the thesis looks at gender disparity in employment rates in rural India. Gender inequality on this dimension is greater with an overall employment rate of females at 40 percent and that of males at 85 percent in the working age. Moreover, there has been a consistent decline in participation rates by females in the last three decades. This chapter provides a theoretical framework to analyze the effect of female education on female labor supply and empirically examines if increase in female education plays any role in explaining the decline in female labor market participation rates in India.

The sections below contain an overview of each of the three essays in this thesis. I describe the motivation, methodology and the main findings for each essay and discuss them in detail in chapters that follow.

1.2 Caste, Female Labor Supply and the Gender Wage Gap in India: Boserup Revisited

The gender wage gap is notable not just for its persistence but also for its variation across regions and countries. A natural question is how greater work participation by women matters to female wages and the gender wage gap. Within India, a seeming paradox is that gender differentials in agricultural wage are the largest in southern regions of India which have otherwise favorable economic outcomes of women. Boserup (1970) hypothesized that this is due to greater labor force participation by women in these regions. Boserup's hypothesis is based on raw correlations drawn from wage data across Indian villages in the 1950s. However, this is not obvious as greater female labor supply could depress male wage as well. Differences in agro-climatic conditions across regions could also lead to differences in relative female efficiency in agricultural tasks due to variation in agricultural technology. Other factors also need to be accounted for especially since women have fewer opportunities for non-farm employment. It is well known that the labor flow from agriculture to other sectors has been much more marked for males than for females (Eswaran et al. 2009). So if men have greater access to non-farm work opportunities, do women working as agricultural labor gain from growth in the non-farm sector? This chapter undertakes a formal test of the Boserup proposition incorporating comprehensive controls for other explanations.

The chapter presents a theoretical framework for the analysis and then undertakes an econometric estimation to test the hypothesis. Econometrically, we need to identify exogenous variation in female and male labor supply to agriculture. The effect of female labor supply on wages is identified by variation in cultural and societal norms that regulate female labor supply (Nayyar 1987; Chen 1995; Bardhan, K 1984; Das 2006; Eswaran, Ramaswami and Wadhwa 2013). The proportion of men employed in large-sized non-farm enterprises is used to identify the effect of male labor supply on daily agricultural wages. As variation in low-caste population and presence of large scale industry might be correlated with variables that directly affect the demand for agricultural labor, we include comprehensive controls to identify the causal impact. These controls include agro-climatic endowment, cropping patterns and infrastructure.

District level wage and employment data from National Sample Survey 2004-05, is used to estimate the aggregate total demand for labor equations for male and female labor in Indian agriculture. The results show that an increase in female labor supply by 10% decreases female wage by 5% (significant) and male wage by 1% (insignificant). This shows that an increase in female labor supply to agriculture by 10% decreases relative female wage (female to male wage ratio) in agriculture by 4%. On the other hand increase in male labor supply decreases both male and female wage equally by approximately 4% and hence as no effect on the gender wage gap.

Decomposing the difference in wage gap between northern and southern states of India, we find that differences in female labor supply are able to explain 55 percent of the difference in gender wage gap between northern and southern regions. The chapter also finds that women gain from greater non-farm employment of men, even if their direct participation in such activity is limited. This happens because of higher wages.

1.3 Rainfall Shocks and the Gender Wage Gap: Evidence from Indian Agriculture

Previous studies have shown that productivity shocks in agriculture like rainfall variability affect wages adversely in developing countries (Jayachandran 2006; Mueller and Quisumbing 2011; Mueller and Osgood 2009; Adhvaryu, Char and Sharma 2013). None of the studies however consider heterogeneity in the impact of these shocks on agricultural wages by gender, a feature which has been studied for demand shocks in urban labor markets for developed countries (O'Neill 1985; Biddle and Hamermesh 2011; Solon, Barsky and Parker 1994; Park and Shin 2005; Marchand and Olfert 2013; Aller and Arce 2001).

Since Indian agriculture is highly dependent on rainfall, the shock to the agricultural labor market which this chapter considers is rainfall variability over the years. It estimates the impact of rainfall shocks on relative female wage. It also examines the heterogeneity of this impact in the rainfed rice growing areas of India. The literature on gender roles in cultivation of rice suggests that there exists a greater demand for female agricultural labor in areas where transplanted rice is grown (Bardhan 1974; Sen 1985). Also, rice is a water-intensive crop.

A theoretical model of wage and employment determination in agriculture shows that aggregate productivity shocks like rainfall variability can have an impact on relative female wage through their asymmetric effects on demand for male and female labor and asymmetric supply responses of male and female labor to these shocks. Using National Sample Survey data for India from 1993 to 2007, I create a district level panel dataset to examine how rainfall shocks affect gender wage gap in Indian agriculture. I find that rainfall shocks affect relative female wage significantly in rainfed rice growing regions of India. As compared to other areas, relative female wage falls (increases) in low (high) rainfall years in the rice growing regions by 5%. This finding is consistent with the documented evidence in the literature that marginal value of female labor is greater in rice cultivation which is also a crop highly sensitive to rainfall variability under rainfed conditions.

I further look into the mechanisms which lead to this result. Either differential demand effects on female labor or differential supply responses of female labor, in comparison to males, in these rainfed rice growing areas, could lead to the observed results. The tests, derived from the existing literature and the theoretical setup, suggest that differential demand side effects on female labor are behind the result that relative female wage responds positively to rainfall shocks in these areas. The chapter concludes that the effect of rainfall shocks on the gender wage gap in agriculture depends upon the gender roles underlying the technology of production which vary across cropping systems.

1.4 Female Labor Supply, Education and Home Production

Indian labor market is characterized by low female labor market participation rates. And it has witnessed a further decline in employment rates of rural women over the past few decades. Female labor force participation rate (LFPR) has fallen from 51% in 1987 to 39% in 2009 in rural areas. This fall has occurred amongst the currently married females and has been accompanied by an increase in domestic work by them. Over the same time period, the country has witnessed a rapid economic growth as well as an increase in education attainment by males and females.

Existing studies, using descriptive statistics, attribute the decline to increase in enrollment of females in higher education resulting in their unavailability for labor market, increase in real incomes of households resulting in a lesser need for women to engage in economic work, and limited growth in employment opportunities for females. We show in the chapter that these reasons cannot fully explain the decline in female LFPR observed over the last three decades in India. The U-shaped behavior of female LFPR with female education, combined with a large increase in primary to secondary level education among females in rural India could be a contributing factor behind the decline in female LFPR.

If so, standard models of female labor supply warrant a revision. In the standard models of female labor supply, female education increases her market wage, and this always results in an increase in female labor supply with her education, at the extensive margin. However, primary education also gives basic skills to mothers, which can improve their productivity in home tasks like raising children (Lam and Duryea 1999). In this chapter, we present a theoretical model of household decision making which shows that in the presence of home production, female labor supply can decline with rising female education levels. This model shows that under certain parametric restrictions one could obtain a fall in female labor supply and an increase in female time spent in home production with increase in female schooling.

We estimate the determinants of female labor market participation and domestic work in India using the National Sample Survey data on employment and unemployment for 1987, 1999 and 2009. Empirical analysis shows that female labor force participation in India decreases and female domestic work increases with female schooling up till secondary education. These results are robust to inclusion of controls for household income, male education and local labor market conditions. Further, parametric and semi-parametric decomposition techniques are used to estimate the proportion of fall in female participation rates which can be explained by the changing socio-economic characteristics of female working age population. The decomposition exercise shows that rising education levels amongst rural married females and males are the most important factors in explaining the decline. The empirical findings are in line with the prediction of the theoretical model that female labor supply falls and her time spent in home production increases with increase in her education, if the marginal return from spending time in the labor market is lower than the marginal return from spending time in home production as female education increases.

Chapter 2

Caste, Female Labor Supply and the Gender Wage Gap in India: Boserup Revisited[1]

2.1 Introduction

The gender gap in wages is a persistent feature of labor markets despite laws mandating equal treatment of women at workplace. What is just as notable is the variation in the gender wage gap across regions and countries, and in some cases, over time as well. In a cross-country context, observable differences in characteristics and endowments, explain only a small portion of the wage gap (Hertz et al. 2009). Since the unexplained component is the dominant one, the geographical variation in the wage gap is commonly attributed to discrimination.

However, discrimination may not be the only reason. If female and male labor are imperfect substitutes, then the wage gap would vary with male and female labor supply. In many regions of the United States, female wages fell relative to male wages during the Second World War (Aldrich 1989; Acemoglu, Autor and Lyle 2004). By exploiting cross-sectional variation in change in female work participation rates that occurred during the World War II, Acemoglu et al. (2004) showed that higher female labor supply increased the gender gap in wages in the United States. In a sample of 22 countries drawn mostly from the

[1] This chapter is based on a paper co-authored with Prof. Bharat Ramaswami, Indian Statistical Institute, Delhi. I would like to thank Richard Palmer-Jones and Kunal Sen for providing district level data on agro-ecological zone composition.

OECD, Blau and Kahn (2003) also explored the idea that higher female labor supply can exacerbate the gender wage gap.

In a developing country context, the role of female labor supply in influencing the gender gap in wages was highlighted by Boserup (1970) in her influential book, *Women's Role in Economic Development*. She pointed to the geographical variation in the female to male agricultural wage ratio that existed in rural India during the 1950s. The gender wage gap in agriculture was greater in southern states of India relative to the states in north India and Boserup ascribed this to the much higher female participation rates in farming in South India. Figure 2.1 maps the ratio of female to male agricultural wage in rural India across the Indian states in year 2004. It is easy to observe a systematic regional pattern – of the same kind as Boserup described 50 years ago.

Boserup's hypothesis is based on raw correlations drawn from wage data across Indian villages in the 1950s. However, the hypothesis is not immediately obvious because variation in female labor supply could affect male wage as well. The extent to which the female and the male labor are substitutes matters. In addition, there are competing explanations. For instance, there could be gender segregation by task where `female' tasks are possibly paid less than supposedly `male' tasks. Second, the relative efficiency of female to male labor in agriculture could vary across regions due to differences in agricultural technology, variation in cropping patterns and agro-climatic conditions. Third, factors that affect the supply of male labor to agriculture, such as non-farm employment, could also matter to the wage gap. The impacts of all these factors must be considered in the analysis. This is what is done in this chapter.

The goal of this chapter is to explain the spatial variation in the gender gap in agricultural wage in rural India. In particular, the chapter asks whether exogenous variations in female as well as male labor supply to agriculture play any part in explaining the gender wage gap.

The effect of male labor supply on the gender wage gap is of independent interest as well. It is well known that labor flow from agriculture to other sectors has been much more marked for males than for females (Eswaran et al. 2009). So if men have greater access to non-farm work opportunities, do women working as agricultural labor gain from growth in the non-farm sector? In trying to understand the impact of economic growth on the economic well being of women, the effect of non-farm employment on the gender wage gap is of immense importance.

Econometrically, we estimate district level inverse demand functions that relate female and male agricultural wages to exogenous variation in female and male labor supply to agriculture. The conceptual challenge is to identify exogenous variation in female and male labor supply to agriculture. The effect of female labor supply on daily wages is identified by variation in cultural and societal norms that regulate female labor supply. In India, the pattern of high female work participation rates in south India relative to north India has persisted over many decades (Nayyar 1987; Chen 1995; Bardhan, K 1984; Das 2006) suggesting the salience of cultural norms. Boserup observed that typically, higher caste Hindu women take no part in cultivation activities while tribal and low caste women have traditions of female farming either on their own land or as a wage laborer.[2] She also points out that tribal and low caste populations are lower in north India relative to other parts of the country. Boserup follows up these observations with its consequences. In her words,

[2] Cultivation and agriculture are used interchangeably in the thesis. They are defined to include cultivation activities of ploughing, sowing, weeding, transplanting, harvesting and other agricultural activities.

> "The difference between the wages paid to women and to men for the same agricultural tasks is less in many parts of Northern India than is usual in Southern India and it seems reasonable to explain this as a result of the disinclination of North Indian women to leave the domestic sphere and temporarily accept the low status of an agricultural wage laborer." (Boserup 1970, 61)

The plausibility of social norms driving the north-south divide in female work participation is consistent with the well-known finding that women have greater autonomy in the southern states of India (Dyson and Moore 1983). Basu (1992) and Jejeebhoy (2001) also find similar patterns in woman's status indicators across India's north and south.[3] Boserup's association of social group membership with female work participation has been confirmed in later work as well (Chen 1995; Das 2006; Eswaran, Ramaswami and Wadhwa 2013). Taking a cue from these studies, we take the proportion of households that are low-caste as an instrument for female labor supply. The idea that social norms determine women's labor supply decisions is not unique to India (Boserup 1970; Goldin 1995; Mammen and Paxson 2000). What is characteristic of India is the variation of these norms along identifiable social groups.[4] As variation in low-caste population might be correlated with variables that directly affect the demand for agricultural labor, we include a comprehensive set of controls to identify the causal impact. These control variables include agro-climatic endowment, cropping pattern and infrastructure.

[3] However, Rahman and Rao (2004) do not find such a distinct differentiation across all indicators of woman's status.

[4] Cross-country variation in women's participation can also be related to cross-country variation in social norms (Cameron, Dowling and Worswick 2001).

The proportion of men employed in large-sized non-farm enterprises instruments male labor supply to agriculture. Large enterprises reflect external demand and are therefore a source of exogenous variation in agricultural labor supply. As we argue later, the possible pitfalls in the use of this variable as an instrument are addressed by inclusion of appropriate controls in the estimating equation.

In the next section we relate this chapter to the relevant literature. In section 2.3, we provide suggestive evidence in support of Boserup hypothesis. Section 2.4 outlines a theoretical framework, which is followed in section 2.5 by a discussion of the empirical strategy. The data is described in section 2.6 and section 2.7 contains the estimation results. To check for robustness, section 2.8 considers alternative specifications. The estimation results are used in section 2.9 to quantitatively decompose the difference in wage gap across the northern and the southern states of India into the proportion explained by each explanatory variable. Concluding remarks are gathered in section 2.10.

2.2 Relation to Literature

Blau and Kahn (2003) analyze the gender wage gap across 22 countries and find evidence that the gender gap in wages is lower when women are in shorter supply relative to their demand. They construct a direct measure of female net supply using data across all occupations and recognize that their estimates might be biased due to reverse causality. Acemoglu et al. (2004) correct for the endogeneity of female labor supply using male mobilization rates during World War II as an instrument for labor supply of females to the non-farm sector in the United States. They find that an increase in female labor supply lowers female wage relative to male wage. In some specifications, the endogenous variable that is instrumented is the female to male labor supply ratio. In other specifications, the female and the male labor supply enter as separate explanatory variables but only the female labor supply is instrumented.

In the Indian context, Rosenzweig (1978) was the first paper to estimate labor demand functions for agricultural labor in India to estimate the impact of land reforms on male and female wage rates. This exercise is embedded within a general equilibrium market clearing model of wage determination. In the empirical exercise, Rosenzweig estimates inverse demand and supply equations for hired labor of males, females and children in agriculture using wage data on 159 districts in India for the year 1960-61. His results show that an increase in female labor supply has a negative effect on both male and female wage rates. Further, the paper is unable to reject the null hypothesis that both effects are of equal magnitude. Thus, the Boserup hypothesis is not supported.

There are several reasons to revisit this analysis. First, the wage data used by Rosenzweig, is not well-suited for capturing cross-sectional variation.[5] The better data set for this purpose (and which is used in this chapter) is the unit level data from the Employment and Unemployment schedule of the National Sample Survey (NSS) which was unavailable to researchers at the time Rosenzweig did his study. [6] Second, as a measure of agricultural labor supply, Rosenzweig uses the percentage of male (or female) agricultural labor force to the total labor force. However, after controlling for agricultural labor supply, changes in total labor supply should not matter to wages. Our specification for the labor demand function

[5] Rosenzweig (1978) uses the wage data reported in Agricultural Wages in India (AWI). The problem with AWI is that no standard procedure is followed by states as the definition of ‘wage’ is ambiguous. Only one village is required to be selected in a district for the purpose of reporting wage data and the prevailing wage is reported by a village official on the basis of knowledge gathered.

[6] See Rao (1972) and Himanshu (2005) for a discussion about the merits of different sources of data. The consensus is that although the AWI data may work well for long-term trend analysis but it is not suitable for a cross sectional analysis since measurement bias may differ across states.

derives from a production function that has land and labor as inputs, and exhibits constant returns to scale. As a result, the relevant labor supply variable is the agricultural employment (male or female) per unit of cultivated land.

Third, Rosenzweig limits the definition of agricultural labor to hired labor alone. This chapter, on the other hand, estimates the demand for total labor and not for hired agricultural labor because it is harder to find instruments that are valid for hired labor demand. Suppose $L^{f,S}$ and $L^{o,S}$ are the aggregate labor supply to the home farm and to the outside farms respectively. Similarly, let $L^{f,D}$ and $L^{h,D}$ be the aggregate demand for family and hired labor respectively. Then equilibrium in the labor market can either be written as $L^{f,S} + L^{o,S} = L^{f,D} + L^{h,D}$ or as $L^{o,S} = L^{h,D}$. However, for econometric estimation, it is preferable to estimate the inverse demand for total agricultural labor than for hired labor alone. This is because the instruments that affect labor supply to outside farms would also affect labor supply to own farm and hence potentially affect the demand for hired labor. For instance, higher caste women may refrain from work outside the home and also limit their work on own farms. Similarly, availability of non-farm work opportunities may reduce the family labor supply of landed households to own farms and increase the demand for hired labor. A simple sum of hired and family labor would, however, contradict the accepted notion that family labor is more efficient than hired labor. Moreover, as we shall see later, the implication of using an un-weighted aggregate is that there might be an omitted variable correlated with the aggregate labor supply. However, we demonstrate that our findings are robust to whichever way the family and the hired labor are weighted to form aggregate labor supply.

Finally, current data allows for more comprehensive controls and better identification strategies than available to Rosenzweig. We are able to employ controls for crop composition, agro-ecological endowments and district infrastructure. For identification, Rosenzweig assumes that the demand for hired labor (whether male, female or child labor) is not affected by proportion of population living in urban areas in the district, indicators of the non-farm economy (factories and workshops per household, percentage of factories and workshops employing five or more workers, percentage of factories and workshops using electricity) and the percentage of population that is Muslim. We do not use urbanization as an instrument because that could be directly correlated with agricultural productivity by determining the access to technology and inputs. We therefore employ urbanization as a control variable in some of our specifications. We improve on the non-farm economy instrument by confining it to the large enterprises in the traded sectors. Section 2.5 argues why such an instrument is plausibly exogenous. We replace the percentage Muslim population variable by the proportion of population that is of low-caste. As we argue in section 2.5, there is a large literature that highlights caste-specific norms of female labor supply in India.

Other studies that estimate structural demand and supply equations for hired agricultural labor in India are Bardhan (1984) and Kanwar (2004). Bardhan (1984) estimates simultaneous demand and supply equations for hired male laborers at village level in West Bengal. He instruments the village wage rate by village developmental indicators, unemployment rate and seasonal dummies. Kanwar (2004) estimates village level seasonal labor demand and supply equations for hired agricultural labor simultaneously accounting for non-clearing of the labor market using ICRISAT data. These studies do not analyze male and female laborers separately and they cover only a few villages in a state. Singh (1996) estimates an inverse demand function for both males and females in agriculture, using state level pooled time series data for 1970 to 1989; however ordinary least squares methods are used and the endogeneity of labor supply is not corrected. Datt (1996) develops an alternative bargaining model of wage and employment determination in rural India. In this model, the gender wage gap is determined by the relative bargaining power of females. The lower wage of female agricultural laborer relative to male laborer in rural India is thus attributed to their lower bargaining power in comparison to males, during the bargaining process with the employers.

2.3 The Gender Gap in Wages and Female Labor Supply: Correlations

Figure 2.2 cross-plots the state-level average of female to male agricultural wage ratio against female labor time in agriculture per unit of cultivable land. This figure is based on data from a national survey in 2004 and is consistent with Boserup's hypothesis that the two variables are inversely related.[7]

If the female and the male labor are perfect substitutes in agricultural production, then a change in female labor supply, say a decline, would raise both female and male wages proportionately and not affect the gender wage gap (which in a world without discrimination would be solely due to gender differences in marginal product). For Boserup hypothesis to hold, female and male labor must not be perfect substitutes so that changes in female labor supply affects female wage more than male wage. The lack of perfect substitutability is closely related to the gender division of labor within agriculture that is often found in many

[7] Kerala, the state with the best human development indicators, is an outlier to the Boserup relation. Like other southern states, its female to male wage ratio is low. Unlike other southern states, however, the agricultural female employment (per unit of land) is also low. This is partly because Kerala uses less labor (female or male) per unit of land than other southern states. So if the female labor supply was measured as a proportion of male labor supply, Kerala is substantially closer to the Boserup line although it remains an outlier.

countries (Burton and White 1984; Doss 1999). For instance, in many societies, weeding is usually seen as a task mostly performed by females while ploughing is a task done mostly by males. Direct evidence on limited substitutability of female and male labor in agriculture has been found in a number of studies in India and other countries (Jacoby 1992; Laufer 1985; Skoufias 1993; Quisumbing 1996).

If some tasks are better paid than others and if males mostly do the better paid tasks and females do the less paying tasks, then that could result in a gender wage gap. In this case, the geographical variation in the gender wage gap could simply be because of variation in the gender division of labor. It is, in fact, true that the gender division of labor is more pronounced in southern states of India.[8] However, this is not the primary reason for either the gender wage gap or its variation.

In table 2.1, individual wage rates are regressed on gender, age, age square, education and marital status.[9] With these control variables, column (1) shows that females get a 35 percent lower daily wage than males in agriculture. In column (2) we add the controls for agricultural task for which the daily wage was recorded. The gender wage gap narrows slightly to 33 percent. Thus, the gender wage gap in Indian agriculture is mostly within tasks.

A direct way of accounting for variation across states in the gender division of labor is to hold it constant and to re-do the plot in figure 2.2. The female to male wage ratio for state ‘s’ is the weighted mean across tasks given by

[8] This was found by computing, for each state, the proportion of agricultural labor days of males and females spent in each task. An index of gender division of labor (in agricultural tasks) for each state was constructed by considering the Euclidean distance measure between female and male labor proportions.

[9] NSS records wage data for each agricultural task done by a person in the preceding week. It is possible that a person is involved in more than one task. The unit of observation hence is individual-task wage. However, the multiplicity of tasks by the same individual is not a grave concern since only 2.5 percent of individuals report undertaking more than one agricultural task in the preceding week.

$$\frac{w_{fs}}{w_{ms}} = \frac{\sum P_{fjs} w_{fjs}}{\sum P_{mjs} w_{mjs}}$$

where $w_{fs}(w_{ms})$ is the female (male) wage in state 's', P_{fjs} (P_{mjs}) is the proportion of females (males) working in task 'j' in state 's' and $w_{fjs}(w_{mjs})$ is the female (male) wage in task 'j' in state 's'. Suppose we replace the state proportions in tasks by females and males by the proportions observed for the southern state of Tamil Nadu (arbitrarily chosen), then the wage ratio in state 's' becomes

$$\frac{w'_{fs}}{w'_{ms}} = \frac{\sum P_{fj,TN} w_{fjs}}{\sum P_{mj,TN} w_{mjs}}$$

Figure 2.3 plots this measure of agricultural wage ratio, which is devoid of variation in gender division of labor across states, against the female employment in agriculture. The negative relationship between female to male wage ratio and female employment still persists, even when we account for differential participation in tasks by males and females across states in India. As shown earlier, this is because the wage difference between males and females in Indian agriculture is mostly within the same task.

2.4 Theoretical Framework

Before proceeding with the empirical strategy it is useful to discuss the theoretical implications of exogenous changes in male and female labor supply on male and female wages. When male and female labor supply changes are exogenous, the resulting impact on wages can be determined by reading off the labor demand curve. Identification of such exogenous changes and estimation of the demand curve is the subject of later sections.

Assume a homogenous, continuous and differentiable agricultural production function with three factors of production – land (A), male labor (L_M) and female labor (L_F).

Returns to each factor are diminishing and land is fixed in the short run. Let w_M and w_F be the male and the female wage rate respectively. The profit function is given by

$$\pi = F(A, L_M, L_F) - w_M L_M - w_F L_F$$

Let F_{L_M} and F_{L_F} denote the marginal product of male and female labor respectively. For given wages, the first order conditions for labor demand satisfy

$$\ln(w_M) = \ln(F_{L_M}) \tag{2.1}$$

$$\ln(w_F) = \ln(F_{L_F}) \tag{2.2}$$

If labor supply were to, say, increase for a reason exogenous to wages, then wages must adjust to increase demand. We derive the own and the cross-price elasticity of male labor demand as

$$\frac{\partial \ln(w_M)}{\partial \ln(L_M)} = \frac{\partial \ln(F_{L_M})}{\partial L_M} \frac{\partial L_M}{\partial \ln(L_M)} = \frac{F_{L_M L_M}}{F_{L_M}} L_M \tag{2.3}$$

$$\frac{\partial \ln(w_M)}{\partial \ln(L_F)} = \frac{\partial \ln(F_{L_M})}{\partial L_F} \frac{\partial L_F}{\partial \ln(L_F)} = \frac{F_{L_M L_F}}{F_{L_M}} L_F \tag{2.4}$$

Similarly, expressions for the own and the cross-price elasticity of female labor demand are given by

$$\frac{\partial \ln(w_F)}{\partial \ln(L_F)} = \frac{F_{L_F L_F}}{F_{L_F}} L_F \tag{2.5}$$

$$\frac{\partial \ln(w_F)}{\partial \ln(L_M)} = \frac{F_{L_F L_M}}{F_{L_F}} L_M \tag{2.6}$$

The diminishing return to factor inputs implies that own-price elasticities, (2.3) and (2.5) are negative. To sign the cross-price elasticity we need to know whether male and female labor are substitutes or complements in the production process.[10] If they are substitutes then (2.4) and (2.6) will also be negative since the marginal product of male labor will decline if female labor increases and vice versa. If they are complements then (2.4) and (2.6) will be positive.

The effect of female employment on relative female wage is given by $\frac{\partial \ln(w_F/w_M)}{\partial \ln(L_F)} = \frac{\partial \ln(w_F)}{\partial \ln(L_F)} - \frac{\partial \ln(w_M)}{\partial \ln(L_F)}$. If the male and the female labor are imperfect substitutes, this expression cannot be signed without further restrictions. If the two kinds of labor are complements, then increase in female labor employment will decrease the female to male wage ratio (or increase the gender wage gap). Similarly, the effect of male labor employment on relative female wage is given by $\frac{\partial \ln(w_F/w_M)}{\partial \ln(L_M)}$. Again, this expression cannot be signed when the male and the female labor are imperfect substitutes. If they are complements, then an increase in the male labor employment will increase the female to male wage ratio (or reduce the gender wage gap). Note that the relative magnitude of the cross-price elasticities can be obtained from (2.4) and (2.6). This is given by

$$\frac{\partial \ln(w_F)}{\partial \ln(L_M)}\frac{\partial \ln(L_F)}{\partial \ln(w_M)} = \frac{L_M}{F_{L_F}}\frac{F_{L_M}}{L_F} = \frac{L_M}{w_F}\frac{w_M}{L_F} = \frac{L_M}{L_F}\frac{w_M}{w_F} \tag{2.7}$$

The relative magnitude of cross-price elasticities can, thus, be expressed as a product of male to female labor employment and male to female wage ratio. In the Indian agricultural labor market, it is seen that labor employment of males is greater than that of females and male

[10] We use the q-definition of complementarity and substitutability which measures the effect on marginal product of one input when the quantity of the other input changes. Two factors inputs are called q-substitutes (complements) when an increase in one factor decreases (increases) the marginal product of the other factor.

wage is also greater than female wage. Therefore, the above expression will be greater than unity which implies that the effect of male labor employment on female wage will be greater than the effect of female labor employment on male wage. Later, in the chapter we see if the estimate of the relative cross-price elasticities, implied by the above theoretical model, holds ground empirically.

2.5 Empirical strategy

For observed levels of female and male employment in agriculture, the inverse demand functions can be written as

$$W_{M,d} = \alpha_0 L_{F,d} + \beta_0 L_{M,d} + \gamma_0 X_d + \varepsilon_{M,d} \qquad (2E.1)$$

$$W_{F,d} = \alpha_1 L_{F,d} + \beta_1 L_{M,d} + \gamma_1 X_d + \varepsilon_{F,d} \qquad (2E.2)$$

The first stage regressions for male and female employment are

$$L_{M,d} = \delta_0 X_d + \pi_0 Z_d + \epsilon_{M,d} \qquad (2E.3)$$

$$L_{F,d} = \delta_1 X_d + \pi_1 Z_d + \epsilon_{F,d} \qquad (2E.4)$$

where 'd' indexes district and subscripts 'M' and 'F' denote male and female respectively, W_d is log of real daily agricultural wage in district 'd', L_d is log of labor employed in agriculture in district 'd', X_d are other district level control variables and Z_d are the instrumental variables used to identify the impact of male and female labor supply on wages. The inverse demand functions are estimated at the level of a district.[11] This requires Indian districts to approximate separate agricultural labor markets. This has also been assumed in

[11] India is a federal republic having a parliamentary system of governance. It is divided into various states having their own elected government. Each state is further divided into different administrative districts.

previous studies on Indian rural labor markets (Jayachandran 2006; Rosenzweig 1978) and is supported by the conventional wisdom that inter-district permanent migration rates are low in India (Mitra and Murayama 2008; Munshi and Rosenzweig 2009; Parida and Madheswaran 2010). While some recent work has questioned this, the evidence here points to rural-urban and out-country migration rather than rural-rural migration (Tumbe 2014). If rural-rural labor mobility across districts is large in India, then, the district level effect of labor supply changes on agricultural wages will be insignificant.

From $(2E.1)$ and $(2E.2)$ it can be seen that the effect of female labor supply on female to male wage ratio is given by $(\alpha_1 - \alpha_0)$. As α_1 is expected to be negative, an increase in female labor supply leads to a greater gender gap in agricultural wages (i.e., the Boserup hypothesis) if $(\alpha_1 - \alpha_0) < 0$. Similarly, the effect of male labor supply on the relative female wage is $(\beta_1 - \beta_0)$. A decline in male labor supply to agriculture due to greater non-farm employment opportunities would increase the gender gap in agricultural wages if $(\beta_1 - \beta_0) > 0$. Identification requires that we relate wages to exogenous variation in female and male labor supply to agriculture.

2.5.1 Identification of the Impact of Female Labor Supply

For female labor supply, this chapter uses the proportion of district population that is low caste as an instrument.[12] The relation between district level female employment in agriculture and the instrument is plotted in figure 2.4. The positive association between the two is

[12] The definition of `low caste' is the following. In the employment survey (which is our data source), households are coded as 'scheduled tribes', 'scheduled castes', 'other backward classes' and 'others'. Scheduled tribes (ST) and scheduled castes (SC) are those social groups, in India, that have been so historically disadvantaged that they are constitutionally guaranteed affirmative action policies especially in terms of representation in Parliament, public sector jobs, and education. Other backward class (OBC) is also a constitutionally recognized category of castes and communities that are deemed to be in need of affirmative action (but not at the cost of the representation of ST and SC groups). 'Others' are social groups that are not targets of affirmative action. We define a household to be low caste if it is ST, SC or OBC.

consistent with earlier work that has established the effect of caste on female labor supply. These studies observe that high caste women refrain from work participation because of 'status' considerations (Aggarwal 1994; Bagchi and Raju 1993; Beteille 1969; Boserup 1970; Chen 1995). Correlations from village level and local studies have been confirmed by statistical analysis of large data sets. Using nationally representative employment data, Das (2006) shows that castes ranking higher in the traditional caste hierarchy have consistently lower participation rates for women. The 'high' castes also have higher wealth, income and greater levels of education. So could the observed effect be due to only the income effect? In an empirical model of household labor supply, Eswaran et al. (2013) show that 'higher' caste households have lower female labor supply even when there are controls for male labor supply, female and male education, family wealth, family composition and village level fixed effects that control for local labor market conditions (male and female wages) as well as local infrastructure.

The exclusion restriction for identification of the impact of female labor supply on wage rates is that caste composition affects wages only through its affect on labor supply of women to agriculture. Could the caste composition of a district directly affect the demand for agricultural labor? Das and Dutta (2008) find no evidence of wage discrimination against low castes in the casual rural labor market of India. An earlier village level study by Rajaraman (1986) also did not find any effect of caste on offered wage in Indian agriculture.

However, the disinclination of higher-caste women to work suggests that their reservation wage ought to be higher. Table 2.2 shows the results for the regression of individual female wages on a dummy for low caste and other controls. The low caste dummy is insignificant controlling for age, education, marital status, type of agricultural operation and district fixed effects. If the district fixed effects are dropped, then the low caste dummy is negative and significant even with other district controls (X_d). These controls do not, however, capture the across district variation in male and female labor supply all of which are impounded in the district fixed effects. Thus, within a district, differential selection into the labor force across castes does not affect their individual wage.[13]

The second concern with caste composition as an instrument is that areas with greater low-caste households may have lower access to inputs, public goods and infrastructure (Banerjee and Somanathan 2007). Such areas may also have agro-ecological endowments which are unfavorable to agriculture. For these reasons, we include a comprehensive set of controls for irrigation, education, infrastructure (roads, electrification, banks), urbanization and agro-climatic endowments.

While there is no ex-ante way of knowing whether our controls are good enough, we can perform the following consistency check. Suppose conditional on our controls, the instrument is still correlated with omitted variables that affect the demand for agricultural labor. Then the caste composition also ought to have an effect on the demand for male labor. This can be easily checked from the first-stage regressions of the instrument variable procedure. As will be shown later, conditional on controls for agro-climatic endowments and infrastructure, caste composition does not have a statistically significant effect on the employment of male labor in agriculture.

A third possibility is that the caste composition in a district reflects long run development possibilities. In this story, the ‘higher’ castes used their dominance to settle in better endowed regions. Once again, this would require adequate controls for agro-ecological conditions. Finally, could caste composition itself be influenced by wages? Anderson (2011)

[13] In another set of regressions, we control for the interaction of caste with the education and the age of an individual. The earnings for low caste women are lower than that of others for educations levels of graduate and higher.

argues that village level caste composition in India has remained unchanged for centuries and location of castes is exogenous to current economic outcomes. This is, of course, entirely consistent with the low levels of mobility in India noted earlier.

2.5.2 Identification of the Impact of Male Labor Supply

For male labor supply, this chapter uses, as instrument, the district proportion of men (in the age group 15-59) employed in non-farm manufacturing and mining units with a workforce of at least 20. The relation between this instrument and district level male employment in agriculture is plotted in figure 2.5. The negative association visible in the graph is consistent with the proposition that competition from non-farm jobs reduces labor supply to agriculture (Lanjouw and Murgai 2009). Rosenzweig's (1978) study of agricultural labor markets also uses indicators of non-farm economy as an instrument for labor supply to agriculture.[14] However, not all non-farm activity can be considered to be exogenous to agriculture. We define our instrument to include employment in manufacturing and mining sectors, and further restrict it to only large scale units. Our case, elaborated below, is that employment in the non-traded sectors and in small enterprises is endogenous to agricultural development but that is not so for large enterprises in traded sectors.

The rural non-farm sector is known to be heterogeneous. Some non-farm activity is of very low productivity and "functions as a safety net – acting to absorb labor in those regions where agricultural productivity has been declining – rather than being promoted by growth in the agricultural sector" (Lanjouw and Murgai 2009). These are typically service occupations

[14] The variables used by Rosenzweig (1978) are the number of factories and workshops per household, percentage of factories and workshops employing five or more people and the percentage of factories and workshops using power.

with self-employment and limited capital. It is clear that such non-farm activity is endogenous to agricultural wages.

The other case is when a prosperous agriculture stimulates demand for non-farm activity. This type of non-farm employment tends to be concentrated in the non-traded sector of retail trade and services and mostly in small enterprises. Using a village level panel data set across India, Foster and Rosenzweig (2003) argue that non-traded sectors are family businesses with few employees while factories are large employers and frequently employ workers from outside the village in which they are located. In a companion paper, they state that on an average non-traded service enterprises consist of 2-3 workers. This is no different from the international experience of developing countries (World Bank 2008, Chapter 9).

Column 1 in table 2.3 presents the sectoral distribution of non-farm employment in production units with workforce of size 20 or more. This can be compared to the sectoral distribution of non-farm employment in production units with workforce of size nine or less in column 2 of table 2.3. It can be seen that the tradable sectors of manufacturing and mining account for a substantially larger proportion of large work units while non-tradable sectors such as trade and hotels, transport and construction are less important. These considerations dictate that a valid instrument that captures withdrawal of labor from the farm sector would measure non-farm employment in large units and in the traded sectors.

Even though the tradable non-farm goods and services sectors do not depend on local demand, this variable could still be invalid if large non-farm enterprises locate in areas of low agricultural wages. This possibility is suggested in the work of Foster and Rosenzweig (2004). They analyze a panel data set over the period 1971-1999 collected by the National Council of Applied Economic Research (NCAER). This data suggests a much higher expansion of rural non-farm activity than that implied by the nationally representative employment survey data of NSS (Lanjouw and Murgai 2009). To see if the non-farm sector gravitates towards agriculturally depressed areas in this data set, Lanjouw and Murgai (2009) estimate the impact of growth in agricultural yields on growth in non-farm sector employment. They take growth in agricultural yields as a proxy for agricultural productivity and do not find a negative relationship between manufacturing employment and yield growth. They find a positive association between the two in the specification with state fixed effects and no other district controls. However, the addition of region fixed effects makes the positive relation also disappear.

Therefore, if anything, the traded non-farm sector grew more in areas that were relatively agriculturally advanced. One explanation for this has been provided by Chakravorty and Lall (2005). They analyze the spatial location of industries in India in the late 1990s and find that private investment gravitates towards already industrialized and coastal districts with better infrastructure. No such pattern is seen for government investment. The significance of geographical clusters is that it makes initial conditions of agricultural productivity and infrastructure important in determining future investments. This implies that estimation of labor demand equation should include adequate controls for infrastructure to sustain the validity of the instrument.

Again, the adequacy of controls that ensures validity of the non-farm employment instrument may be hard to judge ex-ante. However, if non-farm employment instrument is correlated with omitted variables that affect overall agricultural labor demand, then the instrument ought to be significant in the first-stage regression for female employment. As we show later, this consistency check shows that non-farm employment in large manufacturing and mining units is not a significant explanatory variable for female employment in agriculture.

2.6 Data

The key data this chapter uses is from the nationally representative Employment and Unemployment survey of 2004-05 (July 2004 to June 2005) conducted by the National Sample Survey Organization. NSS is a cross-sectional dataset which is representative of India's population. The survey contains labor force participation and earnings details for the reference period of a week (preceding seven days from the date of survey) and follows a two stage sampling design. In rural areas, the first stratum is a district. Villages are primary sampling units (PSU) and are picked randomly in a district over an entire agricultural year (July to June) over quarters to ensure equal spacing of observations across the year. The households are randomly chosen in the selected PSU's. The district level analysis includes 15 major states in the sample: Punjab, Haryana, Uttar Pradesh (includes Uttarakhand), Madhya Pradesh (includes Chattisgarh), Bihar (includes Jharkhand), Gujarat, Rajasthan, West Bengal, Assam, Maharashtra, Andhra Pradesh, Karnataka, Orissa, Tamil Nadu and Kerala.

Daily wage rate for each individual is calculated by dividing the total earnings by the total number of days worked in agriculture in the last week. Average agricultural wage in a district is estimated by calculating an average of daily wage rate for all casual wage laborers in agriculture in a district, weighted by the sampling weights provided in the survey so that the district level average wage rate is representative of the district population. Wages are adjusted for differences in price levels across the states by using the consumer price index for agricultural laborers as the deflator. District level employment is estimated by summing up the total days worked in agriculture in the last week by all individuals in a district weighted by the sampling weights and is divided by total area under cultivation in a district. Some of the other variables including the instruments are also constructed from this data set. The control variables are obtained from a variety of sources (see table 2.5).

The first set of control variables relate to agriculture: irrigation, inequality in land holdings, rainfall, agro-climatic endowments, and land allocation to various crops. The agro-climatic variables are derived from a classification of the country into 20 agro-ecological zones (AEZ) described in table 2.4 (Palmer-Jones and Sen 2003). The independent variables are computed by taking the proportion of area of a district under a particular AEZ. A second set of control variables relate to infrastructure: roads, electrification and banking. A third set of variables relate to education and urbanization. Table 2.5 contains a description of all the variables, their definitions and descriptive statistics.

The district-level regressions are weighted by district population and the standard errors are robust and corrected for clustering at state-region level. In some districts, there are very few wage observations. To avoid measurement error in the dependent variable, the districts where the number of wage observations for either males or females was less than 5 were dropped from the analysis. Dropping districts where either male or female observations are few in number results in a data set with equal observations for males and females. However, this could lead to a biased sample as the districts where female participation in the casual labor market is the least are most likely to be excluded from the sample. To see whether such selection matters, we also estimate male labor demand function for districts in which number of male wage observations are at least five (ignoring the paucity, if any, in the number of female observations) and similarly estimate female labor demand function for districts in which number of female wage observations are at least five (ignoring the paucity, if any, of male wage observations).

2.7 Main Findings

Panel A of table 2.6 shows the system two stage least squares (2SLS) estimates of inverse demand functions for total male and female labor in agriculture. The first specification considers only the agriculture controls of irrigation, land inequality, rainfall, agro-ecological endowments and allocation of land to various crops. In the second specification, we add the infrastructure controls of roads, electrification and banking. The final specification includes the controls for education and urbanization.[15] Panel B of table 2.6 shows the coefficients of the instruments in the first-stage reduced form regressions for each of these three specifications. Panel C of table 2.6, displays the coefficients of the labor supply variables from an ordinary least squares estimation of $(2E.1)$ and $(2E.2)$.

In panel B of table 2.6, for all specifications, we find a significantly positive association between proportion of low caste households in a district and female employment in agriculture. Similarly, a greater presence of large scale non-farm enterprises in manufacturing and mining sectors decreases male employment in agriculture significantly in all the specifications. The F-statistic for the instruments is reported in the bottom of Panel A and it is significant at 5 percent level for female labor supply and at 1 percent level for male labor supply. The first-stage regressions thus confirm the causal story about these variables: status norms govern female labor supply and non-farm opportunities are primarily received by men.

Note also that the proportion of low caste households does not affect employment of male labor in agriculture and presence of large scale non-farm manufacturing and mining enterprises does not affect female labor employment in agriculture significantly. The

[15] Appendix 2.A shows the first-stage coefficients of the other control variables for the final specification in table 2.6.

significance of this observation is that if, despite the controls, the instruments retained some residual correlation with demand for agricultural labor, then we would expect both instruments to be significant in the first-stage reduced form regressions for male and female employment. The fact that this is not so supports the case that these are valid instruments for labor supply to agriculture. Returning to the labor demand equations, the system 2SLS estimates of the effect of female and male labor supply on own wage rates in panel A of table 2.6 are larger in magnitude and statistically more significant than the OLS estimates in panel C of table 2.6, and have the expected negative signs for own effects. [16] The coefficients of the labor supply variables do not change much between the three specifications in Panel A. The agriculture controls seem to be the most important in removing the correlation between agricultural labor demand and the instruments.

The cross effects of labor supply on wage rates are negative in sign. This implies that males and females are substitutes in agriculture. However, male labor and female labor are not perfect substitutes. In the system 2SLS regressions with full set of controls (the third specification), female labor supply has a significant impact on female wage with an inverse demand elasticity of – 0.52. However, the impact of female labor supply on male wage is smaller (around -0.1) and is not significantly different from zero. Thus, an increase in female labor supply by 10% decreases female wage by 5.2%, male wage by 1.3% and decreases the female to male wage ratio by 4%. To test formally that the impact on female wage is greater (in absolute terms) than the impact on male wage, we carry out a chi-square-test. In all of the specifications, the chi-square-test rejects the null that the coefficients are equal against the alternative that the coefficient of female labor supply in the female wage regression is higher than the coefficient of female labor supply in the male wage regression. This is supportive of

[16] By the Durbin-Wu-Hausman test, the null hypothesis that the employment variables can be treated as exogenous is rejected for all specifications (at 10 % significance level).

the Boserup hypothesis that caste driven variation in female labor supply leads to variation in the gender wage gap in agriculture across regions of India. In particular, greater female work participation decreases female wage relative to male wage.[17]

In contrast, the effect of male labor supply variation is significantly negative for both male and female wage rates. In the third specification with the full set of controls, the point estimate of the inverse demand elasticity is -0.37 for female and -0.28 for male wage with respect to male labor supply. Although large scale non-farm employment is dominated by men, non-farm labor demand has favorable effects on female and male wage rates. The point estimates would imply that a 10% decrease in male labor supply increases male wage by 2.8%, female wage by 3.7% and increases the female to male wage ratio by 1%. A chi-square test however, does not reject (in all the specifications) the null of equality of the two coefficients in the male and female inverse demand functions for male labor supply. Hence, a decrease in male labor supply to agriculture has no significant impact on gender wage gap in agriculture.

There is, thus, an asymmetry between the effects of gender specific variation in labor supply on the wage of the opposite gender. Male labor supply matters to female wage but the effect of female labor supply on male wage is small and insignificant. Why is this so? The theoretical model posited in section 2.3 predicts that the elasticity of female wage with respect to male employment relative to the similar cross elasticity of male wage is the product of two ratios: the ratio of male to female labor employment and the male to female wage ratio. The sample estimate for male and female labor employment is 5.17 and 2.57 days per

[17] We also estimated Rosenzweig's specification for our data set with instruments that are as close as possible to those employed by him. In these results, female labor supply has a significant negative impact on both female and male wage but not on the gender wage gap. This matches the finding of Rosenzweig for the 1961 data. We also find that male labor supply does not have a significant impact on the gender wage gap even though the impact of male labor supply on male wage is significant and negative and on female wage is insignificant. In Rosenzweig's earlier analysis, male labor supply had an insignificant impact on male and female wages and therefore did not matter to the gender wage gap.

week per hectare of land respectively while the sample estimate for male and female wage is Rs 47.3 and Rs 36.13 per day respectively. This gives an estimate of relative cross-wage elasticities to be 2.63. The results in panel A of table 2.6, for the specification with the full set of controls, yield an econometric estimate of the ratio of cross-wage elasticities as 2.84 which is close to the prediction from the theoretical model.

The control variables could also have an effect on the gender wage gap. To ascertain this, a chi-square test was conducted to test for the equality of coefficients for each control variable across male and female demand equations. The null hypothesis of equality of coefficients is rejected at 5 % level of significance for rice cultivation, access to roads and landholding inequality. Rice growing areas have a higher demand for female labor which leads to a higher wage rate for women and translates into a lower gender wage gap. Many researchers have documented greater demand for female labor in rice cultivation due to greater demand for females in tasks like transplanting and weeding (Mbiti 2007) and this result validates their observations. On the other hand, access to roads seems to increase demand for only male labor resulting in a larger wage gap between females and males in districts with better access to roads. Landholding inequality measured by the Gini coefficient for a district affects demand for both males and females significantly negatively reflecting the well known feature that large farms use less labor per unit of land than small farms. Moreover, women are more adversely affected by men resulting in a larger gender wage gap in districts with higher land inequality. Theoretically, the effect of landholding inequality on gender wage differential is ambiguous (Rosenzweig 1978).

A concern with the 2SLS results is that the first-stage F-statistic though significant is not very large. Weak-instruments could lead to biased estimates and to finite sample distributions that are poorly approximated by the theoretical asymptotic distribution. While such concerns are greater in an over-identified model, the weak-instrument critique suggests caution in interpreting the 2SLS results. As a check for just identified models with possibly weak-instruments, Angrist and Pischke (2008) and Chernozhukov and Hansen (2008) recommend looking at the reduced form estimates of the dependent variable (male and female wage) on all exogenous variables (X_d, Z_d) since they have the advantage of being unbiased. Chernozhukov and Hansen (2008) formally show that the test for instrument irrelevance in this reduced form regression can be viewed as a weak-instrument-robust test of the hypothesis that the coefficient on the endogenous variable in the structural equation is zero. The sign and the strength of the coefficients in the reduced form regression can provide evidence of whether a causal relationship exists.

Panel D of table 2.6 shows the results for the coefficients of instruments from the reduced form regression of male and female wage on instruments and other covariates. The instruments are significant in this regression and so it can be concluded that the weak-instrument problem does not contaminate the inference from the structural regressions. It can be seen that an increase in proportion of low caste households reduces only the female wage. This is entirely consistent with the 2SLS results where the instrument increases only female labor supply (the first-stage regression) which in turn has a significantly negative impact only on female wage. On the other hand, large scale industrial employment has a significantly positive impact on male and female wage rates. This is also in line with the 2SLS results where the presence of large enterprises in the non-farm sector decreases only male labor supply to agriculture which in turn impacts both male and female wage positively.

2.8 Robustness Checks

The third specification in Panel A of table 2.6 is our baseline and we consider the robustness of its estimates to alternative specifications. Panel A of table 2.7 adds more agriculture controls: fertilizer usage per unit of cultivated land and implements (consisting of tractor and power operated tools) per unit of cultivated land. Including fertilizers (first specification) does not change the impact of female labor supply on male and female wage and a 10% increase in female labor supply increases the gender wage gap by 3.6%. The chi-square test does not reject the equality of male labor supply coefficients across male and female labor demand equations but rejects the equality of female labor supply coefficients. The inclusion of fertilizers does, however, reduce the coefficient of irrigation in both equations to the point that it becomes insignificant in the female labor demand equation. This is possibly because of a high positive correlation (0.4) between irrigation and fertilizer use. Controlling for implements used per unit land cultivated (specification 2) does not change any of the principal findings of the base specification. Again, the chi-square test does not reject the equality of male labor supply coefficients across male and female demand equations but rejects the equality of female labor supply coefficients.

In a second robustness check, we control for male and female health in rural areas. Nutrition status can affect productivity which in turn could impact rural wage. If nutrition status is correlated with our instrumental variable of low caste composition, then it could bias our results as well. Adult measures of health in India are not available at district level. Weight and height measurements are available at state level from the National Family and Health Survey of 2005-06. The measure of under-nutrition is percentage of rural adults with a body mass index of less than 18.5. Panel B of table 2.7 shows the structural estimates for the total demand for labor with state level health controls. The results from the base specification continue to hold. While increase in female labor supply increases the gender wage gap significantly, male labor supply has no impact.

As a third check, we reconsider our sample selection rule. Recall, that we chose districts for which there were at least five observations for female as well as male wages. While this ensures an equal sample size for males and females, it also entails a risk of dropping districts where female participation in wage work is the least. To check robustness, we consider the following alternative. For the male worker sample, we considered all districts where there are at least five observations for male wages. Similarly, for the female worker sample, we included all districts where there are at least five observations for female wages. This increases the number of districts from 279 in the matched sample to 359 for males and to 288 for females. Table 2.8 shows the estimates from the baseline specification on this enlarged sample. The estimates validate our central result that the gender wage gap is sensitive to female labor supply and not to male labor supply. In fact, the effect of female labor supply on gender wage gap in the enlarged sample is greater. A 10% increase in female labor supply results in a 4.8% decline in female to male wage ratio in the enlarged sample compared to 4% in the matched sample.

In a fourth robustness check, we control for differential participation in tasks by males and females across districts. As noted earlier, some agricultural tasks are traditionally deemed as male while others are dominated by women. In section 2.3, we showed that the gender wage gap in Indian agriculture is within tasks. A very small percentage of the wage gap can be attributed to differential participation of men and women across tasks. To address this issue formally, we regress individual wages on individual characteristics (age, age square, education dummies, and marital status dummies), district level female and male labor employment in agriculture (suitably instrumented), other district controls and dummy variables for agricultural tasks for which the wage is recorded. The agricultural tasks are ploughing, sowing, transplanting, weeding, harvesting and other agricultural activities.

The estimates are reported in table 2.9. They show that a 10% increase in female labor supply reduces female wage by 5.5% and has no significant effect on male wage. Male labor supply on the other hand has an identical negative effect on male and female wage.

Lastly, we consider the possibility that hired and family labor may not be equally efficient. Family labor may be more efficient because of better incentives. If this is so, a simple aggregate of family and hired labor is not valid and could lead to inconsistent estimates. Suppose one unit of hired labor is equivalent to θ units of family labor (with θ less than one). Then in terms of efficiency units of family labor, the total labor supply is $L^{f,S} + \theta L^{o,S}$, where $L^{f,S}$ and $L^{o,S}$ are the aggregate labor supply to home farm and to outside farms. In the regressions, we have measured labor supply as $\ln(L^{f,S} + L^{o,S})$. Since, $\ln(L^{f,S} + \theta L^{o,S})$ = $\ln(L^{f,S} + L^{o,S})$ + $\ln[(L^{f,S} + \theta L^{o,S})/(L^{f,S} + L^{o,S})]$, the second term is absorbed in the error term of the regressions. This could lead to inconsistent estimates. The instruments will be correlated with $\ln\left(\frac{L^{f,S}+\theta L^{o,S}}{L^{f,S}+L^{o,S}}\right)$ if they not only affect the total labor supply but also the allocation of labor between own farm and outside farm. It is possible that low caste women have a greater propensity to work outside their family farm due to less social restrictions. Similarly, the opportunity of employment in manufacturing and mining could lead landed households to divert their labor supply to industry and increase hiring of labor on their farms.

To meet these concerns, we estimate the baseline specification for values of θ ={0.5, 0.7, 0.9}, for both male and female labor. The results are shown in table 2.10. The last column shows the results for θ =1 which corresponds to the results of the base specification in table 2.6. As the value of θ decreases, the impact of female labor supply on male wage does not change but the impact of female labor supply on female wage falls in magnitude. The chi-square test for the equality of the impact of female labor supply on female and male wage continues to be rejected for the selected values of θ. A decrease in the value of θ increases the impact of male labor supply on both male and female wage. Once again, the chi-square test for the equality of the impact of male labor supply on male and female wage is not rejected for the selected values of θ.

2.9 Explaining the Difference in Gender Wage Gap between Northern and Southern States of India

While our findings support the Boserup hypothesis, there are other factors as well that matter to the gender wage gap. So to what extent does the Boserup hypothesis, i.e., the difference in female work participation across northern and southern states in India explain the observed difference in the gender wage gap?

From estimation equations $(2E.1)$ and $(2E.2)$, the gender wage gap in a southern state can be written as

$$\overline{W}_{M,s} - \overline{W}_{F,s} = (\hat{\alpha}_0 - \hat{\alpha}_1)\bar{L}_{F,s} + (\hat{\beta}_0 - \hat{\beta}_1)\bar{L}_{M,s} + (\hat{\gamma}_0 - \hat{\gamma}_1)\bar{X}_s + \left(\bar{\varepsilon}_{M,s} - \bar{\varepsilon}_{F,s}\right) \tag{2.8}$$

where, W is the log of wages, L is the log of labor supply and X are other district level covariates included in the empirical analysis. 'M' and 'F' index males and females respectively. Similarly, the gender wage gap in a northern state can be written as

$$\overline{W}_{M,n} - \overline{W}_{F,n} = (\hat{\alpha}_0 - \hat{\alpha}_1)\bar{L}_{F,n} + (\hat{\beta}_0 - \hat{\beta}_1)\bar{L}_{M,n} + (\hat{\gamma}_0 - \hat{\gamma}_1)\bar{X}_n + \left(\bar{\varepsilon}_{M,n} - \bar{\varepsilon}_{F,n}\right) \tag{2.9}$$

Subtracting (2.9) from (2.8), we obtain

$$\left(\overline{W}_{M,s} - \overline{W}_{F,s}\right) - \left(\overline{W}_{M,n} - \overline{W}_{F,n}\right) = (\hat{\alpha}_0 - \hat{\alpha}_1)(\bar{L}_{F,s} - \bar{L}_{F,n}) + (\hat{\beta}_0 - \hat{\beta}_1)(\bar{L}_{M,s} - \bar{L}_{M,n}) + (\hat{\gamma}_0 - \hat{\gamma}_1)(\bar{X}_s - \bar{X}_n) + \left(\bar{\varepsilon}_{M,s} - \bar{\varepsilon}_{F,s}\right) - \left(\bar{\varepsilon}_{M,n} - \bar{\varepsilon}_{F,n}\right) \tag{2.10}$$

The ratio $[(\hat{\alpha}_0 - \hat{\alpha}_1)(\bar{L}_{F,s} - \bar{L}_{F,n})]/\ [\left(\overline{W}_{M,n} - \overline{W}_{F,s}\right) - \left(\overline{W}_{M,n} - \overline{W}_{F,n}\right)]$ is the proportion of the difference in wage gap across north and south that is explained by the difference in female labor supply.

To implement this, we let the variables take the average values of northern and southern states respectively.[18] The mean values are listed in table 2.11. The parameters are the coefficient estimates of the base specification estimated in column 3 of table 2.6 (panel A). Table 2.12 shows the proportion of the gender wage gap explained by each right hand side variable. The proportions for agro-ecological zones have not been shown for brevity. One can see that 55 percent of the regional difference in the gender wage gap is because of the larger female labor supply in the southern states. Greater land inequality and lower area under cultivation of rice in the southern states are other important and significant factors which lead to a greater gender wage gap in the south. On the other hand, greater electrification, lower male supply and the greater importance of coarse cereal crops (sorghum and millets) should lead to a lower gender wage gap in the south but these do not affect the gender wage gap significantly in the regressions.

[18] We classify Andhra Pradesh, Karnataka, Kerala, Maharashtra and Tamil Nadu as southern states while Assam, Bihar, Gujarat, Punjab, Haryana, Rajasthan, Uttar Pradesh, Madhya Pradesh and West Bengal are classified as northern states. Orissa is omitted from the north-south analysis since it does not fall clearly into any of the categories and also is geographically sandwiched between the North and the South.

2.10 Conclusion

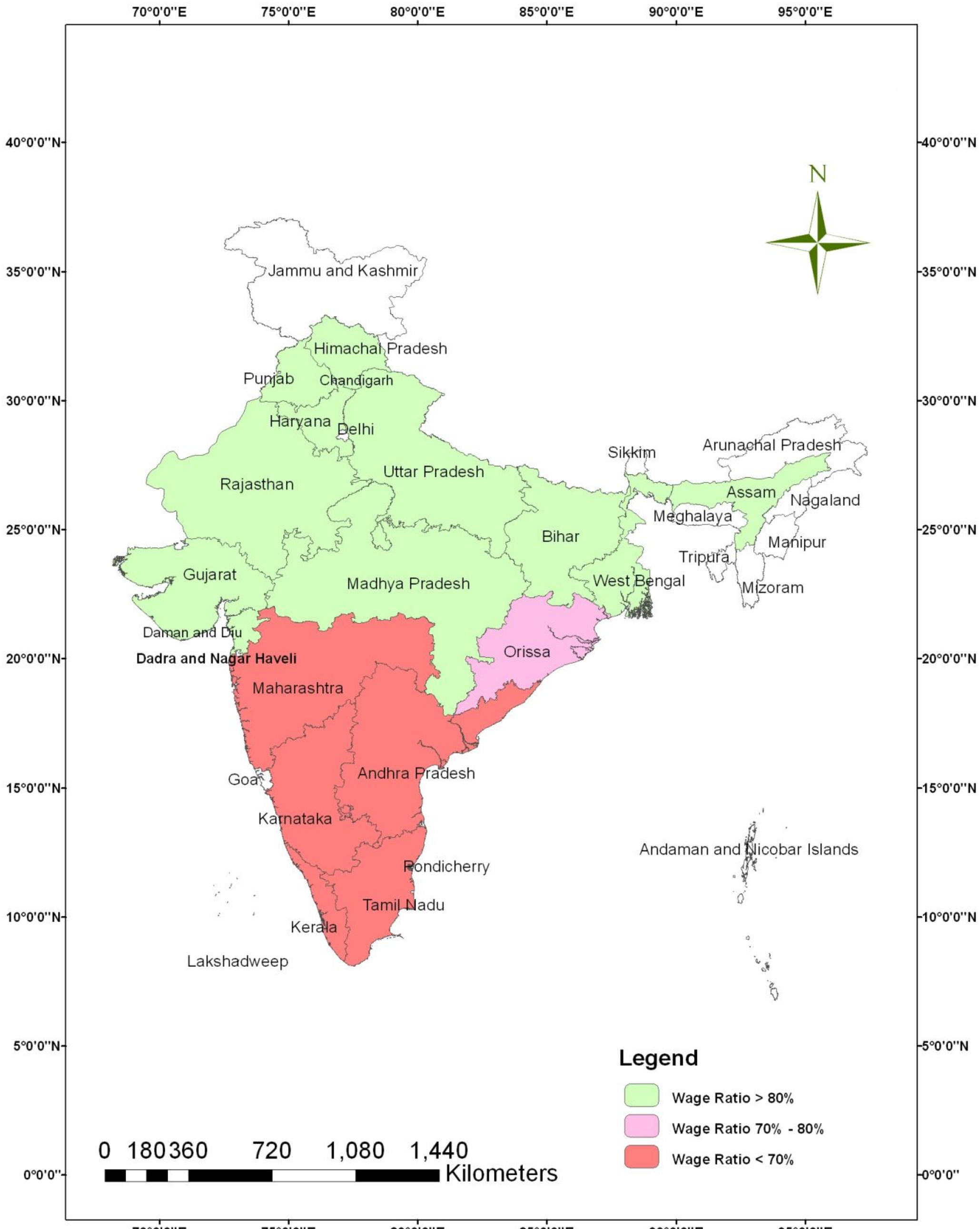
The effect of variation in female work force participation on the gender wage gap in the developed countries has been explored in recent papers. In a developing country context, such a connection was made by Boserup many decades ago. Based on data from 1950s, she posited that the gender wage gap was higher in the southern states of India relative to the northern states because of greater female labor supply in south India, which stemmed from differences in cultural restrictions on women’s participation in economic activity. This chapter confirms the hypothesis within a neo-classical framework of labor markets. Compared to the literature, this chapter also pays attention to the variation in male labor supply and how that impacts the gender wage gap. The exogenous variation in labor supply is identified by spatial variation in caste composition and non-farm employment of men in large units.

We find that female labor supply has a sizeable effect on female wage but not on male wage. This result thus has important implications for the literature on gender wage differentials. It shows that the usual approach of attributing the gender wage gap to only individual characteristics or discrimination is incomplete. The overall labor market structure that determines labor supply and the substitutability between female and male labor may also have a significant impact on gender wage inequality.

The chapter also found that male labor supply has sizeable effects on male as well as female wage. This finding is interesting on three counts. First, it provides a causal effect of withdrawal of males from agriculture due to non-farm employment opportunities on wages of men and women. The chapter, therefore, sheds light on the economic processes that affect agricultural wage (Lanjouw and Murgai 2009; Eswaran et al. 2009; Foster and Rosenzweig 2003). Second, the strong effect of male labor supply on female wage is of independent interest since the sectoral mobility of women from the farm to the non-farm sector is much less marked compared to men (Eswaran et al. 2009). This is because of lower education levels as well as societal constraints that limit female participation in most non-farm jobs. This raises a concern that rapid growth in the non-farm sector does not entail much gain for women. Our finding, however, suggests that there is enough substitutability between men and women in the agricultural production process that a withdrawal of men from agriculture has positive effects on male and female wages.

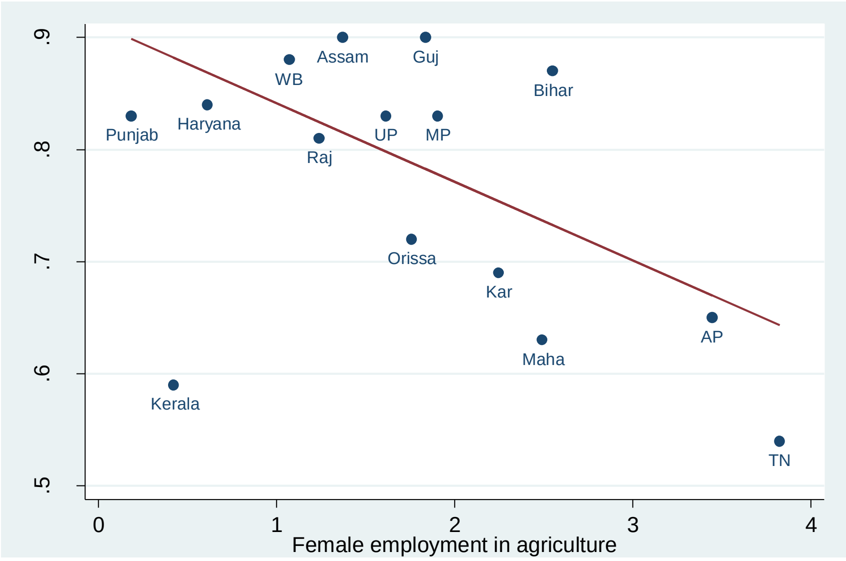
Finally, the findings point to a marked asymmetry between the effects of female and male labor supply. Female labor supply does not impact male wage significantly but male labor supply does move female wage significantly. A standard neo-classical model predicts this asymmetry and its magnitude is determined by the gender gap in wage and the gender gap in labor supply. The findings match the prediction closely.

Figure 2.1 Female to Male Wage Ratio Across the Indian States (2004-05)



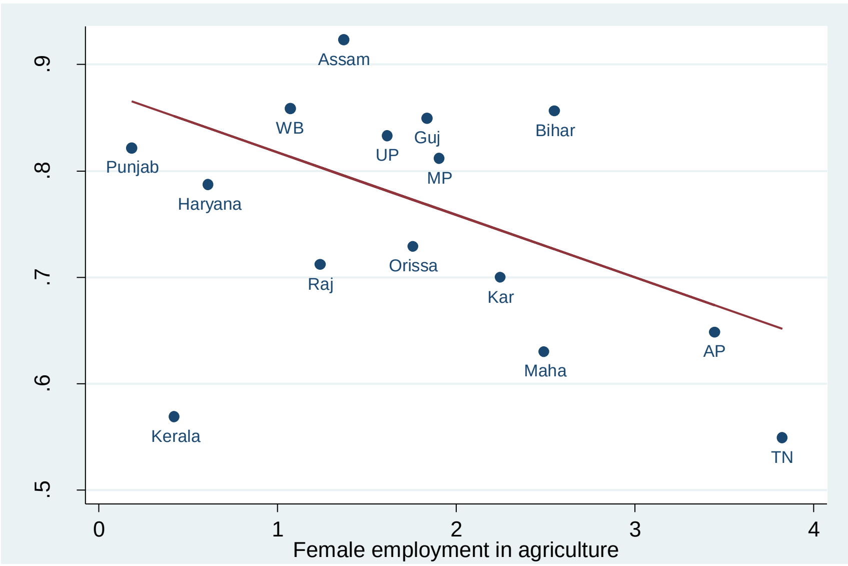
Source: NSS 2004 Employment and Unemployment Schedule (Authors' calculations).
Note: Ratio of female to male daily wage rates for agricultural workers in rural India aged 15-59 is plotted in the above figure.

Figure 2.2 Female Employment in Agriculture and Female to Male Wage Ratio



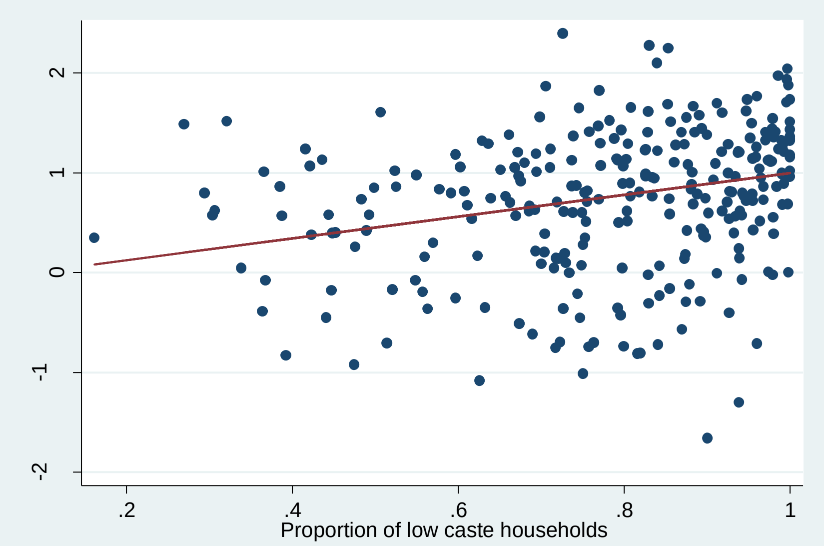
Source: NSS 2004 Employment and Unemployment Schedule (Authors' calculations).
Note: Labor employment is measured as total days worked in a reference week per unit land under cultivation in a state. Population weighted regression lines are fitted to the above plots.

Figure 2.3 Female Employment in Agriculture and the Re-weighted Female to Male Wage Ratio



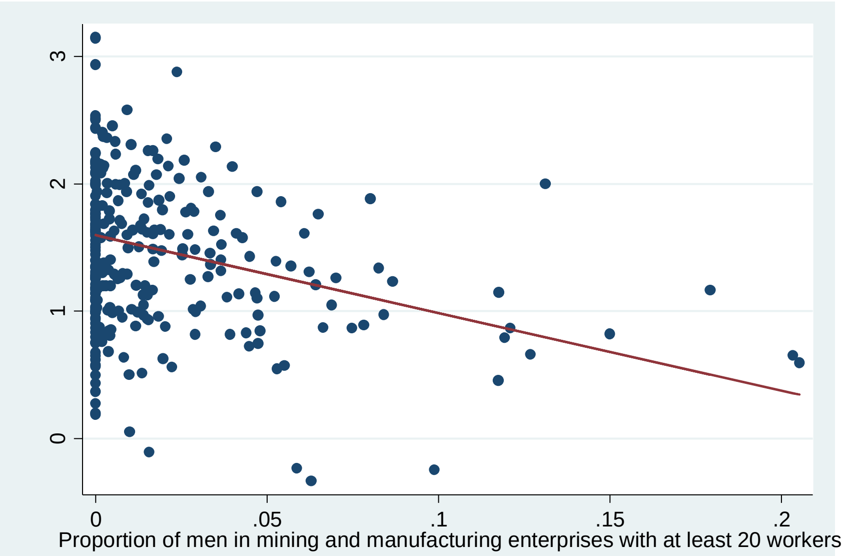
Source: NSS 2004 Employment and Unemployment Schedule (Authors' calculations).
Note: Labor employment is measured as total days worked in a reference week per unit land under cultivation in a state. Population weighted regression lines are fitted to the above plots.

Figure 2.4 Low Caste Households and Female Employment in Agriculture



Source: NSS 2004 Employment and Unemployment Schedule (Authors' calculations).
Note: Labor employment is measured as total days worked in a reference week per unit land under cultivation in a district. Population weighted regression lines are fitted to the above plots.

Figure 2.5 Large Scale Industrial Employment and Male Employment in Agriculture



Source: NSS 2004 Employment and Unemployment Schedule (Authors' calculations).
Note: Labor employment is measured as total days worked in a reference week per unit land under cultivation in a district. Population weighted regression lines are fitted to the above plots.

Table 2.1 Gender Wage Gap in Indian Agriculture

	Wage (1)		Wage (2)	
Female	-0.35***	(0.03)	-0.33***	(0.03)
Age	0.02***	(0.00)	0.02***	(0.00)
Age square	-0.00***	(0.00)	-0.00***	(0.00)
Below primary	0.06***	(0.02)	0.06**	(0.02)
Primary	0.05*	(0.02)	0.05*	(0.02)
Middle	0.03	(0.03)	0.02	(0.03)
Secondary	0.04	(0.03)	0.04	(0.03)
Senior secondary and above	-0.03	(0.03)	-0.03	(0.03)
Married	-0.02	(0.02)	-0.01	(0.02)
Widowed	-0.06**	(0.03)	-0.05	(0.03)
Divorced	-0.13***	(0.04)	-0.11**	(0.05)
Sowing			-0.17**	(0.06)
Transplanting			-0.04	(0.05)
Weeding			-0.20***	(0.04)
Harvesting			-0.12***	(0.04)
Other cultivation			-0.11***	(0.03)
Constant	3.37***	(0.05)	3.50***	(0.06)
Observations	14,190		14,190	
R-square	0.21		0.22	

Source: NSS 2004 Employment and Unemployment Schedule
Note: The above table reports the results from an OLS regression of individual wage on individual characteristics. Log of individual wage for each task that a person undertakes is the dependent variable. Robust standard errors clustered at state-region level are in parenthesis; ***, ** and * indicate significance at the 1, 5 and 10% levels respectively. Districts having at least 5 wage observations for males and females are included.

Table 2.2 Effect of Caste on Female Wage

	Female wage	
	Coefficient	Standard Error
Low caste	-0.00	(0.01)
Age	0.01**	(0.00)
Age square	-0.00**	(0.00)
Below primary	0.01	(0.02)
Primary	0.02	(0.02)
Middle	0.02	(0.02)
Secondary	0.01	(0.04)
Senior secondary and above	0.13***	(0.04)
Married	0.00	(0.02)
Widowed	-0.01	(0.02)
Divorced	-0.05	(0.04)
Sowing	-0.01	(0.08)
Transplanting	0.08	(0.07)
Weeding	-0.03	(0.07)
Harvesting	0.04	(0.07)
Other cultivation	0.02	(0.06)
Constant	3.23***	(0.08)
District fixed effect	Yes	
Observations	6,377	
R-squared	0.49	

Source: NSS 2004 Employment and Unemployment Schedule
Note: The above table reports the results from an OLS regression of individual female wage on individual characteristics. Log of individual wage for each task that a person undertakes is the dependent variable. Robust standard errors clustered at state-region level are in parenthesis; ***, ** and * indicate significance at the 1, 5 and 10% levels respectively. Districts having at least 5 wage observations for females are included.

Table 2.3 Sectoral Distribution of Non-Farm Employment

Industry	Percentage in units with 20 or more workers (1)	Percentage in units with 9 or less workers (2)
Allied activities in agriculture	1%	7%
Fishing	0%	1%
Mining	7%	1%
Manufacturing	44%	20%
Construction	11%	17%
Trade and hotels	3%	28%
Transport	9%	12%
Finance and real estate	3%	2%
Public administration	22%	11%
Domestic services	0%	1%

Source: NSS 2004 Employment and Unemployment Schedule (Authors' calculations).
Note: The above figures are calculated from the usual status activity status of respondents. The sample includes men aged 15-59.

Table 2.4 Agro-Ecological Zones (AEZ)

AEZ	Description
2	Western Plain, Kachch and part of Kathiarwar, peninsular, hot arid ecoregion, with desert and saline soils and LGP (Length of Growing Period) <90 d
3	Deccan Plateau, hot arid ecoregion, with red and black soils and LGP < 90 d
4	Northern Plain and Central Highlands including Aravelli hills, hot semi-arid ecoregion with alluvium derived soils and LGP 90-150 d
5	Central Highlands, Gujarat Plains, Kathiarwar peninsular, hot arid ecoregion, with medium and deep black soils and LGP 90-150 d
6	Deccan Plateau, hot semi arid ecoregion, with mainly shallow and medium but some deep black soils and LGP 90-150 d
7	Deccan Plateau of Telengana and Eastern ghats, hot semi-arid ecoregion with red and black soils and LGP 90-150 d
8	Eastern Ghats, Tamil Nadu uplands and Deccan (Karnataka) Plateau, hot semi arid ecoregion with red loamy soils and LGP 90-150 d
9	Northern Plain, hot subhumid (dry) ecoregion with alluvium derived soils and LGP 150-180 d
10	Central Highlands (Malwa, Bundelkhand, an Eastern Satpura), hot subhumid ecoregion, with black and red soils and LGP 150-180 d up to 210 d in some places
11	Eastern Plateau (Chattisgarh), hot subhumid ecoregion, with red and yellow soils and LGP 150-180 d
12	Eastern (Chotanagpur) plateau and Eastern Ghats, hot subhumid ecoregion with red and lateritic soils and LGP 150-180 to 210 d
13	Eastern Gangetic Plain, hot subhumid (moist) ecoregion, with alluvium derived soils and LGP 180-210 d
14	Western Himalayas, warm subhumid(to humid and perhumid ecoregion) with brown forest & podzolic soils, LGP 180-210+d
15	Bengal and Assam Gangetic and Brahmaputra plains , hot subhumid (moist) to humid (and perhumid) ecoregion, with alluvium derived soils and LGP 210+ d
16	Eastern Himalayas, warm perhumid ecoregion with brown and red hill soils and LGP 210+ d
17	Northeastern Hills (Purvachal), warm perhumid ecoregion with red and lateritic soils and LGP 210+ d
18	Eastern coastal plain, hot subhumid to semi-arid ecoregion, with coastal alluvium derived soils and LGP 210+ d
19	Western ghats and coastal plain, hot humid region, with red, lateritic and alluvium derived soils and LGP 210+d

Source: Gajbhiye and Mandal (2006).
Note: AEZ 1 and AEZ 20 are not present in any district included in our analysis.

Table 2.5 Variable Definition and Summary Statistics

	Variable	Definition	Mean	Standard deviation
Wage	Male wage	ln(Real average male casual manual worker wage in cultivation, aged 15-59 years)	3.82	0.28
	Female wage	ln(Real average female casual manual worker wage in cultivation, aged 15-59 years)	3.54	0.31
Labor supply	Male LS	$\ln\left(\frac{\text{Total days worked in a reference week in cultivation by males aged 15−59}}{\text{Area under cultivation}}\right)$	1.46	0.61
	Female LS	$\ln\left(\frac{\text{Total days worked in a reference week in cultivation by females aged 15−59}}{\text{Area under cultivation}}\right)$	0.73	0.71
Instruments	Low caste	Percentage SC/ST/OBC households	0.75	0.19
	Industry	Percentage men aged 15-59 engaged in a manufacturing or mining unit employing more than 20 workers	0.02	0.03
Agriculture	Irrigation	Percentage cultivated area irrigated	0.43	0.26
	Gini	Gini coefficient for land holding inequality	0.69	0.10
	Rainfall	Rainfall received during June to September 2004 in cms	8.30	5.41
	Coarse cereals	Percentage area under production of coarse cereals	0.16	0.19
	Cotton	Percentage area under production of cotton, jute, mesta, tobacco and sugarcane	0.08	0.11
	Oilseeds and Pulses	Percentage area under production of oilseeds and pulses	0.25	0.20
	Rice	Percentage area under production of rice	0.35	0.29
	Horticulture	Percentage area under production of horticulture crops	0.06	0.12
	Wheat	Percentage area under production of wheat	0.10	0.15
Infrastructure	Paved roads	Percent villages accessible by a paved road	0.66	0.24
	Electrified	Percent villages electrified	0.86	0.23
	Commercial bank	Percent villages having a commercial bank	0.09	0.13
Education and Urbanization	Primary-Middle male	Percentage Primary-Middle educated male aged 15-59	0.36	0.09
	Secondary male	Percentage Secondary or more educated male aged 15-59	0.23	0.09
	Primary-Middle fem	Percentage Primary-Middle educated female aged 15-59	0.25	0.10
	Secondary fem	Percentage Secondary or more educated female aged 15-59	0.11	0.07
	Urban	Percentage population in a district living in urban areas	0.27	0.18

Source: Wages, Labor supply, Gini, Education, Low caste, Large scale non-farm employment in manufacturing and mining units- National Sample survey 2004-05; Irrigation, Land under cultivation- Land Use Statistics 2004-05; Crop composition- Area, Production and Yield statistics 2004-05; Rainfall- India Water Portal 2004-05(data originally collected by Indian Meteorological Department); Agro Ecological Zones- Compiled by Richard Palmer-Jones and Kunal Sen; Urban, Paved roads, Electrified and Commercial banks- Census of India 2001, Village directory

Note: Weighted mean with weights equal to district rural population, Agro-Ecological Regions are described in table 2.4.

Table 2.6 Baseline Specification

Panel A: Aggregate Demand for Total Labor in Agriculture

	District Controls: Agriculture				District Controls: Agriculture District Controls: Infrastructure				District Controls: Agriculture District Controls: Infrastructure District Controls: Education & Urbanization			
	Male wage		Female wage		Male wage		Female wage		Male wage		Female wage	
	(1)				(2)				(3)			
Female LS	-0.08	(0.17)	-0.49*	(0.27)	-0.11	(0.17)	-0.54*	(0.31)	-0.13	(0.15)	-0.52**	(0.25)
Male LS	-0.29***	(0.09)	-0.35***	(0.12)	-0.23***	(0.09)	-0.36***	(0.14)	-0.28***	(0.09)	-0.37**	(0.15)
Irrigation	0.21*	(0.12)	0.30*	(0.17)	0.28**	(0.12)	0.41**	(0.19)	0.31**	(0.12)	0.41**	(0.20)
Gini	-0.52	(0.37)	-1.28**	(0.54)	-0.64*	(0.34)	-1.33**	(0.56)	-0.65*	(0.33)	-1.30**	(0.51)
Rainfall	-0.00	(0.01)	0.01	(0.01)	0.00	(0.00)	0.01	(0.01)	0.00	(0.01)	0.01	(0.01)
Paved roads					0.43***	(0.10)	0.05	(0.25)	0.47***	(0.11)	0.08	(0.23)
Electrified					-0.55***	(0.17)	-0.41*	(0.25)	-0.61***	(0.18)	-0.44*	(0.24)
Commercial bank					0.04	(0.20)	-0.01	(0.21)	0.04	(0.17)	-0.00	(0.21)
Primary-Middle female									-0.01	(0.27)	-0.15	(0.54)
Secondary female									0.39	(0.35)	0.39	(0.66)
Primary-Middle male									-0.28	(0.26)	-0.20	(0.40)
Secondary male									-0.16	(0.24)	0.04	(0.45)
Urban percent									-0.15**	(0.08)	-0.08	(0.16)
Constant	4.50***	(0.37)	4.64***	(0.49)	4.85***	(0.41)	5.08***	(0.69)	5.10***	(0.49)	5.16***	(0.76)
AEZ	Yes				Yes				Yes			
Land Allocation to crops	Yes				Yes				Yes			
Observations	279		279		279		279		279		279	
Under-id (p-val)	0.00		0.00		0.01		0.01		0.01		0.01	
F(excluded instruments) L^S_F	3.93		3.93		3.53		3.53		4.81		4.81	
F(excluded instruments) L^S_M	26.79		26.79		23.90		23.90		17.52		17.52	
Null Female labor supply has equal effect on male and female wage (at 5% level)	Reject				Reject				Reject			
Null Male labor supply has equal effect on male and female wage (at 5% level)	Accept				Accept				Accept			

Panel B: First-stage for Labor Supply by Males and Females to Agriculture

	Male LS		Female LS		Male LS		Female LS		Male LS		Female LS	
	(1)				(2)				(3)			
Low caste	-0.11	(0.19)	0.70**	(0.27)	-0.15	(0.20)	0.66**	(0.26)	-0.22	(0.19)	0.79***	(0.27)
Industry	-3.86***	(0.53)	-0.58	(0.77)	-3.68***	(0.55)	-0.29	(0.89)	-3.33***	(0.59)	-0.26	(0.97)
R-Square	0.69		0.53		0.70		0.54		0.71		0.54	
Observations	279		279		279		279		279		279	

Panel C: Ordinary Least Squares Estimates

	Male wage		Female wage		Male wage		Female wage		Male wage		Female wage	
	(1)				(2)				(3)			
Female LS	-0.07**	(0.03)	-0.15***	(0.04)	-0.06**	(0.03)	-0.15***	(0.04)	-0.06**	(0.03)	-0.15***	(0.04)
Male LS	-0.01	(0.05)	0.04	(0.05)	-0.01	(0.04)	0.05	(0.05)	-0.01	(0.04)	0.06	(0.05)
R-Square	0.62		0.62		0.68		0.63		0.69		0.64	
Observations	279		279		279		279		279		279	

Panel D: Reduced Form Estimates

	Male wage		Female wage		Male wage		Female wage		Male wage		Female wage	
	(1)				(2)				(3)			
Low caste	-0.02	(0.11)	-0.31**	(0.13)	-0.04	(0.10)	-0.30**	(0.13)	-0.04	(0.10)	-0.34**	(0.13)
Industry	1.15***	(0.35)	1.63***	(0.42)	0.89***	(0.33)	1.47***	(0.44)	0.98***	(0.34)	1.37***	(0.48)
R-Square	0.62		0.61		0.68		0.62		0.68		0.63	
Observations	279		279		279		279		279		279	

Note: Panel A reports two stage least squares estimates, instrumenting for labor supply of males and females using low caste and industry employment as defined in table 2.5. Log of wages and log of labor supply are used in the above regressions. Panel B reports the corresponding first-stage. Panel C reports the results from OLS regression of the dependent variable against total labor employed in agriculture with other controls the same as in Panel A. Panel D reports the results from a reduced form regression of the log wage on instruments with other controls the same as in Panel A. Robust clustered standard errors are in parenthesis; ***, ** and * indicate significance at the 1, 5 and 10% levels respectively. The unit of analysis is a district and districts having at least 5 wage observations for male and female each are included here.

Table 2.7 Aggregate Demand for Total Labor in Agriculture with Additional Controls

Panel A: Adding Controls for Fertilizer and Implements

	Male wage		Female wage		Male wage		Female wage	
	(1)				(2)			
Female LS	-0.10	(0.14)	-0.46**	(0.23)	-0.12	(0.15)	-0.52**	(0.26)
Male LS	-0.31***	(0.10)	-0.44***	(0.15)	-0.29***	(0.09)	-0.37**	(0.15)
Irrigation	0.25**	(0.11)	0.27	(0.17)	0.31**	(0.13)	0.40**	(0.20)
Gini	-0.66**	(0.33)	-1.31***	(0.48)	-0.64*	(0.34)	-1.28**	(0.51)
Rainfall	0.00	(0.00)	0.01	(0.01)	0.00	(0.01)	0.01	(0.01)
Paved roads	0.52***	(0.11)	0.18	(0.20)	0.49***	(0.12)	0.09	(0.23)
Electrified	-0.60***	(0.18)	-0.43*	(0.24)	-0.62***	(0.19)	-0.45*	(0.24)
Commercial bank	-0.02	(0.19)	-0.15	(0.19)	0.04	(0.18)	0.00	(0.21)
Primary-Middle female	-0.04	(0.26)	-0.23	(0.52)	-0.02	(0.27)	-0.16	(0.54)
Secondary female	0.07	(0.40)	-0.35	(0.65)	0.36	(0.33)	0.37	(0.65)
Primary-Middle male	-0.24	(0.25)	-0.13	(0.37)	-0.28	(0.26)	-0.20	(0.40)
Secondary male	-0.05	(0.25)	0.30	(0.47)	-0.14	(0.24)	0.06	(0.45)
Urban percent	-0.23***	(0.09)	-0.27	(0.17)	-0.15**	(0.07)	-0.08	(0.15)
Fertilizer	0.04**	(0.02)	0.10***	(0.03)				
Implements					0.08	(0.10)	0.06	(0.12)
Constant	5.13***	(0.50)	5.23***	(0.75)	5.06***	(0.50)	5.13***	(0.76)
AEZ	Yes				Yes			
Land allocation to crops	Yes				Yes			
Observations	279		279		279		279	
Under-id (p-val)	0.01		0.01		0.01		0.01	
F(excluded instruments) L^S_F	4.86		4.86		4.60		4.60	
F(excluded instruments) L^S_M	15.81		15.81		17.06		17.06	

Panel B: Adding Controls for Adult Health				
	Male wage		Female wage	
Female LS	-0.16	(0.16)	-0.53*	(0.28)
Male LS	-0.28***	(0.10)	-0.37**	(0.16)
Irrigation	0.33***	(0.12)	0.39**	(0.19)
Gini	-0.75***	(0.29)	-1.20**	(0.47)
Rainfall	0.00	(0.01)	0.01	(0.01)
Paved roads	0.35***	(0.13)	0.13	(0.26)
Electrified	-0.59***	(0.21)	-0.50*	(0.30)
Commercial bank	-0.04	(0.16)	-0.01	(0.23)
Primary-Middle female	0.04	(0.27)	-0.15	(0.55)
Secondary female	0.38	(0.35)	0.34	(0.66)
Primary-Middle male	-0.29	(0.27)	-0.21	(0.42)
Secondary male	-0.16	(0.25)	0.11	(0.48)
Urban percent	-0.11	(0.08)	-0.09	(0.17)
BMI (Female)	-0.00	(0.01)	-0.01	(0.02)
BMI (Male)	-0.01	(0.01)	0.01	(0.02)
Constant	5.73***	(0.60)	4.91***	(0.87)
AEZ		Yes		
Land allocation to crops		Yes		
Observations	279		279	
Under-id (p-val)	0.01		0.01	
F(excluded instruments) L^S_F	3.957		3.957	
F(excluded instruments) L^S_M	17.25		17.25	

Source: Fertilizer- Fertilizer Association of India 2004-05; Implements- Livestock Census 2003; Body Mass Index- National Family Health Survey 2005-06

Note: Two stage least squares estimates, instrumenting for labor supply of males and females using low caste and industry employment as defined in table 2.5. Log of wages and log of labor supply are used in the above regressions. Robust clustered standard errors are in parenthesis; ***, ** and * indicate significance at the 1, 5 and 10% levels respectively. The unit of analysis is a district and districts having at least 5 wage observations for male and female each are included here.

Table 2.8 Aggregate Demand for Total Labor in Agriculture with All Observations

	Male wage		Female wage	
Female LS	-0.05	(0.06)	-0.53**	(0.24)
Male LS	-0.36***	(0.13)	-0.34**	(0.16)
Irrigation	0.22**	(0.10)	0.42**	(0.19)
Gini	-0.46**	(0.20)	-1.32**	(0.53)
Rainfall	-0.01	(0.01)	0.01	(0.01)
Paved roads	0.40***	(0.12)	0.09	(0.22)
Electrified	-0.60***	(0.20)	-0.47*	(0.24)
Commercial bank	0.06	(0.22)	-0.03	(0.22)
Primary-Middle female	0.08	(0.22)	-0.24	(0.51)
Secondary female	0.20	(0.30)	0.29	(0.64)
Primary-Middle male	-0.21	(0.20)	-0.16	(0.37)
Secondary male	0.11	(0.26)	0.14	(0.42)
Urban percent	-0.16*	(0.09)	-0.01	(0.15)
Constant	5.09***	(0.50)	5.22***	(0.77)
AEZ		Yes		
Land allocation to crops		Yes		
Observations	359		288	
Under-id (p-val)	0.02		0.02	
F (excluded instruments) L^S_F	8.76		5.54	
F (excluded instruments) L^S_M	6.69		17.03	

Note: Two stage least squares estimates, instrumenting for labor supply of males and females using low caste and industry employment as defined in table 2.5. Log of wages and log of labor supply are used in the above regressions. Robust clustered standard errors are in parenthesis; ***, ** and * indicate significance at the 1, 5 and 10% levels respectively. The unit of analysis is a district and districts having at least 5 wage observations for male and female separately are included here for estimating male and female demand equations respectively.

Table 2.9 Two-Stage Least Squares Estimates: Individual Data

Panel A: Impact of Female and Male Labor Supply on Female and Male Wages

	Male wage		Female wage	
Female LS	-0.06	(0.23)	-0.55**	(0.28)
Male LS	-0.39***	(0.13)	-0.40*	(0.20)
Observations	7,812		6,378	
Under-id (p-val)	0.00		0.00	
F(excluded instruments) L^S_F	3.71		5.34	
F(excluded instruments) L^S_M	12.96		13.14	

Panel B: First-Stage Coefficients

	Male LS		Male LS	
Low caste	-0.24	(0.15)	-0.18	(0.16)
Industry	-2.84***	(0.56)	-2.61***	(0.55)
R-square	0.71		0.64	
	Female LS		Female LS	
Low caste	0.61***	(0.23)	0.60***	(0.18)
Industry	0.53	(1.18)	-0.54	(0.85)
R-square	0.57		0.52	

Note: Panel A reports the two stage least squares estimates, instrumenting for labor supply of males and females using low caste and industry employment as defined in table 2.5 and controlling for individual characteristics like age, age square, education dummies, marital status and agricultural task along with all district controls in the base specification (specification 3) of table 2.6. Log of individual wage for each task that a person undertakes is the dependent variable. Log of labor supply at district level is the measure of explanatory variable. Panel B reports the corresponding first-stage. Robust clustered standard errors are in parenthesis; ***, ** and * indicate significance at the 1, 5 and 10% levels respectively. The districts are restricted to the ones included in analysis in table 2.6.

Table 2.10 Aggregate Demand for Total Labor in Agriculture when Total Labor is Measured in Efficiency Units

θ=	0.5		0.7		0.9		1	
Male Wage								
Female LS	-0.12	(0.15)	-0.13	(0.15)	-0.13	(0.15)	-0.13	(0.15)
Male LS	-0.37***	(0.13)	-0.32***	(0.11)	-0.29***	(0.10)	-0.28***	(0.09)
Female Wage								
Female LS	-0.47*	(0.26)	-0.50**	(0.25)	-0.52**	(0.25)	-0.52**	(0.25)
Male LS	-0.58***	(0.22)	-0.47***	(0.18)	-0.40**	(0.16)	-0.37**	(0.15)

Note: Two stage least squares estimates, instrumenting for labor supply of males and females using caste and industry employment as defined in table 2.5 and controlling for all district controls in the base specification (specification 3) of table 2.6. Log of wages and log of labor supply are used in the above regressions. Robust clustered standard errors are in parenthesis; ***, ** and * indicate significance at the 1, 5 and 10% levels respectively. The unit of analysis is a district and districts having at least 5 wage observations for male and female each are included here.

Table 2.11 Summary Statistics of Variables in Northern and Southern States

Variable	Mean	Standard deviation	Mean	Standard deviation
	Northern states		Southern states	
Female LS	0.54	0.73	0.98	0.60
Male LS	1.70	0.61	1.19	0.53
Irrigation	0.52	0.27	0.34	0.22
Gini	0.66	0.10	0.71	0.09
Rainfall	9.21	4.73	7.12	6.11
Paved roads	0.53	0.23	0.83	0.13
Electrified	0.75	0.27	0.99	0.02
Commercial bank	0.06	0.03	0.14	0.17
Primary-Middle female	0.23	0.10	0.27	0.11
Secondary female	0.09	0.05	0.15	0.07
Primary-Middle male	0.36	0.09	0.36	0.10
Secondary male	0.21	0.09	0.25	0.08
Urban percent	0.23	0.18	0.32	0.18
Coarse Cereals	0.09	0.13	0.24	0.22
Cotton	0.08	0.12	0.09	0.11
Oilseeds and Pulses	0.22	0.20	0.30	0.19
Rice	0.39	0.28	0.25	0.25
Horticulture	0.03	0.03	0.10	0.17
Male wage	3.77	0.25	3.88	0.30
Female wage	3.63	0.29	3.43	0.29

Note: Weighted mean with weights equal to district population. Andhra Pradesh, Karnataka, Kerala, Maharashtra and Tamil Nadu are classified as the Southern states while Assam, Bihar, Gujarat, Punjab, Haryana, Rajasthan, Uttar Pradesh, Madhya Pradesh and West Bengal are classified as the Northern states.

Table 2.12 Explained Difference in Gender Wage Gap between Northern and Southern States

Variable	Proportion wage gap explained
Female LS	55%
Paved roads	36%
Rice	29%
Horticulture	10%
Gini	10%
Rainfall	7%
Irrigation	5%
Primary-Middle female	2%
Commercial bank	1%
Secondary female	0%
Primary-Middle male	0%
Cotton	-2%
Urban percent	-2%
Oilseeds and Pulses	-2%
Secondary male	-2%
Electrified	-13%
Male LS	-14%
Coarse Cereals	-22%

Appendix for Chapter 2

2.A First-stage Results for Identification of Labor Demand Equations

Table 2.A.1 First-stage Results for Labor Supply by Males and Females to Agriculture

	Male labor supply		Female labor supply	
Low caste	-0.22	(0.19)	0.79***	(0.27)
Industry	-3.33***	(0.59)	-0.26	(0.97)
Irrigation	0.37*	(0.18)	0.49**	(0.24)
Gini	-0.37	(0.34)	-1.78***	(0.46)
Rainfall	0.01	(0.01)	0.01	(0.02)
Paved roads	0.04	(0.24)	-0.29	(0.38)
Electrified	-0.45*	(0.25)	-0.33	(0.3)
Commercial bank	-0.08	(0.16)	-0.39	(0.4)
Urban percent	-0.22	(0.17)	-0.07	(0.27)
Primary-Middle female	-0.13	(0.58)	0.40	(0.9)
Secondary female	0.41	(0.77)	0.33	(0.88)
Primary-Middle male	-0.27	(0.44)	-0.96	(0.78)
Secondary male	-0.43	(0.42)	0.38	(0.84)
Constant	1.83***	(0.58)	1.14	(0.74)
AEZ		Yes		
Crop Composition		Yes		
R-Square	0.71		0.54	
Under-id (p-val)	0.01		0.01	
Observations	279		279	

Note: The table shows the first-stage estimates for specification 3 in table 2.6. Log of labor supply is the dependent variable. Robust clustered standard errors are in parenthesis; ***, ** and * indicate significance at the 1, 5 and 10% levels respectively. The unit of analysis is a district and districts having at least 5 wage observations for male and female each are included here.

Chapter 3

Rainfall Shocks and the Gender Wage Gap: Evidence from Indian Agriculture[1]

3.1 Introduction

Labor economists have long been concerned with measurement of wage gaps across demographic groups like gender, race and ethnicity. The immediate question of interest which has received much attention is why do these gaps exist? These gaps have been shown to exist due to differences in characteristics, differential returns to similar characteristics, discrimination and prevailing labor demand and supply conditions for males and females. The second question of interest which has received comparatively less attention is how these wage gaps respond to changes in labor market conditions.

Very few studies examine the changing nature of wage gaps between demographic groups due to a tightening or a boom in the labor market and these too are restricted to the developed countries' urban labor markets. O'Neill (1985) and more recently Biddle and Hamermesh (2011) analyze the impact of aggregate changes in unemployment rates in the

[1] I would like to acknowledge a research grant from United Nations Development Program towards the research work in this chapter.

economy on gender gap in wages using data from the United States. O'Neill (1985) finds that gender wage gap increases when high unemployment rates prevail in the economy and attributes it to possibly greater vulnerability of females to layoffs due to lesser training and in general a weaker labor market for females during recessions. Biddle and Hamermesh (2011) also find that gender wage gaps are counter cyclical in nature. They attribute it to greater discrimination against females during downturns.

Contrary evidence to the above is presented in Solon, Barsky and Parker (1994), Park and Shin (2005) and Marchand and Olfert (2013). Using other datasets and time periods for the United States they estimate that gender wage gaps are pro-cyclical in nature. This they argue is due to males being over represented in industries with more pro-cyclical demand. Aller and Arce (2001) also find that gender wage gap reduced during the 1990-94 economic recession in Spain due to a proportional increase in service employment. In the Indian context there is scanty literature on the effects of demand shocks on labor market outcomes. The only related study that the author is aware of is by Bhalotra (2010) who finds that decrease in state level income (a proxy for recessionary conditions in labor market) increases women's labor force participation in the agriculture sector.

As noted in many countries, there is a persistent gap in daily wage rate of females and males in Indian agriculture as well. Table 3.1 shows the female to male wage ratio across 14 Indian states from the National Sample Survey (NSS) rounds of 1993-2007. On an average female daily wage rate in agriculture is 70% of that of male agricultural labor. However, real agricultural wages for males and females are not static over time and change in response to prevailing conditions in the labor market. One of the shocks that the agricultural labor market is continually subject to is rainfall variability across years. These shocks affect profitability and change demand for labor in Indian agriculture. This variability in demand for labor in agriculture due to natural rainfall shocks can have an impact on gender wage gap in agriculture. The relative impact depends on asymmetric effects of rainfall variability on demand for male and female labor in agriculture and asymmetric supply responses of male and female labor to these shocks.

In the past few decades, climate change has made rainfall more variable in India leading to recurrence of drought like situations across different parts of the country in different years. Each year some part of the country suffers from rainfall scarcity leading to low agricultural output and affects livelihood of people dependent on agriculture in rural India. Of the twenty two countrywide major droughts witnessed in India in the last 120 years, seven have occurred in the last three decades. Since agriculture plays a pivotal role in employment generation in developing countries (according to the Ministry of Home Affairs (2011), 73 percent of rural workforce is engaged in agriculture), any disruption in the agricultural labor market has an impact on a large section of working population. It remains the largest sector absorbing the Indian workforce and agricultural laborers constitute 55 percent of the agricultural workforce (Ministry of Home Affairs 2011). In terms of gender composition, about 77 percent of female work force in rural India is engaged in agriculture. Climatic shocks may not have gender neutral effects (WEDO 2008) and could also potentially affect the agricultural wages of males and females differently.

This is the first study which looks into the impact of adverse demand shocks in rural labor markets on gender gap in agricultural wages in a developing country context. In this chapter, we estimate the impact of changing demand conditions in rural agricultural labor market due to rainfall shocks on gender wage gap in agriculture. We construct a district level panel dataset to examine the above question. Also, we specifically look into the heterogeneity in the impact of these shocks in rice growing regions of the country since rice is considered to be a crop in which there is greater demand for female labor (Boserup 1970) and it is also a crop highly susceptible to rainfall shocks under rainfed conditions due to large water requirement.

In the next section we provide a literature review of studies estimating the impact of weather variability on labor market outcomes. Section 3.3 outlines a simple theoretical framework for the chapter and section 3.4 discusses the data and the variables constructed. In section 3.5 we estimate the impact of rainfall shocks on agricultural yields and in section 3.6 discuss the empirical strategy and the main findings. Section 3.7 provides a few robustness checks for the main findings and in section 3.8 we discuss the mechanisms which could possibly lead to the results obtained. In section 3.9 we gather concluding remarks.

3.2 Effect of Weather Shocks on Labor Market Outcomes

The literature on weather shocks and labor market outcomes can be broadly divided into two strands. One strand of work looks into the effect of rainfall shocks on wages (Jayachandran 2006; Mueller and Osgood 2009; Mueller and Quisumbing 2011; Burgess et al. 2011; Adhvaryu, Char and Sharma 2013). The second strand of work looks into the effect of these shocks on time use in different activities by agricultural households as a means to cope with the shock (Rose 2001; Maitra 2001; Badiani and Safir 2008; Ito and Kurosaki 2009). This chapter contributes to the first body of work by looking at the gender differentiated impact of rainfall shocks on wages. There is no study which looks into the heterogeneity of wage impacts by gender for aggregate shocks like rainfall variability.[2]

[2] Halliday (2012) is the only paper which uses an aggregate natural shock of earthquakes in El Salvador to study their impact on male and female time use on own farm-agricultural field work, livestock labor, domestic labor and off-farm labor. The earthquake increased domestic work and decreased livestock production by women. He

We briefly discuss the studies in this literature. Jayachandran (2006) measures the impact of productivity shocks on district level daily agricultural wage for men in India and concludes that a negative productivity shock lowers the real wage significantly. This effect is less pronounced in districts which are more developed in terms of access to roads and banks. Mueller and Quisumbing (2011) study the impact of 1998 Bangladesh flood on agricultural and non-agricultural wages. They find a short term reduction in real wage and a smaller reduction in wage for agricultural workers who were able to find non-agricultural employment. Mueller and Osgood (2009) look at not only short term effects of droughts on wages but long term effects as well. They argue that the long term effect can be due to selling of productive assets during the shock to sustain consumption and it may take a long time to replenish them. They use data on both rural and urban areas in Brazil and find that adverse rainfall shocks upto five years in the past can have negative effects on individual wages. Adhvaryu et al. (2013) quantify the impact of rainfall shocks on total manufacturing sector employment and output in India and find that there is a fall in both when rainfall shock is negative, particularly more in pro-employer states and for small factories and agro-based industries in India. The literature on the heterogeneous effect of weather shocks on labor market outcomes by gender is thus largely non-existent. This chapter aims to fill this gap by examining the dimension of wages and the consequent effect of weather shocks on male and female wage gap in agriculture.

3.3 Theoretical Framework

Before proceeding with the empirical analysis it is useful to discuss the theoretical implications of productivity shocks like ‘rainfall’ on male and female wages. In the model

argues that this is compatible with a theoretical framework in which household members are allocated to sectors according to their comparative advantage when it is affected by an adverse income shock.

below, we assume a competitive agricultural labor market and three factors of production – land (A), male labor (l_M^D) and female labor (l_F^D) - and rainfall shock (r) which affects agricultural productivity. The production function is continuous and differentiable with diminishing returns to each factor and amount of land is fixed in the short run. The male and the female labor are assumed to be imperfect substitutes in production. Also, hired and own farm labor are assumed to be perfectly interchangeable in the production. The profit function is given by

$$\pi = F(A, l_M^D, l_F^D, r) - W_M l_M^D - W_F l_F^D$$

Where W_M and W_F are male and female wage rate respectively. In a competitive equilibrium all factors are paid their marginal products. The first order conditions for profit maximization are

$$W_M = F_{L_M}(A, l_M^D, l_F^D, r) \quad (3.1)$$

$$W_F = F_{L_F}(A, l_M^D, l_F^D, r) \quad (3.2)$$

If we assume labor supply to be constant, then the effect of rainfall shocks on equilibrium male and female wages will be

$$\frac{dW_M^*}{dr} = \frac{dF_{l_M^D}}{dr} \quad (3.3)$$

$$\frac{dW_F^*}{dr} = \frac{dF_{l_F^D}}{dr} \quad (3.4)$$

The above expressions show that rainfall shocks affect wages through their impact on the value of marginal productivity of labor. A positive shock increases value of marginal productivity for both males and females and hence increases male and female wages. The above expressions are thus unambiguously positive. In the above model female wage will be affected more than male wage if the impact of rainfall shocks on the value of marginal productivity of females is greater than that on the value of marginal productivity of males.

However, in general, labor supply is unlikely to be fixed. Relaxing this assumption, in a general equilibrium framework, consider two types of households- landless (N) and landed (H). Landed households supply farm labor and also hire in labor on their farms, while landless households only hire out labor to landed households. Both the male and the female labor are supplied by the two types of households. Male and female labor are imperfect substitutes in production but labor of each gender from different household types are perfectly substitutable. Male and female leisure are also assumed to be imperfect substitutes in household utility function. All households are assumed to be price takers but wage rates are determined endogenously by the market clearing equilibrium conditions. This framework assumes separability in agricultural household production and consumption decisions which holds when markets are perfect (Singh, Squire and Strauss 1986).

The two types of households maximize an identical, twice differentiable utility function with respect to a consumption commodity (x), leisure of male members (l_M) and leisure of female members (l_F) subject to the budget constraint. The total time available with each member of the household is one unit. Consider a household with one male and one female member. The budget constraint facing each household is given by

$$x^i = W_M(l_M^{S,i}) + W_F(l_F^{S,i}) + v^i \qquad \text{when } i = \text{N (Landless)}$$

$$x^i = W_M\left(l_M^{S,i}\right) + W_F\left(l_F^{S,i}\right) + v^i + \pi \qquad \text{when } i = \text{H (Landed)}$$

Where $l_M^{S,i}$ is male labor supplied by the household, $l_F^{S,i}$ is female labor supplied by the household, v is non-labor income of the household and π is profits equal to $F(A, l_M^D, l_F^D, r) - W_M l_M^D - W_F l_F^D$. Since we assume that each member of the household has one unit of time available, we can further substitute for labor in the above budget constraints using the relation $l_M^{S,i} = 1 - l_M^i$. Re-writing the budget constraints in terms of leisure, the household maximization problem is given by

$$\text{Max } U(x^i, l_M^i, l_F^i)$$

s.t.

$x^i + W_M l_M^i + W_F l_F^i = W_M + W_F + v^i$ when $i =$ N (Landless)

$x^i + W_M l_M^i + W_F l_F^i = W_M + W_F + v^i + \pi$ when $i =$ H (Landed)

The first order conditions with interior solutions are

x: $U_x^i = \lambda^i$

l_M: $U_{l_M}^i = \lambda^i W_M$

l_F: $U_{l_F}^i = \lambda^i W_F$

λ^i: $x^i + W_M l_M^i + W_F l_F^i = W_M + W_F + v^i$ when $i =$ N (Landless)

$x^i + W_M l_M^i + W_F l_F^i = W_M + W_F + v^i + \pi$ when $i =$ H (Landed)

For the profit maximizing landed households, the first order condition for maximization will also include (3.1) and (3.2). In equilibrium total labor demand will be equal to total labor supply for both males and females

$$L_M^D(W_M, W_F, r, A) = L_M^{S,N}(W_M, W_F) + L_M^{S,H}(W_M, W_F, r, A) \quad (3.5)$$

$$L_F^D(W_M, W_F, r, A) = L_F^{S,N}(W_M, W_F) + L_F^{S,H}(W_M, W_F, r, A) \quad (3.6)$$

Here, L_M^D (L_F^D) is the aggregate demand for male (female) labor and and L_M^S (L_F^S) is the aggregate supply of male (female) labor. Totally differentiating the above market equilibrium conditions and using Cramer's Rule to get expressions for change in equilibrium wage due to the rainfall shock ($\frac{dW_M^*}{dr}$ and $\frac{dW_F^*}{dr}$)

$$\frac{dW_M^*}{dr} = \frac{\left[\varepsilon_{F,W_F}{}^*\left(L_{M,r}^{S,H} - L_{M,r}^D\right) - \varepsilon_{M,W_F}{}^*\left(L_{F,r}^{S,H} - L_{F,r}^D\right)\right]}{\varphi} \quad (3.7)$$

$$\frac{dW_F^*}{dr} = \frac{\left[\varepsilon_{M,W_M}{}^*\left(L_{F,r}^{S,H} - L_{F,r}^D\right) - \varepsilon_{F,W_M}{}^*\left(L_{M,r}^{S,H} - L_{M,r}^D\right)\right]}{\varphi} \quad (3.8)$$

where, ε_F and ε_M refer to excess demand for females and males respectively and $\varepsilon_{F,W_F}$is the differential of excess demand for females with respect to female wage. Similar interpretations hold for ε_{F,W_M}, ε_{M,W_M} and ε_{M,W_F}. $L_{M,r}^{S,H}$ and $L_{F,r}^{S,H}$ is the differential of labor supply of males and females respectively in the landed households to change in rainfall conditions. This will be negative for both males and females, as a positive productivity shock increases profit income thus reducing labor supply by landed households. $L_{M,r}^D$ and $L_{F,r}^D$ are change in male and female labor demand respectively due to the rainfall shock. These are positive since by definition demand for both types of labor increases with higher rainfall. φ has to be positive for the multimarket Hicksian stability condition to hold.[3] Given the assumptions on production and utility functions $\frac{dW_M^*}{dr}$ and $\frac{dW_F^*}{dr}$ will be unambiguously positive. However, it is not possible to arrive at the effect of rainfall shocks on relative female wage $\left(\frac{dW_F^*}{dr} - \frac{dW_M^*}{dr}\right)$

[3] The expression for φ is $\left(\varepsilon_{M,W_M} * \varepsilon_{F,W_F}\right) - \left(\varepsilon_{F,W_M} * \varepsilon_{M,W_F}\right)$.

since it will depend on relative magnitude of the impact of rainfall shock on labor demand for males and females and labor supply of males and females to agriculture as well as the existing labor demand and supply elasticities. Intuitively, the expression in 3.7 shows that the effect of rainfall shock on male wage in equilibrium depends on the effect of the rainfall shock on shifts in demand for and supply of male labor, and the effect of the rainfall shock on the shifts in demand for and supply of female labor. The effect of a given magnitude of this shift on equilibrium wages depends on labor demand and supply responsiveness (both own and cross) of male and female labor.

3.4 Data and Variable Construction

The primary dataset used in this chapter are the Employment and Unemployment rounds (1993-94, 1999-00, 2004-05, 2007-08) of NSS in India. NSS is a cross-sectional dataset which is representative of India's population. The survey contains labor force participation and earnings details for the reference period of a week and follows a two stage sampling design. In rural areas, the first stratum is a district. Villages are primary sampling units (PSU) and are picked randomly in a district over an entire agricultural year (July to June) over quarters to ensure equal spacing of observations across the year. The households are randomly chosen in the selected PSU's. The district level analysis includes 14 major states in the sample: Punjab, Haryana, Uttar Pradesh (includes Uttarakhand), Madhya Pradesh (includes Chattisgarh), Bihar (includes Jharkhand), Gujarat, Rajasthan, West Bengal, Maharashtra, Andhra Pradesh, Karnataka, Orissa, Tamil Nadu and Kerala. Average agricultural wage in an agricultural year is estimated by calculating an average of daily wage rate (total earnings divided by total days worked in the last week) for all casual wage laborers in agriculture in a district, weighted by sampling weights provided in the survey so that the district level average wage rate is representative of the district population. Wages are adjusted for changes in price levels over-time by using the consumer price index for agricultural laborers as the deflator. Districts with both male and female wage observations are included in the analysis.

Data on area and yield of crops at district level is obtained from Area, Production and Yield statistics (1999-2008) published by Ministry of Agriculture, India. Rainfall data used in the chapter comes from the gridded dataset of the Center of Climatic Research at the University of Delaware, which includes monthly precipitation values on 0.5 degree intervals in longitude and latitude centered on 0.25 degree. This grid value is achieved by spatial interpolation using data from nearby weather stations and other sources of rainfall data.[4] District level monthly rainfall estimates were arrived at by averaging the monthly precipitation value of all the grid points lying within the geographic boundaries of a district in a year.[5] The geographic boundaries refer to the boundaries of the parent districts as Indian districts have been split into two or more districts over time. Districts across NSS rounds (1993-94, 1999-00, 2004-05, and 2007-08) have hence been merged into their parent districts according to district boundaries in the 1991 census.[6]

About 75% of rainfall in India is received during the monsoon season from June to September.[7] The monsoon precipitation levels are very critical for agricultural yields during

[4] For further information on the dataset and the methodology used for interpolation please refer to the below link http://climate.geog.udel.edu/~climate/html_pages/Global2_Ts_2009/README.global_p_ts_2009.html

[5] Hilly regions of Jammu and Kashmir, Himachal Pradesh and North-East are excluded from the analysis as measurement of rainfall in hilly areas using gridded dataset induces a lot of measurement error since precipitation varies considerably by altitude.

[6] To match the districts into their parent districts two sources of information were used which track the evolution of Indian districts over time. Kumar and Somanathan (2009) document changes in district boundaries over the census years and http://www.statoids.com/yin.html which compiles changes in Indian district boundaries from 1982 to the present.

[7] Typically, March to May are summer months and the winter spans over December to February in India: Indian Meterological Department http://www.imd.gov.in/section/nhac/dynamic/FAQ_monsoon.htm. Also, a report by

the agricultural year not only for the kharif crop (June to October) but also for the rabi crop (November to May) since they help recharge the aquifers and also replenish the moisture content in the soil. The methodology used for constructing rainfall shocks is similar to Jayachandran (2006). She argues that in the Indian context above normal level rainfall improves agricultural productivity while below normal level rainfall reduces it and this has an effect on agricultural wages.[8] Thus, excess rain is treated as a good shock while a shortfall is taken to be a bad shock.

The "RainShock" variable equals one if the monsoon rainfall is above the eightieth percentile for a district, zero if it is between the eightieth and the twentieth percentile and minus one if it is below the twentieth percentile. Rainfall data from 1971-2008 is used to construct the shocks. Using this definition, 67% districts experienced a drought and 50% experienced a good rainfall in at least one year in our dataset. Similar to her finding, when regressing wage is regressed on dummy variables for rainfall above the eightieth percentile and for rainfall below the twentieth percentile separately, I cannot reject that they have equal and opposite magnitude.[9] This restriction helps to improve power in the regressions. The definitions and summary statistics for other variables are presented in table 3.2.

IMD on rainfall pattern in India can be accessed at
http://www.imdpune.gov.in/ncc_rept/RESEARCH%20REPORT%202.pdf

[8] Table 3.B.1 shows the effect of rainfall shocks when the rainfall shock is defined as the standardized value of rainfall in a district. Standardized rainfall is calculated by subtracting the long term average rainfall in a district (1970-2008) from the actual rainfall in a given year and dividing by the long term standard deviation of rainfall in that district. The results show that in our data as well, the effect of rainfall shocks is not linear. Only extreme deviations from the normal rainfall have an effect on agricultural wages.

[9] Table 3.B.2 shows the results from this specification.

3.5 Rainfall Variability and Its Effect on Agricultural Yield in India

Weather varies across regions in India, but broadly India experiences four seasons- winter (January to February), summer (March to May), monsoon (June to September) and the post-monsoon season (October to December) (Ribot, Najam and Watson 1996). Across India, regions greatly differ in their precipitation levels. It ranges from scanty to heavy in different parts of the country. The rainfall pattern not only differs spatially but temporally as well. Alternating sequence of multi-decadal periods of thirty years having frequent droughts are observed in the all India monsoon rainfall data. There has been no overall trend in rainfall observed in India, but the frequency and the intensity of rainfall appears to be changing (Kumar 2009). Agricultural output in India depends on monsoon as nearly 55.7 per cent of area sown is dependent on rainfall.[10] Sources of irrigation like canal and groundwater also get depleted during drought years.

More than 80% of the production and area sown in India is under food-grains. Rice and wheat are the most important food-grains cultivated in India. Figure 3.1 plots the coefficients from regressing log of yield of rice and wheat at district level on the deciles of rainfall with the first decile as the omitted category, district and time fixed effects. A concern with the rainfall shock measure constructed is that higher levels of rainfall may result in flood like situations, which could destroy crops, thus reducing yields. The plot of coefficients shows that this is not the case. For both rice and wheat we do not observe any yield dips in

[10] National Portal of India, website http://india.gov.in/sectors/agriculture/index.php?id=2

the ninth and the tenth deciles of rainfall which would be the case if floods were affecting these crops adversely. All the coefficients in the regressions are significant at 1% level.[11]

Table 3.3 shows the effect of the rainfall shock variable constructed on the yields for five crops- rice, wheat, bajra, gram and maize. For all the crops, the shock has a significantly positive effect on yields. A positive shock results in approximately 8% higher yields than in a normal rainfall year for rice, bajra and gram. While for wheat this effect is 5.7% and is further lower for maize at 1.8%. The magnitude is highest for rice at 8.7%.

Rice crop requires a lot of water and it takes 3,000–5,000 liters to produce one kilogram of rice, which is about two to three times more than to produce one kilogram of other cereals (Singh, Choudhury and Bouman 2002). Rice is best grown under flooded conditions. Though rice can be grown in both dry or semi-dry conditions and wet conditions, the cultivation practices are very different. In dry areas, the soil is ploughed in summer and the seed is sown by broadcasting or by sowing the seed behind the plough. In wet cultivation, transplanting in puddled fields is adopted. Wetland conditions are created in semi-dry areas by impounding rain water.

Climatically, areas which receive high average rainfall and have tropical climate are most suitable for rice cultivation. Among the Indian states in the analysis, Kerala, West Bengal, Orissa and Bihar receive the highest monsoon rainfall. Table 3.4 shows monsoon rainfall levels and area under rice cultivation in the fourteen states of India included in the analysis. It shows that the states having at least fifty percent area under rice cultivation are also the states which receive the highest levels of rainfall in the country. These states are Kerala, located in the western coastal plain and West Bengal, Orissa and Bihar located in the

---

[11] Another concern with the constructed rainfall shocks may be that they do not sufficiently explain variability in yields since timing of rainfall would also be important. ANOVA of district level standardized yields on rainfall deciles shows that rainfall shocks explain 10 percent variation in year to year yield variation across districts of India. This is statistically significant at 1 percent level.

Eastern Gangetic plains, which receive higher amounts of rainfall (more than 1000 mm during the monsoon season). The climatic conditions in these states are thus highly suitable for wet cultivation of rice.

Gender roles also differ across crops (Boserup 1970). It has been well documented by researchers that cultivation of rice involves greater labor-days. Women labor, particularly, is in greater demand for tasks like transplanting (Aggarwal 1986). Wet cultivation of rice requires rice to be grown in flooded fields. Since rice seeds cannot germinate under flooded conditions they need to be grown in nurseries and then transplanted manually in the flooded rice fields. Women are considered more efficient than males in transplanting since it requires long periods of back-bending work and delicate hands and women seem to have a comparative advantage in these skills. Bardhan (1974) noted that female labor demand is more in rice growing areas

> "Transplantation of paddy is an exclusively female job in the paddy areas; besides, female labor plays a very important role in weeding, harvesting and threshing of paddy. By contrast, in dry cultivation and even in wheat cultivation, under irrigation, the work involves more muscle power and less of tedious, often back-breaking, but delicate, operations." Bardhan (1974, 1304)

There is also a distinct sequential nature of tasks by gender in rice production. Men are involved in land preparation, sowing and seed-bed preparation. The next stage involves transplanting of rice by women. Thereafter, men are only involved with irrigation and application of fertilizer while women undertake most manual weeding operations, harvesting and post-harvesting operations including threshing of crops (IRRI 2000). Tasks like transplanting and weeding in which females supposedly have comparative advantage account for one-third of total labor percentage used in rice farming (Mbiti 2007). Using this observation he finds that agricultural households in rice growing areas do not marry their daughters in high rainfall years and attributes this to greater relative value of female labor in these years. If this were indeed true then the positive effect of rainfall shock on equilibrium wage must be greater for females in rice growing areas as compared to male wage, assuming that labor supply responses by males and females are not able to completely offset the changing demand for female labor.

This discussion motivates a construction of an indicator for rainfed rice cultivation intensity since we expect the heterogeneous wage impact by gender to be the greatest if rice is grown under rainfed conditions. If a state has at least 50% area under cultivation of rice then districts in that state are classified as rainfed rice growing areas.[12] We use a state level indicator of rice cultivation since rice can be grown in select districts within a state having irrigated agricultural systems. In southern states of India and in the state of Punjab, due to lower rainfall, rice cultivation is only done in a few districts under well irrigated conditions. Thus, around 90% of area under rice is irrigated in these states (Sen 1985).[13] In such areas rice cultivation may not occur due to better climate suitability but rather due to better irrigation facilities, which also mitigates the effect of rainfall shocks on rice yields. Hence, we use a state level indicator for rice cultivation, since we expect the heterogeneous gender impact of rainfall shocks to be the greatest in rainfed rice systems. To check the robustness of

[12] This indicator has been arrived at by averaging the area under rice and other crops (excluding horticulture) over a span of six years, from 1999-2004, for which data for all states is available. This proxy is used since district level dataset on rainfed rice cultivation is not collected for al l districts of India.

[13] More recent data from 1999-2004, on crop level irrigated area (Directorate of Economics and Statistics, Ministry of Agriculture, India) in states of India also shows that more than 90 percent under rice cultivation in Punjab and southern states of India is irrigated.

the results I use alternative definitions of rainfed rice cultivation in a region by using district level area under rainfed rice cultivation.

3.6 Estimation Strategy and Findings

To examine the effect of rainfall shocks on male and female wage gap in agriculture, I create a district level panel dataset using the NSS rounds of 1993-94, 1999-00, 2004-05 and 2007-08. Equation $(3E.1)$ below estimates the effect of rainfall shocks on male and female wages and equation $(3E.2)$ estimates the effect on relative female wage in agriculture.

$$W_{dt}^{k} = \beta_0 + \beta_1 RainShock_{dt} + \beta_2 X_{dt} + \beta_3 D_d + \beta_4 T_t + \epsilon_{dt} \quad (3E.1)$$

$$W_{dt}^{F} - W_{dt}^{M} = \gamma_0 + \gamma_1 RainShock_{dt} + \gamma_2 X_{dt} + \gamma_3 D_d + \gamma_4 T_t + \epsilon_{dt} \quad (3E.2)$$

Here, $k = M, F$ indexes male and female respectively and W_{dt}^{k} is log of average wage in cultivation in district 'd' in year 't'. '*RainShock*' is the rainfall shock experienced in district 'd' in year 't'. X_{dt} are time varying district characteristics. D_d are district fixed effects that control for time invariant characteristics of the district such as agro-ecological conditions, average crop composition, culture, norms, labor force characteristics and the initial level of development. T_t is a vector of time dummies. The identification of the parameter of interest is based on over time variation in wages and rainfall in a district. The parameter of interest γ_1 gives the impact of rainfall shocks on relative female wage in a district. Given that realized rainfall shocks are random the regressions will give an unbiased estimate of γ_1.

A district is a smaller administrative unit within states in India and is taken as the unit of analysis. This is based on the assumption that districts constitute separate agriculture labor markets in the country. Previous studies by Jayachandran (2006) and Rosenzweig (1984) also make a similar assumption since districts can be considered as distinct labor markets due to low mobility of rural labor across districts in India. Following Jayachandran (2006), we allow for clustering of standard errors within a region-year since rainfall shocks are likely to be spatially correlated.[14] The district level regressions are estimated with analytical weights equal to district population since the dependent variable is an average estimated at district level.

Table 3.5 shows the estimates for the impact of rainfall shocks on wages for specification $(3E.1)$ and $(3E.2)$. Both male and female wages seem to be affected equally by the shocks. A positive shock increases female wage by 3.6% and male wage by 2.4% relative to a normal year. There is however, no significant impact of rainfall shocks on relative female wage.

As discussed earlier, there could be a differential impact of the rainfall shocks on relative female wage in rainfed rice cultivating regions. I further investigate the heterogeneity in the impact of rainfall shocks on relative female wage in rainfed rice cultivating areas, using the below specification.

$$W_{dt}^{k} = \beta_0 + \beta_1 RainShock_{dt} + \beta'_1 RainShock_{dt} * Rice_d + \beta_2 X_{dt} + \beta_3 D_d + \beta_4 T_t + \epsilon_{dt} \qquad (3E.3)$$

$$W_{dt}^{F} - W_{dt}^{M} = \gamma_0 + \gamma_1 RainShock_{dt} + \gamma'_1 RainShock_{dt} * Rice_d + \gamma_2 X_{dt} + \gamma_3 D_d + \gamma_4 T_t + \epsilon_{dt} \qquad (3E.4)$$

[14] The NSS groups districts within an Indian state into agro-climatically similar regions. On an average, a region comprises of five to ten districts.

Here, $Rice_d$ is the indicator variable for rainfed rice cultivation and γ'_1 gives the differential impact of rainfall shocks on relative female wage in rice cultivating areas relative to the others. Chin (2011) exploits a similar strategy of using the interaction of rice cultivation with rainfall shock, as an instrument for labor force participation of females. While she uses cross-sectional data, our estimation is based on panel data and that allows us to control for district fixed effects.

One possible concern may be that rice cultivation can be determined by local demand and supply conditions that also affect labor market outcomes. Since we arrive at the average percentage area under rice cultivation by taking an average of area under rice cultivation over six years, the rice indicator variable does not represent crop choice as a response to agricultural conditions prevailing in any given year. It reflects the average intensity of rice cultivation in a district. Any concerns about the endogeneity of rice indicator variable can then be allayed by inclusion of district fixed effects. Thus, in all the econometric specifications we include district fixed effects. Even so, we have an additional instrumental variable check in the next section.

Table 3.6 shows the results for specifications $(3E.3)$ and $(3E.4)$. Here, we find that gender wage gap increases with a negative rainfall shock (low rainfall) and decreases with a positive rainfall shock (high rainfall) in rice growing areas. This is reflected in a lower female to male wage ratio (relative female wage) during negative rainfall shocks and a higher female to male wage ratio during positive rainfall shocks in rice growing areas. A positive rainfall shock in rainfed rice growing areas increases the relative female wage by 4.8% relative to other areas. On the whole, a positive rainfall shock in rainfed rice growing areas increases the relative female wage by 4.7 percent (0.048-0.001).[15] This is driven by a lower impact of rainfall shocks on male wages in these rice growing areas.[16]

Table 3.7 augments the specifications $(3E.3)$ and $(3E.4)$ with additional control variables. To control for other variables which may affect wage responsiveness to shocks in a district we use percentage literate population in a district as an indicator of education and interact it with the rainfall shock. Generally, non-farm sector enterprises locate in areas with better infrastructure and education levels. Higher education levels, by improving access to non-farm jobs can insulate agriculture wages from rainfall shocks. The technological advancement in agriculture may also be greater in areas with more educated workforce. So we expect that in districts with larger literate population agricultural wages will be better protected against rainfall variations. The percentage of irrigated area in a district may lead to a differential effect of rainfall shocks on wages as well. To the extent that well irrigated areas experience smaller effects of rainfall shocks on yield of crops, the wages in irrigated areas tend to be insulated against rainfall variations. However, cultivation of riskier crops and greater dependence of local population on agriculture in well-irrigated areas can make local agricultural wages more responsive to rainfall shocks.[17] Morduch (1990) shows that Indian farm households which are more susceptible to income shocks are less likely to use riskier

[15] This is obtained by adding the coefficients from RainShock and the interaction of RainShock with Rice.

[16] Appendix 3.B shows the results where alternate definitions of rainfall shock are used. The last column of table 3.B.1 shows the heterogeneous effect of rainfall shocks in rainfed rice growing areas when the rainfall shock is defined as the standardized value of rainfall in a district. We still find that a one standard deviation increase in rainfall increases the female wage more than the male wage in rainfed rice growing areas of India relative to the other areas, the significance levels though fall. The last column of table 3.B.2 shows the heterogeneous effect of the rainfall shock when positive and negative shocks are considered separately. Again, the results show that a positive rainfall shock increases relative female wage by 5.9 percent in years of positive rainfall shocks and decrease it by 4.5 percent in years of negative rainfall shocks. These estimates are however insignificant. F-test shows that the null of these coefficients having equal and opposite effects cannot be rejected. Thus, we prefer the specification in table 3.6 where rainfall shock is constructed as a continuous variable which takes on values: -1, 0 and 1.

[17] In India among the irrigated districts in the period of study, 90 per cent have less than 75 per cent area under irrigation. If there are few irrigated pockets, workers could travel from villages within a district to these irrigated pockets in search of work, hence decreasing the wage rate.

seed varieties. Using ICRISAT data from three Indian villages, Wadood and Lamb (2006) also show that area under risky crop varieties increases with irrigation availability. In well irrigated areas high value but risky water-intensive crops and varieties are a major source of income. While rainfall shocks may not affect yield of cultivated crops, lower groundwater levels and less availability of canal water due to insufficient rainfall can result in less area planted under high value crops in a low rainfall year, affecting the value of marginal labor input.[18]

Table 3.7 shows the results conditional on the above variables. Both male and female wages are less responsive to rainfall shocks in areas with more literate population since the interaction term is negative but there is no impact on gender wage gap.[19] On the other hand, both male and female wages are more responsive to rainfall shocks in irrigated areas since the interaction term is significantly positive. Thus, wages fall more in response to negative rainfall shocks in irrigated areas. As discussed earlier this could be due to the practice of high risk cultivation strategies or a greater labor supply variation in the irrigated pockets within a district. The addition of the above controls does not, however, change the impact of rainfall shocks in rice growing areas on the gender wage gap. On the whole, a positive rainfall shock in rainfed rice growing areas increases the relative female wage by 10 percent (0.052+0.051). The specification with education and irrigation controls is used as the baseline and we add further controls to check the robustness of the above results.

---

[18] The accessed report documents crop composition shifts to less water intensive crops like gram and oilseed advised by the government in irrigated areas during years of drought especially in the rabi season when the alternative seeds could be made available by the government to the farmers. Sometimes the irrigation water is diverted for drinking purposes as well (http://www.empowerpoor.org/downloads/drought1.pdf). Though, overall irrigation will improve welfare of households since even with a greater percentage fall, the average wage in irrigated areas will be greater than in un-irrigated areas in drought years.

[19] Specification with separate female and male district level education was also estimated, but due to high collinearity between the two only male education was significant in male and female wage equations when both were added. However, when entered alone female education was also significant. No effect on gender wage gap was observed for any education variable. The specification with average education was hence preferred. The other results did not change though.

Jayachandran (2006), using data from 1960 to 1987, finds that the impact of rainfall shocks on wages is lower in more developed areas. Our period of analyses is more recent. To check for the robustness of our results, we control for a host of other developmental indicators used in her paper like accessibility of villages in a district by bus, road, rail, closeness of villages to a town and banking activity in a district. The data source for these variables is Census of India. Indicators of poverty like level of per capita expenditure and percentage landless households in a district are also entered as controls. We examine the robustness of our results to inclusion of the above development indicators.

Table 3.8, panel A, shows the results for indicators of accessibility like bus, road and rail. Panel B of table 3.8 shows the results controlling for banking activity and mean distance to town with the last column incorporating all the indicators. Panel C of table 3.8, controls for other indicators of poverty like percentage landless households and monthly per capita consumption of households in a district in 1993, along with the previous infrastructure indicators.[20] Adding controls for these indicators of development and poverty does not change our earlier result that in rice growing areas the fall in female wage is greater relative to male wage in low rainfall years. Also, in our estimation results, none of the developmental indicators or their interaction with the rainfall shock has a significant effect on either absolute wages or relative female wage.

3.7 Robustness Checks

In the previous section we argued that the interaction of rice cultivation with rainfall shock is exogenous since we have controlled for district fixed effects. To further allay any

[20] The coefficients of the added control variables in the third specification of Panel B and in the specifications presented in Panel C are provided in the appendix, table 3.B.3.

endogeneity concerns I instrument for the rice growing states indicator using the average state level rainfall from 1971-2008. The interaction of '$RainShock$' with mean monsoon rainfall is used as an instrument for the interaction of '$RainShock$' with '$Rice$'. Long term mean rainfall is unlikely to be correlated with labor market outcomes once district fixed effects are controlled for since it is an exogenously determined measure of climatic endowment. Also, it is likely to be highly correlated with average rainfed rice cultivation in a region since rice is a water intensive crop (table 3.4).

Table 3.9 shows the results of the two-stage least squares estimates along with the F-statistic of the first stage. The F-stat shows that the instrument is significant at 1% level of significance. The first stage results show that the interaction of rainfall shock and mean rainfall level is a significantly positive predictor of the interaction of rainfall shock and rice cultivation, which would be the case if mean rainfall level acts as a determining factor behind the choice of farmers to cultivate rice in a particular state. The second stage results confirm the findings of the previous reduced form regressions that the gender wage gap reduces in rice growing regions in years of high rainfall and increases in years of low rainfall. The magnitude of the impact is only slightly higher than the reduced form regression indicating that endogeneity is not a major concern.

As a second robustness check, I define the '$Rice$' variable differently by constructing a district level indicator of rainfed rice cultivation. The percentage area under rainfed rice cultivation in a district is calculated by multiplying the district level percentage area under rice cultivation with state level percentage area of rice crop which is rainfed.[21] We expect this to be a good indicator at district level due to uniformity within the states in rainfed rice

[21] The district level percentage area under rice cultivation has been arrived at by averaging the area under rice and other crops (excluding horticulture) over a span of six years, from 1999-2004, for which data for all states is available. The district level crop-wise irrigated area is not reported for all states and hence to overcome this data constraint I use state level crop wise irrigated area(1999-2004) which is available for all states from Directorate of Economics and Statistics, Ministry of Agriculture, India.

cultivation. In states like Punjab, Haryana, Andhra Pradesh and Tamil Nadu, where rice is intensively cultivated in a few districts, the entire area cultivated under rice is irrigated. In other states, there is very little area where irrigated rice is cultivated. Rice can be grown under both rainfed and irrigated conditions and it is crucial to make this distinction since in rainfed areas rainfall shocks are more likely to affect the demand for labor in rice cultivation. Table 3.10 shows the results with this alternative district level indicator of rice cultivation. The results support the previous findings that a positive rainfall shock decreases the gender wage gap (increases relative female wage) while a negative rainfall shock increases the gender wage gap (decreases relative female wage) as the percentage area under rainfed rice cultivation in a district increases. This is driven by a lower effect of rainfall shocks on male wages in rainfed rice cultivating regions. The above result is not affected when district level development indicators in table 3.8 are also controlled for. The results have not been shown for brevity.

In the third robustness check, I add state time fixed effects to the above specification in order to control for any state-year specific factors in wages which could have been omitted. The results after controlling for state specific time fixed effects are shown in the last column of table 3.10. As can be seen, our previous result is robust to the inclusion of these trends as well.

As the last robustness check, I conduct the analysis at individual level. The district level estimates do not take into account the differences in labor force characteristics of males and females in agriculture. To the extent that differences in characteristics are constant over time within a district, district fixed effects take into account the effect of these differential characteristics which are not changing over time, on male-female wage gap. To see if our results are robust to inclusion of individual characteristics I estimate the following equation.

$$W_{idt} = \alpha_0 + \alpha_1 RainShock_{dt} + \alpha_2 Female_{idt} + \alpha_3 RainShock_{dt} * Female_{idt} + \alpha_4 C_{idt} + \alpha_5 D_d + \alpha_6 T_t + \varepsilon_{idt} \qquad (3E.5)$$

Here, 'i' refers to an individual in district 'd' at time 't'. W_{idt} is log of individual wage, $Female_{idt}$ is an indicator variable for females, C_{idt} are individual characteristics, D_d are district fixed effects and T_t are time fixed effects. The differential impact of rainfall shocks on gender wage gap is given by α_3. A significantly positive value of α_3 indicates that relative female wage is greater in years of high rainfall and lower in years of low rainfall.

To check for the heterogeneous effect of rainfall shocks in rice cultivating areas, I also consider a specification where I interact the rice indicator with the product of rainfall shock and female dummy. Here α'_3 gives the differential impact of rainfall shocks on gender gap in wages in rice growing areas. A significantly positive value of α'_3 indicates that relative female wage is greater in years of high rainfall and lower in years of low rainfall in rice growing areas relative to the other areas. The overall effect of a rainfall shock in rice growing areas on relative female wage is given by $\alpha_3 + \alpha'_3$.

$$W_{idt} = \alpha_0 + \alpha_1 RainShock_{dt} + \alpha_2 Female_{ijt} + \alpha_3 RainShock_{dt} * Female_{idt} + \alpha'_1 RainShock_{dt} * Rice_d + \alpha'_2 Female_{ijt} * Rice_d + \alpha'_3 RainShock_{dt} * Female_{idt} * Rice_d + \alpha_4 C_{idt} + \alpha_5 D_d + \alpha_6 T_t + \varepsilon_{idt} \qquad (3E.6)$$

Table 3.11 shows the individual level results. Females earn on an average 26% less than males in agriculture. Wages increase with positive rainfall shocks and there is no differential impact on male and female wages at the overall level (α_3 is insignificant) in column (1). Column (2) includes the interaction of the rice indicator with the rainfall shock and a female dummy. A positive rainfall shock in rice growing areas increases relative female wage by 8 percent (0.034+0.046) which is significant at 5 % level. Column (3) in table 3.11 shows the individual regressions with the alternative definition of rainfed rice cultivation in a district and the previous findings continue to hold. In general, the wage gap between males and females is smaller in rice growing regions but it exacerbates during negative rainfall shocks and decreases during positive rainfall shocks.

3.8 Possible Mechanisms

As discussed in the theoretical section 3.3, both demand side and supply side factors can result in a differential impact of rainfall shocks on relative female wage. The estimations show that rainfall shocks have a positive effect on relative female wage in rainfed rice growing areas. To understand the gender dynamics in an agricultural labor market, the side of the labor market that is driving the above finding must be determined.

We re-write the expressions obtained for change in male and female equilibrium wage due to a rainfall shock, under the assumptions of the theoretical model constructed in section 3.3, below

$$\frac{dW_M^*}{dr} = \frac{\varepsilon_{F,W_F}\left(L_{M,r}^{S,H} - L_{M,r}^{D}\right) - \varepsilon_{M,W_F}\left(L_{F,r}^{S,H} - L_{F,r}^{D}\right)}{\varphi}$$

$$\frac{dW_F^*}{dr} = \frac{\varepsilon_{M,W_M}\left(L_{F,r}^{S,H} - L_{F,r}^{D}\right) - \varepsilon_{F,W_M}\left(L_{M,r}^{S,H} - L_{M,r}^{D}\right)}{\varphi}$$

Let L_M^S denote total labor supply by males and L_F^S denote total labor supply by females belonging to landed and landless households. In the above model, under what conditions the effect of rainfall shock on relative female wage can be positive (i.e. $\frac{dW_F^*}{dr} - \frac{dW_M^*}{dr} > 0$)? Below we derive the sufficient conditions for this.

Case (i): Given that labor demand and supply elasticities are same for males and females and own price effect is greater than the cross price effect on excess labor demand (i.e. $\varepsilon_{F,W_F} = \varepsilon_{M,Ww_M}$, $\varepsilon_{F,W_M} = \varepsilon_{M,W_F}, |\varepsilon_{F,W_F}| > |\varepsilon_{F,W_M}|$), then only differential shifts in demand and supply of males and females due to a rainfall shock can impact relative female wage. In this case, either shift in female demand due to a rainfall shock is greater ($\frac{\partial L_F^D}{\partial r} > \frac{\partial L_M^D}{\partial r}$) or shift in female supply due to a rainfall shock is greater ($\left|\frac{\partial L_F^S}{\partial r}\right| > \left|\frac{\partial L_M^S}{\partial r}\right|$), for the demand and supply shifts due to a rainfall shock to have a positive impact on relative female wage.

Case (ii) Even if the demand and the supply shifts are equal for males and females, differences in the labor demand and the labor supply elasticities across gender could drive the increase in relative female wage due to a rainfall shock. The effect of a rainfall shock will be positive on relative female wage when demand and supply elasticities differ by gender such that either $\left(\frac{\partial L_F^D}{\partial W_F} + \frac{\partial L_F^D}{\partial W_M} - \frac{\partial L_M^D}{\partial W_M} - \frac{\partial L_M^D}{\partial W_F}\right) > 0$ or $\left(\frac{\partial L_F^S}{\partial W_F} + \frac{\partial L_F^S}{\partial W_M} - \frac{\partial L_M^S}{\partial W_M} - \frac{\partial L_M^S}{\partial W_F}\right) < 0$.

The available literature points at the mechanism of change in female labor demand in rainfed rice growing areas due to changes in crop composition and rice cultivating practices when a rainfall shock occurs. Pandey, Bhandari and Hardy (2007) document in their report the coping strategies of rice growing farmers in regions of eastern India during years of low rainfall. They report that in low rainfall years, yield of other crops are not affected while the yield of rice falls in these rain-fed rice growing regions. As a consequence, around 40% of farmers report changes in cropping pattern in low rainfall years with either an early sowing of the next season crop or replanting with a different crop (Pandey et al. 2007). Usually in rainfed rice growing areas, the rice-wheat or rice-pulse cultivation pattern is followed. Maize, oilseeds, pulses or cash crops like vegetables are usually grown to compensate for loss of rice crop depending on soil suitability and seed availability. Even if farmers do not alter their crop-composition, the demand for female labor could still fall since farmers also change the method of rice cultivation. Pandey et al. (2007) find that 10% - 30% of the farmers in states of eastern India change crop establishment method of rice from transplanting to broadcasting when farmers have not already sown the rice crop. Farmers who had already sown the crop before a drought also replanted the rice crop again if possible (using broadcasting). The technique of broadcasting absorbs less female labor as compared to the technique of transplanting, which can only be conducted in flooded fields (Sen 1985). In high rainfall years farmers use transplanting since broadcasting method of sowing rice gives a lower rice yield (Pathak et al. 2011 and Ehsanullah, Ahmad and Randhawa 2000).

The above documented evidence by Pandey, Bhandari and Hardy (2007) suggests that rice yield should fall relatively more than yield of other crops in these rainfed rice growing regions. As a test of this hypothesis, we estimate the below equations, which estimate the heterogeneous impact of rainfall shocks on yields of major crops.

$$Y_{dt}^{r} = \theta_0 + \theta_1(RainShock)_{dt} + \theta'_1(RainShock)_{dt} * Rice_d + \theta_3 D_d + \theta_4 T_t + \epsilon_{dt} \qquad (3E.7)$$

$$Y_{dt}^{o} = \vartheta_0 + \vartheta_1(RainShock)_{dt} + \vartheta'_1(RainShock)_{dt} * Rice_d + \vartheta_3 D_d + \vartheta_4 T_t + \epsilon_{dt} \qquad (3E.8)$$

Here, Y_{dt}^{r} is the log of yield of rice crop and Y_{dt}^{o} is the log of yield of other crops like wheat, bajra, gram and maize. We expect the crop yields to respond positively to rainfall shocks ($\theta_1 > 0, \vartheta_1 > 0$). If the documented evidence holds in the data, then in the rainfed rice growing areas the yield of rice should be more responsive to rainfall shocks ($\theta'_1 > 0$) while the yields of other crops should be less responsive to rainfall shocks ($\vartheta'_1 < 0$) in comparison to areas not cultivating rainfed rice.

Table 3.12 shows the estimation results for specification $(3E.7)$ and $(3E.8)$. In panel A, '$Rice$' is a dummy variable for major rice growing states whereas in panel B, '$Rice$' is defined as the percentage crop area under rainfed rice cultivation in a district. In panel A, $\theta'_1 > 0$ but is insignificant. In panel B, the estimates show that as area under rainfed rice cultivation increases, the yield of rice is affected significantly more by rainfall shocks ($\theta'_1 > 0$). For other crops - wheat, gram, bajra and maize - their yield is affected significantly less by rainfall shocks in rainfed rice growing regions ($\vartheta'_1 < 0$) in comparison to other regions, in both panels A and B of table 3.12.

The results in panel A show that in rainfed rice growing areas the impact of rainfall shocks on rice yield is 10% ($\theta_1 + \theta'_1$) while that on yield of wheat, bajra, gram and maize is 4%, -6%, 2.5% and -3% respectively ($\vartheta_1 + \vartheta'_1$).[22] A chi-square test shows that the magnitude of the impact of rainfall shocks on rice yield is greater than on yield of any other crop in these areas. In rainfed rice growing areas, even if the rainfall is below the twentieth percentile, the crop yields for wheat, gram, bajra and maize fall less in comparison to rice. These crops require less water than rice and hence can be grown even in drought like conditions in these regions, which on an average receive higher rainfall.

The above findings on crop yields and the documented evidence on farmers' coping strategies suggests that in rainfed rice growing areas rice is transplanted in high rainfall years whereas in low rainfall years, rice is either grown by broadcasting or is substituted with other crops. Both the coping strategies would lead to a greater fall in female labor demand than male labor demand in low rainfall years. Thus, the observed positive impact of rainfall shocks on relative female wage in rainfed rice growing areas, could be a result of an increase in demand for female labor in these areas during years of higher rainfall.

[22] We cannot reject the null that the impact of rainfall shock on yield of bajra, gram and maize is equal to zero in rainfed rice growing regions.

But could it also be that supply shifts are larger for females or supply elasticities differ by gender such that it leads to a lower (larger) gender wage gap in high (low) rainfall years in these rice growing areas? To elucidate graphically, the diagrams below depict the case when supply and demand for male and female labor is a function of own wage only. In this case, either labor supply of females shifts more than that of males ($\left|\frac{\partial L_F^S}{\partial r}\right| > \left|\frac{\partial L_M^S}{\partial r}\right|$) or, female labor supply has a lower own wage elasticity than male labor supply $\left(\frac{\partial L_F^S}{\partial W_F} < \frac{\partial L_M^S}{\partial W_M}\right)$ for the supply side to drive the result that relative female wage increases in high rainfall years.[23]

Diagram 3.1 shows the case when female labor supply decreases more than male labor supply during a positive rainfall shock. Female labor supply is assumed to have the same wage elasticity as male labor supply in this case. Diagram 3.2 shows the case when female labor supply has a lower own wage elasticity than male labor supply. There is no shift in female or male labor supply when a rainfall shock occurs. In both the diagrams we assume that labor demand elasticity, and labor demand shift when a rainfall shock occurs, do not differ between males and females. When a positive rainfall shock occurs, the demand for males and females increases equally from L_M^D to L'^D_M. The initial equilibrium in a normal year is E for both males and females, and the equilibrium when rainfall is high is E'_M and E'_F for males and females respectively. In both the diagrams (3.1 and 3.2), it can be seen that while relative female wage increases due to the positive shock $\left(\frac{dW_F^*}{dr} - \frac{dW_M^*}{dr} > 0\right)$, the relative female employment decreases due to the positive shock $\left(\frac{dL_F^*}{dr} - \frac{dL_M^*}{dr} < 0\right)$, as compared to a normal year.

[23] As discussed in section 3.3, the sign for $\frac{\partial L_F^S}{\partial r}$ and $\frac{\partial L_M^S}{\partial r}$ is negative due to the income effect of rainfall shocks on landed households. In a year of high rainfall, the profits of landed households increase, which lowers their labor supply due to the income effect.

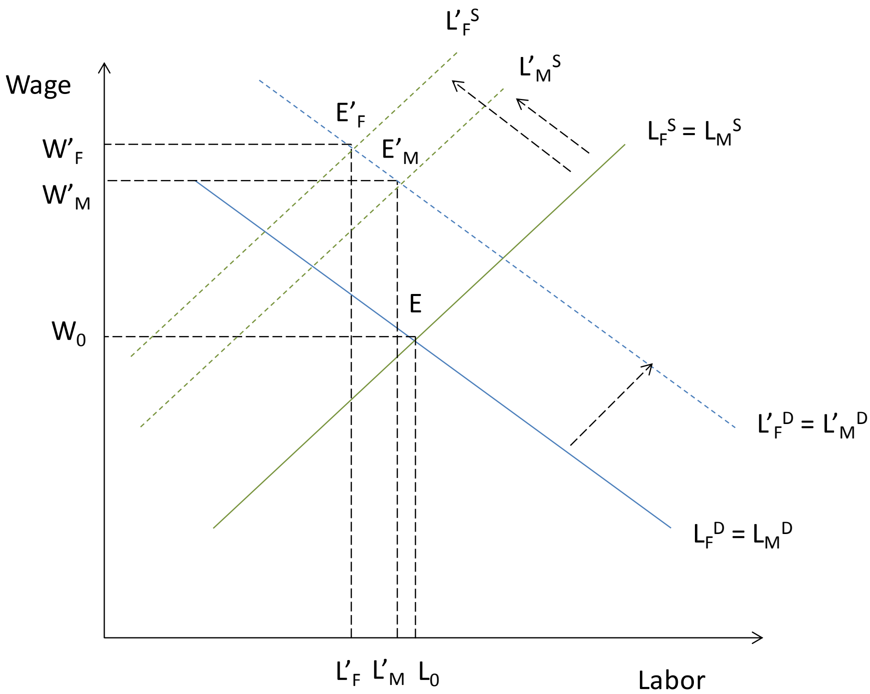


Diagram 3.1 Change in Equilibrium Wage and Employment when Female Labor Supply Falls More than Male Labor Supply and Labor Demand Increases Equally for Females and Males when a Positive Rainfall Shock Occurs

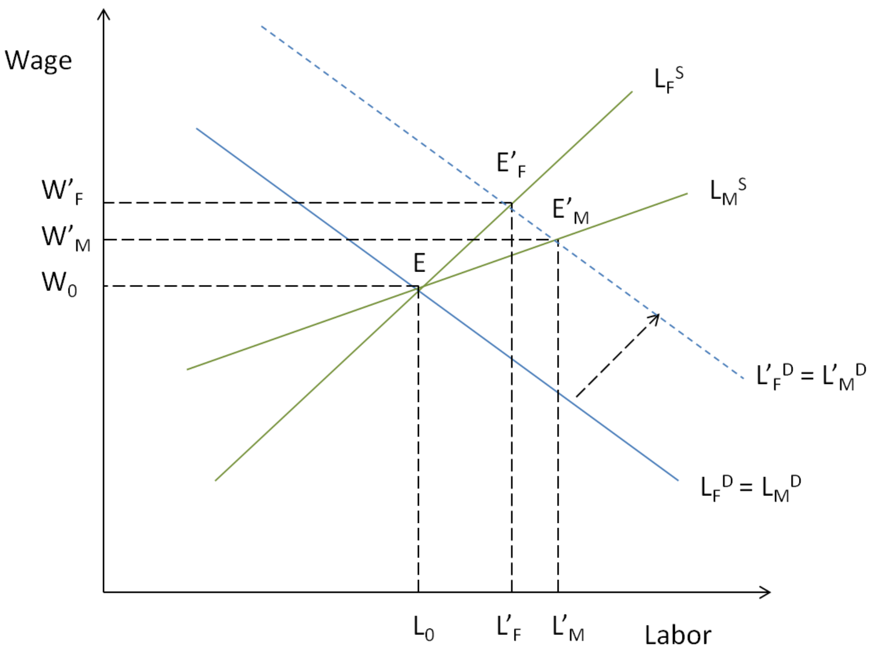


Diagram 3.2 Change in Equilibrium Wage and Employment when Female Labor Supply Elasticity is Smaller than Male Labor Supply Elasticity and Labor Demand Increases Equally for Females and Males when a Positive Rainfall Shock Occurs

In contrast, diagram 3.3 shows the case when demand for female labor increases more than demand for male labor when a positive rainfall shock occurs. As in diagram 3.2, female labor supply is assumed to have a lower own wage elasticity than male labor supply. There is no shift in female or male labor supply when a rainfall shock occurs. We further assume that labor demand elasticity is same for males and females. In this case, when a positive rainfall shock occurs, the demand for males increases from L_M^D to L'^D_M and demand for females increases from L_F^D to L'^D_F. It can be seen from the diagram that now relative female wage increases due to the positive shock $\left(\frac{dW_F^*}{dr} - \frac{dW_M^*}{dr} > 0\right)$ and the relative female employment also increases due to the positive shock $\left(\frac{dL_F^*}{dr} - \frac{dL_M^*}{dr} > 0\right)$, as compared to a normal year. This is because the relative shift in female labor demand is large enough for relative female employment to increase when a positive rainfall shock occurs.

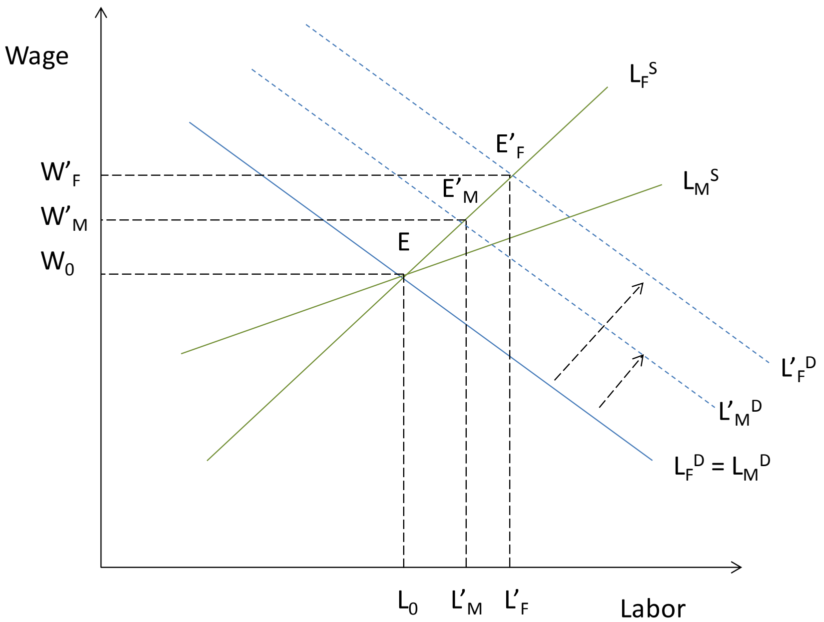


Diagram 3.3. Change in Equilibrium Wage and Employment when Female Labor Supply Elasticity is Smaller than Male Labor Supply Elasticity and Female Labor Demand Increases More than Male Labor Demand when a Positive Rainfall Shock Occurs

In diagrams 3.1 and 3.2, where only differences on the supply side (either differential shifts in labor supply or difference in labor supply elasticities) drive the increase in relative female wage when a positive rainfall shock occurs, relative female employment falls. Only when the demand side factors (in diagram 3.3 this is depicted through a larger shift in female labor demand) also play a role in increasing the relative female wage, that changes in relative female wage and relative female employment due to the rainfall shock, move in the same direction. In appendix 3.A, we relax the assumption that cross wage responsiveness is zero for female and male labor demand and supply. We then show that under certain assumptions, an increase in relative female wage due to a rainfall shock will be accompanied by an increase in relative female employment, when differences on the demand side operate and are large enough to offset the differences on the supply side. When only differences on the labor supply side operate during a rainfall shock and result in an increase in relative female wage, then relative female employment will fall when the rainfall shock occurs.

The above discussion suggests a test to ascertain whether differences on the supply side are driving the positive relationship between relative female wage and rainfall shock in rice growing regions. We evaluate the heterogeneous impact of rainfall shocks on quantity of male and female labor employed in agriculture by estimating the below equations.

$$L^F_{dt} - L^M_{dt} = \delta_0 + \delta_1(RainShock)_{dt} + \delta'_1(RainShock)_{dt} * Rice_d + \delta_2 D_d + \delta_3 T_t + \epsilon_{dt} \quad (3E.9)$$

here $k = M, F$ indexes male and female respectively. 'L' is defined as the log of employment in agriculture by persons aged 15-60 in rural areas of district 'd' at time 't'.[24] If only differences on the labor supply side (differential supply elasticities or differential supply shifts for males and females) are driving the positive relation between relative female wage and rainfall shocks in rice growing regions, then the sign of δ'_1 will be negative.

Table 3.13 shows the estimation results for the specification in equation $3E.9$. In columns (1) and (2), '$Rice$' is defined as a dummy variable for states having at least 50 percent crop area under rice cultivation, whereas in columns (3) and (4), it is defined as proportion of crop area in a district under rainfed rice cultivation. Columns (2) and (4) include other baseline district controls which were incorporated in the wage regressions. The impact of rainfall shocks on relative female employment is positive and significant in rainfed rice growing areas relative to other areas across all the specifications (δ'_1 is positive).[25] This evidence is suggestive of demand side factors (either a larger shift in female labor demand relative to that of males or a lower own wage female labor demand elasticity in comparison to males in absolute terms) operating in rainfed rice growing regions, which result in an increase in relative female wage when a positive rainfall shock occurs.

3.9 Conclusion

In this chapter we examine if shocks to agriculture like rainfall variability over the years affect the daily wages received by men and women differently. This is an important question

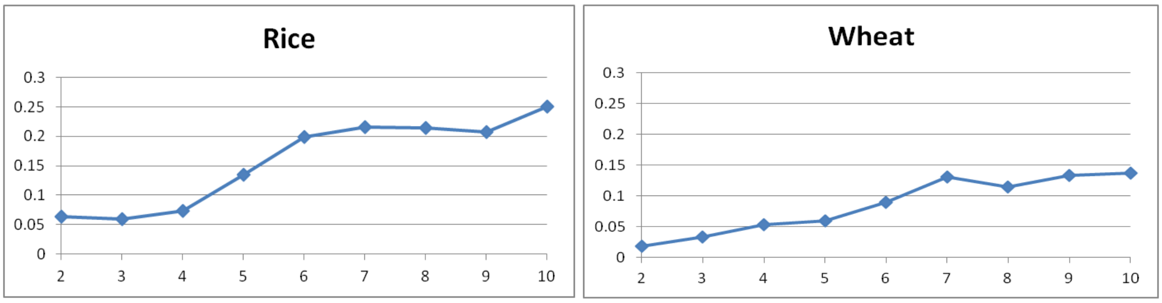
[24] Employment is calculated by taking a district level weighted average of the proportion of days worked in agriculture in the preceding week by each individual. Sampling weights are used to calculate the average so that employment is representative of the district population.

[25] In columns (2) and (4), when other district controls are included, rainfall shock has a significantly positive impact on relative female employment in all areas.

since a large proportion of labor force in developing countries is still engaged in agriculture and the effects of climate change resulting in greater variability of rainfall pose an important challenge to future course of development in these countries. These natural shocks may not have gender neutral effects and can have a differentiated impact on livelihoods of men and women. To my knowledge, this is the first study which undertakes a formal analysis of differentiated impacts of these shocks on male and female wages.

We construct a district level panel data to examine the impact of rainfall shocks on gender wage gap. We find a significantly positive impact of rainfall shocks on female wage relative to male wage as area under rainfed rice cultivation increases in India. In regions where rainfed rice is cultivated, females suffer a greater loss in their wages as compared to males, thus increasing the gender wage gap during low rainfall years. We try to find the mechanism through which the differential effect on wages may occur. The results indicate that the differential effect of rainfall shocks on female and male wages in rainfed rice growing areas can be due to underlying gender roles in agriculture production technology. Creation of non-farm employment opportunities for rural females can mitigate the negative effect of low rainfall on their relative wage and make them less vulnerable to labor market losses when agricultural productivity falls.

Figure 3.1 Impact of Rainfall Deciles on Yield



Source: Center of Climatic Research, University of Delaware; Area, production and Yield statistics, Ministry of Agriculture, India.

Table 3.1 Female to Male Wage Ratio for Casual Laborers in Agriculture

State	1993	1999	2004	2007
Andhra Pradesh	72%	67%	65%	70%
Bihar	87%	88%	87%	89%
Gujarat	98%	89%	90%	99%
Haryana	85%	90%	84%	83%
Karnataka	73%	68%	69%	70%
Kerala	70%	63%	59%	63%
Madhya Pradesh	83%	85%	83%	86%
Maharashtra	63%	65%	63%	68%
Orissa	73%	79%	72%	77%
Punjab	99%	94%	83%	88%
Rajasthan	75%	80%	81%	89%
Tamil Nadu	57%	58%	54%	52%
Uttar Pradesh	75%	78%	83%	84%
West Bengal	88%	89%	88%	94%
All India	72%	72%	70%	74%

Source: National Sample Survey (1993, 1999, 2004, 2007) Employment and Unemployment Schedule (Author’s calculations).

Table 3.2 Variable Definition and Summary Statistics

Variable	Definition	Mean	Standard deviation	Source
Male wage	Real average wage of male casual laborers aged 15-60 in cultivation	10.21	8.55	NSS
Female wage	Real average wage of female casual laborers aged 15-60 in cultivation	8.02	4.48	NSS
Literate	Percentage literate population	0.51	0.13	NSS
Irrigation	Percentage area under irrigation	0.35	0.27	Census
Bus	Percentage villages connected by bus	0.41	0.33	Census
Road	Percentage villages connected by paved roads	0.51	0.29	Census
Rail	Percentage villages connected by rail	0.02	0.02	Census
Bank	Percentage villages having a commercial bank branch	0.09	0.12	Census
Town	Mean distance of a village from a Town (km)	21.34	11.48	Census
Landless	Percentage landless households	0.13	0.12	NSS
Per capita expenditure	Mean monthly per capita expenditure of a household in 1993	292.75	67.69	NSS

Note: Weighted mean with weights equal to district rural population.

Table 3.3 Impact of Rainfall Shocks on Yield of Major Crops

	Rice	Wheat	Bajra	Gram	Maize
RainShock	0.087***	0.057***	0.080***	0.083***	0.018*
	(0.009)	(0.007)	(0.014)	(0.009)	(0.011)
Constant	0.519***	0.656***	-0.217***	-0.288***	0.425***
	(0.014)	(0.011)	(0.023)	(0.015)	(0.017)
District and Year fixed effects	Yes	Yes	Yes	Yes	Yes
Observations	4,112	3,548	2,628	3,560	3,697

Note: The dependent variable is log of yield of a crop in a district according to 2001 census district boundaries for the 14 major states in the analyses. Area, production and yield statistics for years 1999-2008 have been used in the above analysis.

Table 3.4 Rainfall and Cropping Pattern

State	Long term average monsoon rainfall (mm)	Area under rice cultivation
Kerala	1,902.69	95%
West Bengal	1,426.71	74%
Orissa	1,144.99	80%
Bihar	1,011.28	55%
Maharashtra	999.43	8%
Madhya Pradesh	979.59	24%
Uttar Pradesh	863.33	26%
Karnataka	850.24	13%
Gujarat	703.17	7%
Andhra Pradesh	613.68	30%
Punjab	510.39	39%
Haryana	457.95	19%
Tamil Nadu	381.60	39%
Rajasthan	380.28	1%

Source: Center of Climatic Research, University of Delaware. Area, production and yield statistics, Ministry of Agriculture, India.
Note: Rainfall data from 1971-2008 is used to calculate long term average monsoon rainfall for the states. Area, production and yield statistics for years 1999 to 2004 were used to construct the state level crop composition figure since for these years data for all states was complete.

Table 3.5 Impact of Rainfall Shocks on Relative Female Wage

	Female wage	Male wage	Wage ratio
RainShock	0.036**	0.024*	0.012
	(0.014)	(0.014)	(0.013)
Constant	1.259***	1.492***	-0.233***
	(0.017)	(0.017)	(0.014)
Observations	1,216	1,216	1,216
R-squared	0.874	0.903	0.493
District and Year Fixed Effects	Yes	Yes	Yes

Note: Log of wages and log of female to male wage ratio is the dependent variable. The unit of analysis is a district and analytical weights equal to district rural population are used. Robust clustered standard errors are in parenthesis; ***, ** and * indicate significance at the 1, 5 and 10% levels respectively.

Table 3.6 Heterogeneous Impact of Rainfall Shocks on Relative Female Wage

	Female wage	Male wage	Wage ratio
RainShock	0.037***	0.039**	-0.001
	(0.014)	(0.015)	(0.014)
RainShock*Rice	-0.005	-0.053*	0.048*
	(0.032)	(0.032)	(0.025)
Constant	1.259***	1.490***	-0.231***
	(0.017)	(0.017)	(0.014)
Observations	1,216	1,216	1,216
R-squared	0.874	0.903	0.495
District and Year Fixed Effects	Yes	Yes	Yes

Note: Log of wages and log of female to male wage ratio is the dependent variable. Rice is defined as a dummy for states having at least 50 percent crop area under rice cultivation. The unit of analysis is a district and analytical weights equal to district rural population are used. Robust clustered standard errors are in parenthesis; ***, ** and * indicate significance at the 1, 5 and 10% levels respectively.

Table 3.7 Heterogeneous Impact of Rainfall Shocks on Relative Female Wage Conditional on District Controls

	Female wage	Male wage	Wage Ratio
RainShock	0.153***	0.101*	0.051
	(0.056)	(0.055)	(0.052)
RainShock*Rice	-0.005	-0.057*	0.052**
	(0.029)	(0.029)	(0.025)
Literate	0.127	0.042	0.085
	(0.177)	(0.144)	(0.173)
RainShock*Literate	-0.287***	-0.200**	-0.087
	(0.103)	(0.099)	(0.092)
Irrigation	0.034	0.008	0.026
	(0.058)	(0.061)	(0.045)
RainShock* Irrigation	0.108*	0.129**	-0.022
	(0.058)	(0.057)	(0.056)
Constant	1.207***	1.476***	-0.269***
	(0.079)	(0.064)	(0.072)
Observations	1,216	1,216	1,216
R-squared	0.876	0.905	0.496
District and Year Fixed Effects	Yes	Yes	Yes

Note: Log of wages and log of female to male wage ratio is the dependent variable. Rice is defined as a dummy for states having at least 50 percent crop area under rice cultivation. The unit of analysis is a district and analytical weights equal to district rural population are used. Robust clustered standard errors are in parenthesis; ***, ** and * indicate significance at the 1, 5 and 10% levels respectively.

Table 3.8 Heterogeneous Impact of Rainfall Shocks on Relative Female Wage Conditional on Other Development Indicators

Panel A: Development Indicators - Bus, Road, Rail

	Wage Ratio	Wage Ratio	Wage Ratio
Indicator	Bus	Road	Rail
RainShock	0.054	0.057	0.048
	(0.052)	(0.051)	(0.052)
RainShock*Rice	0.059**	0.063**	0.052**
	(0.027)	(0.027)	(0.025)
Literate	0.088	0.097	0.083
	(0.174)	(0.173)	(0.174)
RainShock*Literate	-0.124	-0.132	-0.076
	(0.097)	(0.100)	(0.091)
Irrigation	0.031	-0.016	0.023
	(0.046)	(0.058)	(0.046)
RainShock* Irrigation	-0.029	-0.041	-0.019
	(0.054)	(0.055)	(0.057)
Indicator	-0.079	0.084	0.319
	(0.108)	(0.075)	(0.675)
RainShock* Indicator	0.034	0.040	-0.241
	(0.040)	(0.047)	(0.700)
Constant	-0.241***	-0.296***	-0.273***
	(0.087)	(0.077)	(0.073)
Observations	1,216	1,216	1,216
R-squared	0.496	0.497	0.496
District and Year Fixed Effects	Yes	Yes	Yes

Panel B: Development Indicators - Bank, Town

	Wage Ratio	Wage Ratio	Wage Ratio
Indicator	Bank	Town	All
RainShock	0.052	0.067	0.077
	(0.052)	(0.064)	(0.063)
RainShock*Rice	0.052**	0.053**	0.076***
	(0.025)	(0.024)	(0.028)
Literate	0.085	0.085	0.099
	(0.173)	(0.173)	(0.175)
RainShock* Literate	-0.090	-0.089	-0.142
	(0.091)	(0.093)	(0.102)
Irrigation	0.026	0.025	-0.014
	(0.045)	(0.046)	(0.059)
RainShock* Irrigation	-0.021	-0.029	-0.049
	(0.056)	(0.059)	(0.062)
RainShock* Indicator	0.007	-0.001	
	(0.103)	(0.001)	
Constant	-0.269***	-0.269***	-0.260***
	(0.072)	(0.072)	(0.090)
Observations	1,216	1,216	1,216
R-squared	0.496	0.496	0.498
District and Year Fixed Effects	Yes	Yes	Yes

Panel C: Development Indicators – Landless, Per Capita Expenditure		
	Wage Ratio	Wage Ratio
Indicator	Landless	Per capita expenditure
RainShock	0.082	0.066
	(0.063)	(0.076)
RainShock*Rice	0.074***	0.078***
	(0.028)	(0.029)
Literate	0.088	0.101
	(0.176)	(0.174)
RainShock* Literate	-0.136	-0.146
	(0.104)	(0.109)
Irrigation	-0.014	-0.012
	(0.059)	(0.060)
RainShock*Irrigation	-0.046	-0.051
	(0.061)	(0.062)
Indicator	-0.047	
	(0.094)	
RainShock*Indicator	-0.082	0.000
	(0.099)	(0.000)
Constant	-0.247***	-0.261***
	(0.091)	(0.089)
Observations	1,216	1,216
R-squared	0.498	0.498
District and Year Fixed Effects	Yes	Yes

Note: Log of female to male wage ratio is the dependent variable. Rice is defined as a dummy for states having at least 50 percent crop area under rice cultivation. The data for number of banks in a district and mean distance to nearest town are available for Census 2001 only, resulting in no variation over time in these indicators. District level per capita expenditure is estimated from NSS for the year 1993-94, which controls for initial prosperity levels in a district. Panel C includes all the controls in the last column of Panel B along with the poverty indicators. The unit of analysis is a district and analytical weights equal to district rural population are used. Robust clustered standard errors are in parenthesis; ***, ** and * indicate significance at the 1, 5 and 10% levels respectively.

Table 3.9 Heterogeneous Impact of Rainfall Shocks on Relative Female Wage: Two-Stage Least Squares Estimates

	Female wage	Male wage	Wage Ratio
RainShock	0.153***	0.102**	0.051
	(0.048)	(0.048)	(0.045)
RainShock*Rice	-0.008	-0.069**	0.061**
	(0.033)	(0.032)	(0.030)
Literate	0.128	0.045	0.082
	(0.151)	(0.122)	(0.147)
RainShock* Literate	-0.286***	-0.196**	-0.090
	(0.089)	(0.085)	(0.079)
Irrigation	0.033	0.007	0.026
	(0.049)	(0.052)	(0.039)
RainShock*Irrigation	0.109**	0.132***	-0.024
	(0.050)	(0.050)	(0.048)
Constant	0.909***	1.693***	-0.784***
	(0.060)	(0.128)	(0.117)
Observations	1,216	1,216	1,216
F-Stat for First Stage	177.26	177.26	177.26
District and Year Fixed Effects	Yes	Yes	Yes

Note: Log of wages and log of female to male wage ratio is the dependent variable. Rice is defined as a dummy for states having at least 50 percent crop area under rice cultivation. Table 3.B.4 shows the first-stage regression results. The unit of analysis is a district and analytical weights equal to district rural population are used. Robust clustered standard errors are in parenthesis; ***, ** and * indicate significance at the 1, 5 and 10% levels respectively.

Table 3.10 Heterogeneous Impact of Rainfall Shocks on Relative Female Wage: Alternative Definition of Rainfed Rice Cultivation

	Female wage	Male wage	Wage Ratio	Wage Ratio
RainShock	0.152***	0.099*	0.053	0.070
	(0.057)	(0.054)	(0.052)	(0.067)
RainShock*Rice	-0.064	-0.169**	0.104*	0.128*
	(0.067)	(0.065)	(0.056)	(0.074)
Constant	1.202***	1.467***	-0.265***	-0.228***
	(0.079)	(0.064)	(0.072)	(0.080)
Observations	1,216	1,216	1,216	1,216
R-squared	0.877	0.905	0.495	0.525
Baseline District Controls	Yes	Yes	Yes	Yes
District and Year Fixed Effects	Yes	Yes	Yes	Yes
State Time Trend	No	No	No	Yes

Note: Log of wages and log of female to male wage ratio is the dependent variable. 'Rice' is defined as the proportion of crop area in a district under rainfed rice cultivation. The unit of analysis is a district and analytical weights equal to district rural population are used. The regressions include baseline controls for education and irrigation and their interaction with rainfall shock. Robust clustered standard errors are in parenthesis; ***, ** and * indicate significance at the 1, 5 and 10% levels respectively.

Table 3.11 Heterogeneous Impact of Rainfall Shocks on Relative Female Wage: Individual Data

	(1)	(2)	(3)
	Wage	Wage	Wage
RainShock	0.034***	0.170***	0.170***
	(0.004)	(0.019)	(0.019)
Female	-0.268***	-0.226***	-0.235***
	(0.004)	(0.014)	(0.014)
RainShock*Female	0.002	0.046	0.034
	(0.006)	(0.028)	(0.029)
Female*Rice		0.151***	0.447***
		(0.009)	(0.019)
RainShock*Rice		-0.034***	-0.093***
		(0.009)	(0.021)
RainShock* Female *Rice		0.034**	0.068**
		(0.015)	(0.033)
Constant	1.164***	1.099***	1.085***
	(0.016)	(0.024)	(0.024)
Observations	84,793	84,793	84,793
R-squared	0.619	0.624	0.625
Baseline District Controls	No	Yes	Yes
Individual controls	Yes	Yes	Yes
District and Year Fixed Effects	Yes	Yes	Yes

Note: Log of wage is the dependent variable. The baseline controls for education and irrigation and all their interactions with rainfall shock and a gender dummy are included in columns (2) and (3). Individual controls include age, age square and education. In column (1) and (2) rice is defined as a dummy for states having at least 50 percent crop area under rice cultivation. In column (3) rice is defined as the proportion of crop area in a district under rainfed rice cultivation. The regressions are weighted using the sampling weights provided in National Sample Survey. Robust clustered standard errors are in parenthesis; ***, ** and * indicate significance at the 1, 5 and 10% levels respectively.

Table 3.12 Heterogeneous Impact of Rainfall Shocks on Yield of Major Crops

	Rice	Wheat	Bajra	Gram	Maize
Panel A					
RainShock	0.081***	0.062***	0.092***	0.098***	0.032***
	(0.010)	(0.008)	(0.015)	(0.010)	(0.012)
RainShock* Rice	0.020	-0.024	-0.153***	-0.073***	-0.061**
	(0.019)	(0.015)	(0.049)	(0.021)	(0.024)
Constant	0.518***	0.657***	-0.213***	-0.283***	0.428***
	(0.014)	(0.011)	(0.023)	(0.015)	(0.017)
Observations	4,110	3,548	2,628	3,560	3,697
R-squared	0.809	0.878	0.711	0.655	0.681
District and Year Fixed Effects	Yes	Yes	Yes	Yes	Yes
Panel B					
RainShock	0.056***	0.071***	0.117***	0.095***	0.044***
	(0.011)	(0.008)	(0.016)	(0.011)	(0.013)
RainShock* Rice	0.188***	-0.089***	-0.546***	-0.086*	-0.168***
	(0.042)	(0.032)	(0.107)	(0.046)	(0.052)
Constant	0.515***	0.658***	-0.207***	-0.285***	0.430***
	(0.014)	(0.011)	(0.023)	(0.015)	(0.017)
Observations	4,110	3,548	2,628	3,560	3,697
R-squared	0.810	0.879	0.713	0.654	0.682
District and Year Fixed Effects	Yes	Yes	Yes	Yes	Yes

Note: The dependent variable is log of yield of a crop in a district according to Census 2001 district boundaries for the 14 major states in the analyses. Data for year 1999-2008 is used in the above analysis. In Panel A, 'Rice' is defined as a dummy for states having at least 50 percent crop area under rice cultivation. In Panel B, 'Rice' is defined as the proportion of crop area in a district under rainfed rice cultivation. Robust standard errors are in parenthesis; ***, ** and * indicate significance at the 1, 5 and 10% levels respectively.

Table 3.13 Heterogeneous Impact of Rainfall Shocks on Relative Female Employment in Agriculture

	(1)	(2)	(3)	(4)
	Relative Female Employment	Relative Female Employment	Relative Female Employment	Relative Female Employment
RainShock	0.004	0.047**	0.000	0.048**
	(0.008)	(0.020)	(0.008)	(0.019)
RainShock*Rice	0.019*	0.024**	0.059**	0.075***
	(0.010)	(0.010)	(0.026)	(0.026)
Constant	-0.284***	-0.221***	-0.284***	-0.217***
	(0.007)	(0.025)	(0.006)	(0.025)
Observations	1,216	1,216	1,216	1,216
R-squared	0.820	0.823	0.820	0.823
Baseline District Controls	No	Yes	No	Yes
District and Year Fixed effects	Yes	Yes	Yes	Yes

Note: Log of female to male employment is the dependent variable. In columns (1) and (2), rice is a dummy for states having at least 50 percent crop area under rice cultivation. In columns (3) and (4) it is defined as the proportion of crop area in a district under rainfed rice cultivation. Baseline district controls include education and irrigation and their interaction with rainfall shock. The unit of analysis is a district and analytical weights equal to district rural population are used. Robust clustered standard errors are in parenthesis; ***, ** and * indicate significance at the 1, 5 and 10% levels respectively.

Appendices for Chapter 3

3.A Change in Relative Female Employment Due to a Rainfall Shock

Assume a competitive agricultural labor market and three factors of production – land (A), male labor (L_M) and female labor (L_F) - and rainfall shock (r). The production function is continuous and twice differentiable with diminishing returns to each factor and land is fixed in the short run. The male and the female labor are assumed to be imperfect substitutes in production. Both the male and the female labor are supplied by the households. Male and female leisure are imperfect substitutes in the household utility function. Households are price takers and equilibrium wage and employment is determined by the market clearing conditions.

Let W_M^* be the equilibrium male wage, W_F^* be the equilibrium female wage, L_M^* be the equilibrium male employment and L_F^* be the equilibrium female employment. In this model, male labor supply (L_M^s) and demand (L_M^D) is a function of male wage (W_M), female wage (W_F) and rainfall shock (r). Similarly, for females, their labor supply (L_F^s) and demand (L_F^D) is a function of their own wage (W_F), male wage (W_M) and rainfall shock (r).

In this model, the male and the female labor demand are positively related to the rainfall shock.

$$\frac{\partial L_F^D}{\partial r} > 0, \frac{\partial L_M^D}{\partial r} > 0$$

On the other hand, the male and the female labor supply are negatively related to the rainfall shock i.e. labor supply increases when there is a negative rainfall shock (because landed households increase their labor supply, due to an income effect of decline in their profits, when the rainfall shock is negative).

$$\frac{\partial L_F^S}{\partial r} < 0, \frac{\partial L_M^S}{\partial r} < 0$$

Under the assumptions of the model, the own and the cross price derivates for female and male labor demand take the below signs:

$$\frac{\partial L_F^D}{\partial W_F} < 0, \frac{\partial L_M^D}{\partial W_M} < 0, \frac{\partial L_F^D}{\partial W_M} > 0, \frac{\partial L_M^D}{\partial W_F} > 0,$$

The signs for labor supply response to a change in own wage depend on further assumptions. Labor supply can either increase with increase in own wage (when substitution effect dominates) or decrease with increase in own wage (when income effect dominates). Hence, we consider the below cases.

Case I: Male and female labor supply are upward sloping

$$\frac{\partial L_F^S}{\partial W_F} > 0, \frac{\partial L_M^S}{\partial W_M} > 0, \frac{\partial L_F^S}{\partial W_M} < 0, \frac{\partial L_M^S}{\partial W_F} < 0$$

Case II: Male labor supply is upward sloping and female labor supply decreases with own wage.[1]

$$\frac{\partial L_F^S}{\partial W_F} < 0, \frac{\partial L_M^S}{\partial W_M} > 0, \frac{\partial L_F^S}{\partial W_M} < 0, \frac{\partial L_M^S}{\partial W_F} < 0$$

[1] Since the existing evidence on the direction of female labor supply elasticity in rural areas is mixed we only discuss the case when female labor supply decreases with own wage. Rosenzweig (1984) attempts to estimate labor supply elasticity for males and females in Indian agriculture and finds female labor supply to be positive with respect to own wage at household level but to be irresponsive to wages at aggregate level. On the other hand, Dasgupta and Goldar (2006) find a negative response of female labor supply in rural India to own wage, at low wage levels. Goldberg (2011) in another developing country context using an experimental setting, which takes care of identification issues, finds that male and female labor supply in rural areas is equally elastic and positive at the aggregate level.

3.A.1 Case I: Male and Female Labor Supply are Upward Sloping in Own Wage

To arrive at change in relative female employment due to a rainfall shock, we consider below sub-cases under which relative female wage increases due to a rainfall shock. In case I(a) we assume that a rainfall shock shifts only demand for labor such that female labor demand is affected more than male labor demand. In case I(b) we assume that a rainfall shock shifts only supply of labor such that female labor supply is affected more than male labor supply. In case I(c) we derive change in equilibrium employment gap due to a rainfall shock when both case I(a) and case I(b) hold.

3.A.1.1 Case I(a): The rainfall shock shifts only demand for labor

$$\left(\frac{\partial \mathbf{L_F^S}}{\partial \mathbf{r}} = \mathbf{0}, \frac{\partial \mathbf{L_M^S}}{\partial \mathbf{r}} = \mathbf{0}, \frac{\partial \mathbf{L_F^D}}{\partial \mathbf{r}} > \mathbf{0}, \frac{\partial \mathbf{L_M^D}}{\partial \mathbf{r}} > \mathbf{0}, \frac{\partial \mathbf{L_F^D}}{\partial \mathbf{r}} > \frac{\partial \mathbf{L_M^D}}{\partial \mathbf{r}}\right)$$

To arrive at change in equilibrium employment, we move along the supply curve of labor. In equilibrium, employment is equal to the labor supplied at equilibrium wage.

$$L_M^* = L_M^S(W_M^*, W_F^*) \qquad (3.A.1)$$

$$L_F^* = L_F^S(W_M^*, W_F^*) \qquad (3.A.2)$$

We differentiate $(3.A.1)$ and $(3.A.2)$ with respect to rainfall shock and arrive at expressions for $\frac{dL_M^*}{dr}$ and $\frac{dL_F^*}{dr}$.

$$\frac{dL_M^*}{dr} = \frac{\partial L_M^S}{\partial W_M}\frac{dW_M^*}{dr} + \frac{\partial L_M^S}{\partial W_F}\frac{dW_F^*}{dr} \qquad (3.A.3)$$

$$\frac{dL_F^*}{dr} = \frac{\partial L_F^S}{\partial W_F}\frac{dW_F^*}{dr} + \frac{\partial L_F^S}{\partial W_M}\frac{dW_M^*}{dr} \qquad (3.A.4)$$

Under the assumptions of the model,

$$\frac{dW_F^*}{dr} \geq 0, \frac{dW_M^*}{dr} \geq 0, \frac{dL_F^*}{dr} \rightarrow Ambiguous, \ \frac{dL_M^*}{dr} \rightarrow Ambiguous$$

The sign of $\frac{dL_F^*}{dr}$ depends on movement along the supply curve due to a change in female and male labor demand when a rainfall shock occurs. This is ambiguous since there is a positive effect on female employment due to an own-price effect on female labor supply while there is a negative effect on female employment due to a cross-price effect on female labor supply. The change in relative female employment in equilibrium due to a rainfall shock can be written as

$$\frac{dL_F^*}{dr} - \frac{dL_M^*}{dr} = \underbrace{\left(\frac{\partial L_F^S}{\partial W_F} - \frac{\partial L_M^S}{\partial W_F}\right)\left(\frac{dW_F^*}{dr} - \frac{dW_M^*}{dr}\right)}_{\text{First Term}} + \underbrace{\left(\frac{\partial L_F^S}{\partial W_F} + \frac{\partial L_F^S}{\partial W_M} - \frac{\partial L_M^S}{\partial W_M} - \frac{\partial L_M^S}{\partial W_F}\right)\frac{dW_M^*}{dr}}_{\text{Second Term}} \quad (3.A.5)$$

When relative female wage is affected positively by a rainfall shock (i.e. $\frac{dW_F^*}{dr} - \frac{dW_M^*}{dr}$ is positive), the first term is positive. However, the second term is ambiguous. When the own and cross price responsiveness of the male and the female labor supply are equal (i.e. $\frac{\partial L_F^S}{\partial W_F} = \frac{\partial L_M^S}{\partial W_M}, \frac{\partial L_F^S}{\partial W_M} = \frac{\partial L_M^S}{\partial W_F}$), the second term will be equal to zero. When labor supply elasticities are such that they result in an increase in relative female wage when a positive rainfall shock occurs i.e. $\left(\frac{\partial L_F^S}{\partial W_F} + \frac{\partial L_F^S}{\partial W_M} - \frac{\partial L_M^S}{\partial W_M} - \frac{\partial L_M^S}{\partial W_F}\right) < 0$ (refer to section 3.8), then the second term will be negative. Then, for relative female employment to increase due to a rainfall shock, the positive change in relative female employment in the first term, must dominate any negative change in relative female employment due to differences in labor supply elasticities (second term).

3.A.1.2 Case I(b): The rainfall shock shifts only supply of labor

$\left(\frac{\partial L_F^D}{\partial r} = \mathbf{0}, \frac{\partial L_M^D}{\partial r} = \mathbf{0}, \frac{\partial L_F^S}{\partial r} < 0, \frac{\partial L_M^S}{\partial r} < 0, \left|\frac{\partial L_F^S}{\partial r}\right| > \left|\frac{\partial L_M^S}{\partial r}\right|\right)$

We now move along the demand curve to arrive at change in equilibrium employment due to a rainfall shock. In equilibrium, employment will be equal to the labor demanded at equilibrium wage.

$$L_M^* = L_M^D(W_M^*, W_F^*) \qquad (3.A.6)$$

$$L_F^* = L_F^D(W_M^*, W_F^*) \qquad (3.A.7)$$

We differentiate $(3.A.6)$ and $(3.A.7)$ with respect to rainfall shock and arrive at expressions for $\frac{dL_M^*}{dr}$ and $\frac{dL_F^*}{dr}$.

$$\frac{dL_M^*}{dr} = \frac{\partial L_M^D}{\partial W_M}\frac{dW_M^*}{dr} + \frac{\partial L_M^D}{\partial W_F}\frac{dW_F^*}{dr} \qquad (3.A.8)$$

$$\frac{dL_F^*}{dr} = \frac{\partial L_F^D}{\partial W_F}\frac{dW_F^*}{dr} + \frac{\partial L_F^D}{\partial W_M}\frac{dW_M^*}{dr} \qquad (3.A.9)$$

Under the assumptions of the model,

$\frac{dW_F^*}{dr} \geq 0, \frac{dW_M^*}{dr} \geq 0, \frac{dL_F^*}{dr} \rightarrow Ambiguous, \ \frac{dL_M^*}{dr} \rightarrow Ambiguous$

The sign of $\frac{dL_F^*}{dr}$ now depends on movement along the demand curve due to a change in female and male labor supply when a rainfall shock occurs. This is ambiguous since there is a negative effect on female employment due to an own-price effect on female labor demand while there is a positive effect on female employment due to a cross-price effect on female labor demand. The change in relative female employment in equilibrium due to a rainfall shock can be written as

$$\frac{dL_F^*}{dr} - \frac{dL_M^*}{dr} = \underbrace{\left(\frac{\partial L_F^D}{\partial W_F} - \frac{\partial L_M^D}{\partial W_F}\right)\left(\frac{dW_F^*}{dr} - \frac{dW_M^*}{dr}\right)}_{\text{First Term}} + \underbrace{\left(\frac{\partial L_F^D}{\partial W_F} + \frac{\partial L_F^D}{\partial W_M} - \frac{\partial L_M^D}{\partial W_M} - \frac{\partial L_M^D}{\partial W_F}\right)\frac{dW_M^*}{dr}}_{\text{Second Term}} \qquad (3.A.10)$$

When relative female wage is affected positively by a rainfall shock (i.e. $\frac{dW_F^*}{dr} - \frac{dW_M^*}{dr}$ is positive), the first term is negative. However, the second term is ambiguous. When the own and cross price responsiveness of the male and the female labor demand are equal (i.e. $\frac{\partial L_F^D}{\partial W_F} = \frac{\partial L_M^D}{\partial W_M}, \frac{\partial L_F^D}{\partial W_M} = \frac{\partial L_M^D}{\partial W_F}$), the second term will be equal to zero and relative female employment will fall due to the rainfall shock. When labor demand elasticities are such that they result in an increase in relative female wage when a positive rainfall shock occurs i.e. $\left(\frac{\partial L_F^D}{\partial W_F} + \frac{\partial L_F^D}{\partial W_M} - \frac{\partial L_M^D}{\partial W_M} - \frac{\partial L_M^D}{\partial W_F}\right) > 0$ (refer to section 3.8), then the second term will be positive. Then, for relative female employment to increase due to a rainfall shock, the positive change in relative female employment due to differences in labor demand elasticities(second term) must dominate the fall in relative female employment in the first term.

3.A.1.3 Case I(c): The rainfall shock shifts both the demand and the supply of labor

$\left(\frac{\partial L_F^D}{\partial r} > 0, \frac{\partial L_M^D}{\partial r} > 0, \frac{\partial L_F^S}{\partial r} < 0, \frac{\partial L_M^S}{\partial r} < 0, \frac{\partial L_F^D}{\partial r} > \frac{\partial L_M^D}{\partial r}, \left|\frac{\partial L_F^S}{\partial r}\right| > \left|\frac{\partial L_M^S}{\partial r}\right|\right)$

To arrive at change in equilibrium employment due to a rainfall shock, we can either move along the initial demand curve and then along the shift in demand or move along the initial supply curve and then along the shift in supply. We derive the expressions for change in employment when we move along the supply curve. In equilibrium, employment is equal to the labor supplied at equilibrium wage.

$$L_M^* = L_M^S(W_M^*, W_F^*, r) \qquad (3.A.11)$$

$$L_F^* = L_F^S(W_M^*, W_F^*, r) \qquad (3.A.12)$$

We differentiate $(3.A.11)$ and $(3.A.12)$ with respect to rainfall shock and arrive at expressions for $\frac{dL_M^*}{dr}$ and $\frac{dL_F^*}{dr}$.

$$\frac{dL_M^*}{dr} = \frac{\partial L_M^S}{\partial W_M}\frac{dW_M^*}{dr} + \frac{\partial L_M^S}{\partial W_F}\frac{dW_F^*}{dr} + \frac{\partial L_M^S}{\partial r} \qquad (3.A.13)$$

$$\frac{dL_F^*}{dr} = \frac{\partial L_F^S}{\partial W_F}\frac{dW_F^*}{dr} + \frac{\partial L_F^S}{\partial W_M}\frac{dW_M^*}{dr} + \frac{\partial L_F^S}{\partial r} \qquad (3.A.14)$$

Under the assumptions of the model,

$\frac{dW_F^*}{dr} \geq 0, \frac{dW_M^*}{dr} \geq 0, \frac{dL_F^*}{dr} \rightarrow Ambiguous, \ \frac{dL_M^*}{dr} \rightarrow Ambiguous$

Comparing $(3.A.14)$ with $(3.A.4)$, an additional term in the expression for change in equilibrium female employment is the change in female labor supply with respect to rainfall shock $\left(\frac{\partial L_F^S}{\partial r}\right)$. The change in relative female employment in equilibrium due to the rainfall shock can be written as

$$\frac{dL_F^*}{dr} - \frac{dL_M^*}{dr} =$$

$$\underbrace{\left(\frac{\partial L_F^S}{\partial W_F} - \frac{\partial L_M^S}{\partial W_F}\right)\left(\frac{dW_F^*}{dr} - \frac{dW_M^*}{dr}\right)}_{\text{First Term}} + \underbrace{\left(\frac{\partial L_F^S}{\partial W_F} + \frac{\partial L_F^S}{\partial W_M} - \frac{\partial L_M^S}{\partial W_M} - \frac{\partial L_M^S}{\partial W_F}\right)\frac{dW_M^*}{dr}}_{\text{Second Term}} + \underbrace{\left(\frac{\partial L_F^S}{\partial r} - \frac{\partial L_M^S}{\partial r}\right)}_{\text{Third Term}} \qquad (3.A.15)$$

The above expression is same as that obtained in $(3.A.5)$, along with an addition of a third term. The third term captures the change in relative female employment due to a change in relative female labor supply when a rainfall shock occurs. When a rainfall shock results in a larger change in relative female labor supply such that, $\left|\frac{\partial L_F^S}{\partial r}\right| > \left|\frac{\partial L_M^S}{\partial r}\right|$, then $\left(\frac{\partial L_F^S}{\partial r} - \frac{\partial L_M^S}{\partial r}\right) < 0$. The third term is now negative. When relative female wage is affected positively by a rainfall shock (i.e. $\frac{dW_F^*}{dr} - \frac{dW_M^*}{dr}$ is positive), the first term is positive. The second term is ambiguous. The second term will be negative when labor supply elasticities are such that they also result in an increase in relative female wage when a positive rainfall shock occurs i.e. $\left(\frac{\partial L_F^S}{\partial W_F} + \frac{\partial L_F^S}{\partial W_M} - \frac{\partial L_M^S}{\partial W_M} - \frac{\partial L_M^S}{\partial W_F}\right) < 0$.

Thus, for relative female employment to increase in response to a positive rainfall shock, the positive change in relative female employment in the first term, must dominate the negative change in relative female employment due to differences in labor supply elasticities (second term) and the change in relative female supply due to the rainfall shock (third term).

3.A.2 Case II: Male Labor Supply is Upward Sloping and Female Labor Supply is Downward Sloping in Own Wage

If only supply side factors operate (i.e. female labor supply shifts and female and male labor supply elasticities are such that they results in a positive relation between rainfall shock and relative female wage) then even under this case relative female employment will fall when a positive rainfall shock occurs. Consider the expressions in Case I(b). The first term is negative and the second term is zero.

However, for the relative female employment to increase when relative female wage increases due to operation of demand side factors, we need more stringent assumptions. For example consider the expressions derived in case 1(c). Under the assumptions of Case II,

$$\frac{dW_F^*}{dr} \geq 0, \frac{dW_M^*}{dr} \geq 0, \frac{dL_F^*}{dr} \leq 0, \ \frac{dL_M^*}{dr} \rightarrow Ambiguous$$

The sign of the first term in expression $(3.A.15)$ changes. When relative female wage is affected positively by a rainfall shock (i.e. $\frac{dW_F^*}{dr} - \frac{dW_M^*}{dr}$ is positive), the first term is now ambiguous. The second term has an ambiguous sign. It is negative when labor supply elasticities are such that they also result in an increase in relative female wage when a positive shock occurs i.e. $\left(\frac{\partial L_F^S}{\partial W_F} + \frac{\partial L_F^S}{\partial W_M} - \frac{\partial L_M^S}{\partial W_M} - \frac{\partial L_M^S}{\partial W_F}\right) < 0$.The third term is negative when a rainfall shock results in a relatively larger shift in female labor supply.

When the supply curve is downward sloping for females, for relative female employment to increase due to a rainfall shock the first term needs to be positive and dominate the remaining terms. The first term is positive only when $\left|\frac{\partial L_F^S}{\partial W_F}\right| < \left|\frac{\partial L_M^S}{\partial W_F}\right|$. Thus, in comparison to case 1(c), additional conditions on own-wage and cross-wage responsiveness of female and male labor supply are required under which the change in relative female employment due to a rainfall shock is positive. When the first term is negative due to $\left|\frac{\partial L_F^S}{\partial W_F}\right| > \left|\frac{\partial L_M^S}{\partial W_F}\right|$, then despite a greater increase in demand for female labor during a positive shock, relative female employment will fall.

3.A.3 Summarizing the Results in 3.A.1 and 3.A.2

The results in case I(a), I(b) and I(c) show that the change in relative female employment and the change in relative female wage move in the same direction, if the employment changes due to relative demand shifts or relative demand elasticities dominate the employment changes due to supply shifts or relative supply responsiveness. Additional assumptions on labor supply behavior are required to obtain this result when the supply curve for female labor is downward sloping. Notably, if only supply side factors result in a positive association between relative female wage and rainfall shocks, then relative female employment will be negatively associated with the rainfall shocks.

3.B Additional Specifications

Table 3.B.1 Effect of Standardized Rainfall on Relative Female Wage

	Female wage	Male wage	Wage ratio	Wage ratio
RainStandardized	0.023**	0.012	0.011	0.027
	(0.010)	(0.009)	(0.008)	(0.035)
RainStandardized*Rice				0.025*
				(0.014)
Constant	1.259***	1.491***	-0.232***	-0.280***
	(0.017)	(0.017)	(0.014)	(0.070)
Observations	1,216	1,216	1,216	1,216
R-squared	0.874	0.902	0.493	0.495
Baseline District controls	No	No	No	Yes
District and Year Fixed effects	Yes	Yes	Yes	Yes

Log of wages and log of female to male wage ratio is the dependent variable. Rice is defined as a dummy for states having at least 50 percent crop area under rice cultivation. Baseline district controls include education and irrigation and their interaction with standardized rainfall. The unit of analysis is a district and analytical weights equal to district rural population are used. Robust clustered standard errors are in parenthesis; ***, ** and * indicate significance at the 1, 5 and 10% levels respectively.

Table 3.B.2 Effect of Positive and Negative Rainfall Shocks on Relative Female Wage

	Female wage	Male wage	Wage ratio	Wage ratio
RainShock Positive	0.021	0.022	-0.022	-0.084
	(0.022)	(0.022)	(0.017)	(0.076)
RainShock Negative	-0.046**	-0.025	-0.002	-0.016
	(0.023)	(0.022)	(0.023)	(0.112)
RainShock Positive*Rice				0.059
				(0.039)
RainShock Negative*Rice				-0.045
				(0.039)
Constant	1.263***	1.492***	-0.229***	-0.253***
	(0.017)	(0.017)	(0.015)	(0.075)
Observations	1,216	1,216	1,216	1,216
R-squared	0.874	0.903	0.493	0.497
Baseline District controls	No	No	No	Yes
District and Year Fixed effects	Yes	Yes	Yes	Yes

Log of wages and log of female to male wage ratio is the dependent variable. Rice is defined as a dummy for states having at least 50 percent crop area under rice cultivation. Baseline district controls include education and irrigation and their interaction with positive and negative rainfall shocks. The unit of analysis is a district and analytical weights equal to district rural population are used. Robust clustered standard errors are in parenthesis; ***, ** and * indicate significance at the 1, 5 and 10% levels respectively.

Table 3.B.3 Heterogeneous Impact of Rainfall Shocks on Relative Female Wage Conditional on Other Development Indicators (Detailed Results)

	Wage Ratio	Wage Ratio	Wage Ratio
Indicator	All	Landless	Per capita expenditure
RainShock	0.077	0.082	0.066
	(0.063)	(0.063)	(0.076)
RainShock*Rice	0.076***	0.074***	0.078***
	(0.028)	(0.028)	(0.029)
Literate	0.099	0.088	0.101
	(0.175)	(0.176)	(0.174)
Literate*RainShock	-0.142	-0.136	-0.146
	(0.102)	(0.104)	(0.109)
Irrigation	-0.014	-0.014	-0.012
	(0.059)	(0.059)	(0.060)
Irrigation*RainShock	-0.049	-0.046	-0.051
	(0.062)	(0.061)	(0.062)
Bus	-0.120	-0.119	-0.121
	(0.115)	(0.117)	(0.115)
Bus*RainShock	0.052	0.062	0.051
	(0.080)	(0.080)	(0.080)
Road	0.094	0.090	0.093
	(0.080)	(0.080)	(0.080)
Road*RainShock	0.024	0.024	0.024
	(0.089)	(0.089)	(0.089)
Rail	0.223	0.148	0.203
	(0.657)	(0.673)	(0.664)
Rail*RainShock	-0.612	-0.700	-0.608
	(0.815)	(0.813)	(0.816)
Town*RainShock	-0.001	-0.001	-0.001
	(0.001)	(0.001)	(0.001)
Bank*RainShock	-0.015	-0.030	-0.020
	(0.106)	(0.110)	(0.112)
Landless		-0.047	
		(0.094)	
Landless*RainShock		-0.082	
		(0.099)	
Per Capita Expenditure*RainShock			0.000
			(0.000)
Constant	-0.260***	-0.247***	-0.261***
	(0.090)	(0.091)	(0.089)
Observations	1,216	1,216	1,216
R-squared	0.498	0.498	0.498
District and Year Fixed effects	Yes	Yes	Yes

Note: Log of female to male wage ratio is the dependent variable. Rice is defined as a dummy for states having at least 50 percent crop area under rice cultivation. The data for number of banks in a district and mean distance to nearest town are available for Census 2001 only, resulting in no variation over time in these indicators. District level per capita expenditure is estimated from NSS for the year 1993-94, which controls for initial prosperity levels in a district. Panel C includes all the controls in the last column of Panel B along with the poverty indicators. The unit of analysis is a district and analytical weights equal to district rural population are used. Robust clustered standard errors are in parenthesis; ***, ** and * indicate significance at the 1, 5 and 10% levels respectively.

Table 3.B.4 Heterogeneous Impact of Rainfall Shocks on Relative Female Wage: Two-Stage Least Squares Estimates (First-Stage Estimates)

First-Stage for Interaction of Rainfall Shock with Rice	
	RainShock*Rice
RainShock	-0.492***
	(0.128)
RainShock*Mean Rainfall	0.001***
	(0.000)
Literate	-0.099
	(0.174)
Literate*RainShock	-0.623**
	(0.293)
Irrigation	-0.021
	(0.091)
Irrigation*RainShock	0.475***
	(0.13)
Constant	0.029
	(0.077)
Observations	1,216
F-Stat for First stage	177.26
District and Year Fixed effects	Yes

Note: The interaction of Rainfall shock with rice is the dependent variable in the first stage regression. Rice is defined as a dummy for states having at least 50 percent crop area under rice cultivation. The unit of analysis is a district and analytical weights equal to district rural population are used. Robust clustered standard errors are in parenthesis; ***, ** and * indicate significance at the 1, 5 and 10% levels respectively.

Chapter 4

Female Labor Supply, Education and Home Production

4.1 Introduction

Women's participation in labor market has often been associated with their access to economic opportunities. The Middle East, South Asia and parts of Northern Africa have historically reported the lowest rates whereas East Asia and the Pacific have the highest rates of female participation in the labor force. Over the last quarter of a century though, the female labor market participation rates have increased in the Middle East, South Asia and Northern Africa while those in East Asia and the Pacific have fallen. At the global level women are increasingly joining the labor force in larger numbers. The gender gap in labor force participation rate has consequently declined from 32 percentage points to 26 percentage points between 1980 to 2009 (World Bank 2012). However, India over the past few decades has witnessed a decline in the labor market participation by women.

Figure 4.1 shows two measures of labor force participation rates (LFPR) for males and females aged 15-65 in rural and urban areas of India.[1] The graphs show that participation

[1] Labor force participation rate is defined as the proportion of people currently working or seeking work. Various measures of labor force participation are captured in the National Sample Survey. In this chapter we use

in labor force has been declining for males and females in both rural and urban areas. The largest decline has been for females in rural India. By the Usual Primary and Subsidiary Status (UPSS), female LFPR has significantly fallen from 51% in 1987 to 39% in 2009 in rural areas.[2] These declines are statistically significant. Since the age group 15-24 includes school and college going population, figure 4.2 plots the LFPR for males and females in the age 25-65. In these graphs we do not observe a fall in male LFPR. The participation rates have fallen only for females when we restrict the population to above 25 years of age. The fall in female LFPR is again substantial in the rural sector. It shows a statistically significant fall from 54 percent in 1987 to 51 percent in 1999 and further to 44 percent in 2009 in rural areas.

We further examine the change in LFPR for rural women by current marital status. Figure 4.3 plots the female LFPR for each marital status category. Only currently married females have reduced their labor force participation over time. For never married females there has in fact been an increase in LFPR between 1987 and 2009. The fall in female LFPR is thus driven by a fall in participation by currently married females, who form approximately 85 percent of the total female population in the age 25-65 (proportion of married women have

---

two definitions: Usual Principal and Subsidiary Status (UPSS) and Daily Status. Appendix 4.A provides a detailed description of these measures.

[2] According to the daily status, LFPR for females (rural) was 36 percent in 1987 which increased to 38 percent in 1999 and fell to 33 percent in 2009 (The fall is statistically significant during 1999 and 2009). The daily status employment rates from 1987 are not comparable to 1999 and 2009 due change in survey methodology. There were a few changes in the NSS survey design for capturing employment in the 1990's. While these changes are unlikely to affect the employment definition by UPSS status but can increase the employment figures by daily status. In the earlier surveys, the daily time disposition for the last week was only captured for those who reported being gainfully employed or seeking work for major time in the week. From 1993 onwards, the daily time disposition for the last week was collected for all individuals.

been increasing over time due to a lower proportion of widowed women since mortality rates in India have been falling over time).[3]

The next question which arises is what are rural married females who are out of the labor force doing? The female unemployment rates have remained unchanged over time. Increase in unemployment does not seem to be at play behind the observed decline in female LFPR. Figure 4.4 shows the change in female domestic work over time in India. The fall in female LFPR has been accompanied by an increase in domestic work (the increase in domestic work is statistically significant). In fact there is a one-to-one relationship between falling female LFPR and rising domestic work by females. What explains the fall in labor force participation by married females in rural India along with an increase in domestic work by them? This chapter aims to answer this question.

In the international context, there is an extensive body of literature which looks into the determinants of female labor force participation. This literature suggests a possible role for two factors behind the female LFPR decline in the Indian context. One is the rising income levels and the other is increasing female education.

Real per capita income has been rising in India. Existing evidence suggests that female labor force participation is the highest in the low income countries and the highly developed countries, dipping down for the middle income countries. This has been referred to as the 'U' shaped relationship of female labor force participation with national income. This relationship came into prominence in the seminal work by Goldin (1995) in a cross-section of countries. Goldin (1995) attributed this to structural shifts in the economy with initial economic development. Initial industrialization increases the availability of blue collar jobs

[3] This is all the more surprising when positioned against rising contraceptive use (40% in 1992 to 55% in 2005) and declining total fertility rate (3.7 in 1992 to 3 in 2005) amongst married females aged 15-49. Source: National Family Health Survey, various rounds.

and women do not work in these jobs due to social norms. As the economy develops further, service sector jobs become available which are more attractive for women.[4] Longitudinal analysis involving panel data models however finds mixed evidence for this relationship. While Tam (2011) finds evidence in support of the hypothesis using data from 1950-1980 for 130 countries, Gaddis and Klasen (2014) using data for additional years are not able to establish the ‘U’ shape. In the Indian context, Lahoti and Swaminathan (2013) also conclude that the ‘U’ shaped relationship hypothesized between income and female labor force participation does not seem to hold when over time variation is incorporated.

Another relationship which has received attention in recent years is the ‘U’ shape of female labor force participation with increase in female education (Lincove 2008). Evidence from developed countries suggests that increase in female education increases female LFPR (Schultz 1994). But in developing countries the evidence is mixed (Aromolaran 2004). In an analyses of five Asian countries namely, Indonesia, Korea, the Philippines, Sri Lanka and Thailand, Cameron, Dowling and Worswick (2001) find that female labor force participation rates respond differently to education in different countries.[5] In Sri Lanka there is initially a fall in probability of working for females with increase in their education but in Thailand, the Philippines, Korea and Indonesia female participation in labor force increases or stays constant with increase in female education. A U-shaped relationship between female labor force participation and female education has been observed for India as well (Das and Desai 2003). Figure 4.5 plots the female labor force participation rates in rural India at different female education levels. The graphs show that female LFPR in rural India exhibits a ‘U’

[4] In the Indian context, service sector expansion has been more than that of the industry in the last few decades.

[5] Theoretically, they attribute the effect of education on female labor supply to the positive wage effect and bargaining effect. They argue that if increase in female education increases bargaining power of women and female preferences are such that they do not prefer to work then this might lead to a fall in female labor force participation with their education.

shaped pattern with female education. The labor force participation by educated females falls till women attain a graduate degree. Similarly, figure 4.6 plots the percentage days spent in domestic work within each education category for rural women. Female domestic work in rural India exhibits an inverted 'U' shaped pattern with female education

The U-shaped behavior of female LFPR with female education, combined with a large increase in primary to secondary education among females in rural India (figure 4.7) could be a contributing factor behind the decline in female LFPR in India. If so, standard models of female labor supply warrant a revision. In the standard models of female labor supply, female education increases her market wage, and this always results in an increase in female labor supply with her education, at the extensive margin. However, primary education also gives basic skills to mothers, which can improve their productivity in home tasks like raising children (Lam and Duryea 1999). This observation leads us to posit a theoretical model that incorporates home production (Gronau 1980). Women engage in home production and their productivity in this activity increases with their education. As a consequence, female labor supply may fall with increase in her education even though education also increases the opportunity cost of not working.[6] The stylized fact that female education has increased and fall in employment has occurred only among married females in rural India, along with an increase in their domestic work, suggests that education has increased their productivity in home production.

The theoretical model is followed by a reduced form econometric analysis that relates female labor force participation and domestic work to its determinants including education. The reduced form regressions show that female LFPR (domestic work) falls (increases) with

[6] Cameron, Dowling and Worswick (2001) suggest an alternative explanation: that females could have a preference for not working with increase in their initial years of education and that this preference could vary across countries. Aromolaran (2004) suggests that increase in female education could increase their reservation wage more than market wage, leading to a fall in female labor market participation.

increase in initial years of female schooling. The findings are consistent with the predictions of the theoretical model and establish the robustness of the relationship between female labor force participation and female schooling (observed in figure 4.5) to the inclusion of several controls including income, education of male members of the household and local labor market structures.

The importance of education in explaining the fall (increase) in female LFPR (domestic work) in India is quantified by a decomposition analysis that specifically looks at the changes in observed female characteristics which can explain the fall in female LFPR (increase in domestic work). We decompose the changes in employment rates into those attributable to changing characteristics, and those due to changes in employment rates of females of given characteristics. The decomposition into explained and unexplained changes is robust to whether we use parametric or non-parametric methods. The parametric decomposition enables us to quantify the economic significance of each socio-economic characteristic in explaining the fall in female LFPR.

To the best of our knowledge, no other study has highlighted the role of education in reducing female labor force participation in rural India. In the context of the literature, the contribution of this chapter is two-fold. First, through the development of a theoretical model, the chapter examines the mechanisms by which female education impacts female work participation. Second, through the empirical exercises, the chapter validates the negative association between female LFPR and education (up to college) and quantifies its relative importance in explaining the decline in female LFPR in India.

In the empirical analyses, we use the data from the Employment and Unemployment rounds of National Sample Surveys (NSS) conducted in India during the years 1987-88, 1999-00 and 2009-10. It is a cross-sectional dataset which covers different households across years but is representative of the population outcomes. NSS are the largest representative national surveys conducted in the country for analyzing the labor force participation and earnings. We restrict our sample to married women in rural India since the fall in participation rates is observed for this group only. Further, we only include women aged 25-65 to exclude the effect of enrollment in higher education on female LFPR.

In the next section we examine the hypotheses explored by previous studies to explain the fall in female LFPR in India. Section 4.3 presents a theoretical model of female labor supply. Section 4.4 describes the reduced form strategy and results. The decomposition methodology used to quantify the contribution of changing socio-economic characteristics of female working age population in explaining the decline and the decomposition results are discussed in section 4.5. Concluding remarks are gathered in section 4.6.

4.2 Existing Literature: Falling Female LFPR in India

A few existing studies (Sen, Neff and Kling 2012; Kannan and Raveendran 2012; Chowdhury 2011; Himanshu 2011) explore possible reasons behind the fall in female LFPR in rural India.[7] In the existing literature the commonly stated factors include the following: increase in enrollment of females in higher education resulting in their unavailability for labor market, increase in real incomes of households resulting in a lesser need for women to engage in economic work, and limited growth in employment opportunities for females. These studies do not quantify the relative importance of these factors in explaining the decline. They also do not examine increasing educational attainment in India as a possible

[7] A concurrent study by Klasen and Pieters (2013) examines the stagnation in labor force participation by females in urban India. They attribute the stagnation to demand and supply side factors. Using decomposition analysis they find that on the supply side, rising household incomes and rising household head education reduced female LFPR in urban India while increase in female education had a limited impact.

factor behind the fall in female LFPR. Below we examine the reasons put forth by existing studies to explain the falling female labor force attachment in India.

4.2.1 Rise in Enrollment in Higher Education

As previously discussed in the introduction, the rise in enrollment does not explain the decline in LFPR for females aged 25 and above. The next question is whether rising enrollment in higher education adequately explain the fall in LFPR observed for age group 15-24? To see this, figure 4.8 plots the proportion of males and females currently enrolled in an educational institution in the age groups of 15-19 and 20-24 and the corresponding LFPR across years. For males, there has been an absolute fall in LFPR by 20 percent and 10 percent in the age groups of 15-19 and 20-24 respectively between 1987 and 2009.[8] This is matched by a 20 percent and a 10 percent increase in current enrollment rate in the age groups 15-19 and 20-24 respectively during the same time period. The increase in education enrollment thus completely explains the fall in male LFPR in the age group 15-24.

For females, there has been an absolute fall in LFPR in the age groups of 15-19 and 20-24 by 20 percent between 1987 and 2009. While there has been a 35 percent and a 7 percent increase in current enrollment rate in the age groups 15-19 and 20-24 respectively during 1987-2009. Thus, for females in the age group 20-24, increase in current enrollment is not an adequate explanation for falling LFPR.[9]

[8] The absolute percentage increase in enrollment and LFPR are calculated e.g. if 37 percent and 60 percent males are enrolled in education in 1987 and 2009 respectively in the age group 15-19 then the increase in enrollment rate is 23 percent. Similarly, the decrease in LFPR is 23 percent when LFPR for males declines from 63 percent in 1987 to 39 percent in 2009. Thus, the change in enrollment rate and LFPR is comparable since they are estimated over the same base population.

[9] This explanation has been recognized as having a limited role in explaining the fall in female LFPR in India by previous studies (Sen, Neff and Kling 2012; Kannan and Raveendran 2012; Chowdhury 2011).

4.2.2 Rise in Household Incomes

The second hypothesis which has been proposed is that decline of females in work force is a consequence of rising household incomes. Household real incomes have increased over time and females from households which have now become better off have withdrawn from the labor force. A few papers have argued that female participation in the labor market increases during times of agrarian distress (Himanshu 2011). Theoretically, a rise in household income and wealth will tend to decrease labor market participation by women. In the Indian context, social norms could also be at play at higher household income levels since female participation in labor market is considered to be a low status activity (Eswaran, Ramaswami and Wadhwa 2013).

Sen, Neff and Kling (2012) find support for this effect by examining the LFPR by male wage-income quintiles across NSS rounds. This approach leads to a substantial reduction in sample since only those households in which male members earn a wage are included. Also, their calculations do not adjust for changes in real wage income over time. In our analysis, we take the monthly per capita expenditure (MPCE) as a proxy for income since this data is available for all the households.[10] This measure of income is however not completely exogenous to participation in the labor force by females. Participation of females in the labor market will increase the per capita expenditure of a household. But this is the best measure available in the data.

To capture the income effect we divide the sample households in deciles according to the household MPCE for the year 1987. The cut-off for each decile in 1987 is then adjusted using consumer price index for agricultural laborers and converted into its nominal value for

[10] NSS does not capture income from self-employment. It also does not contain information on household assets which are a better proxy for wealth.

the years 1999 and 2009. As an illustration, the first decile is defined as those households having MPCE less than Rs 76 in 1987. In nominal terms Rs 76 in 1987 is equivalent to Rs 213 and Rs 531 in 1999 and 2009 respectively. Rs 213 is then defined to be the cut-off for the first decile in 1999. Similarly, Rs 531 is defined to be the cut-off for the first decile in 2009. Between 1987 and 2009, due to increasing real incomes, the proportion of households in lower income deciles has fallen and in the upper income deciles has increased.

Figure 4.9 plots the female LFPR for each decile by year. In 1987, LFPR is the highest for females in the lowest income decile and decreases with each higher decile. In 1999 a similar pattern is observed. However, in 2009, the negative association between female LFPR and MPCE deciles is no longer observed. This is because the largest fall in female LFPR between 1999 and 2009 has been for the lowest income households. Thus, with real income held constant across years, females have decreased their participation in work force. If only the income effect was at play then the decline in female LFPR should have been only due to an increase in number of households in upper income deciles. But a large reduction in female LFPR is among the low income households.[11] This warrants a further investigation into the possible reasons behind the fall in female LFPR in rural India.

4.2.3 Demand Side Effects: Suggestive Evidence from Wage Growth

Sen, Neff and Kling (2012) were not able to find conclusive evidence for demand side effects. They looked at changes in female LFPR at state level and changes in state domestic product. They also analyzed growth rate in sectors where females are mostly employed. They could not find evidence for a greater decline in female LFPR in states where domestic

[11] It could also be that withdrawal of women from the labor force in these households has led to low monthly consumption expenditure.

product growth was low. Even at a sectoral level, agriculture and manufacturing, the largest employers of females in India experienced growth in their value added. However, greater income growth at sectoral level and yet a fall in labor demand can occur when production technology changes such that demand for capital is increased. A concurrent study by Chatterjee, Murgai and Rama (2014) also examines the demand side explanations for fall in female LFPR.[12] In this chapter we do not study demand side changes empirically. We analyze the data on real wage rate to show that it cannot just be the demand side factors which have resulted in a relative female LFPR decline. Changes in relative female supply have contributed to the decline observed in female participation in the labor market.

The descriptive statistics show that maximum decline in female LFPR has occurred during 1999-2009. If greater withdrawal of females from the labor force is a result of diminishing relative demand for females in the labor market, then growth in equilibrium wage rates is likely to be relatively lower for females. Table 4.1 shows the compound annual growth rate (CAGR) in real daily wage rate for males and females during 1999-2009 in rural areas.[13] The overall wage growth has been much higher for females than males. The results however vary by sector. The highest wage growth for females has been in the casual labor sector, which has outpaced male wage growth in this sector. It is notable that withdrawal of females from the labor force has also been greater in the casual labor sector versus the salaried sector. The participation rate of females in casual employment fell from 21 percent to 18 percent between 1999 and 2009. While in salaried employment the participation rate of

[12] They control for local employment structures by including district level controls for persons employed in agriculture, non-farm self-employment, non-farm regular employment and casual work as a proportion of total population. They interpret these variables as capturing availability of jobs. The coefficients of these variables are positive and significant. However, these can potentially reflect supply side factors also, since what we observe in the data is employment and not demand.

[13] Individual wage rate is calculated as the total earnings divided by the total days worked in the last week for an individual. The earnings data is only reported for casual laborers or salaried. For self-employed no earnings data is captured in NSS. The district level wage rate is estimated by averaging over all the individuals who report earnings data in a district by using sampling weights. The wages are then converted into real terms by using consumer price index for agricultural laborers.

females has remained constant at 2 percent. Figure 4.10 further plots the change in wage ratio and change in labor force participation rate for the major states of India. Except for Haryana, Rajasthan and West Bengal, all other states witness a decline in female LFPR. Among the states which have witnessed a decline in female LFPR, only four states, namely, Assam, Himachal Pradesh, Kerala and Punjab have experienced a decline in female to male wage ratio. The remaining states (nine) have experienced an increase in relative female wage along with a decline in female LFPR.

The above evidence supports the hypothesis of relative female wage responding to changing relative female labor supply. However, the role of labor demand cannot be ruled out completely. In the absence of a fall in female labor demand, the relative female wage growth could have been even higher with a smaller fall in female employment.

4.3 Theoretical Framework

In India, only married females have withdrawn from the labor force. They are spending their time away from the labor market on increased domestic duties. Increasing education levels may be one of the reasons behind the observed fall in female LFPR and rise in domestic work. We provide a theoretical framework under which female education can reduce female labor supply. The empirical analysis, later, shows whether increasing female education plays any role in explaining the declining labor force attachment of females in India.

Under the assumption that preferences are constant across education levels, what could be the possible mechanism behind withdrawal of married females from labor force and their greater participation in household chores, with increase in their education? Increase in female education will increase labor market wage rate for a female. A simple model of labor supply would indicate that female labor supply decreases with increase in female education, when the income effect is larger than the substitution effect of increase in wage rate. The income effect from own wage will be large at a very high wage. Primary to secondary levels of education do not result in a massive jump in wage in India (Kharbanda 2012). Mehta et al. (2013) also find that the relationship between schooling and wages is convex in India. Moreover, income effect of own wage cannot explain complete withdrawal of females from labor force at the extensive margin. It can only lead to a reduction in hours at the intensive margin.

One of the mechanisms which could lead to a fall in female labor supply with her education is if female time is also used by a household in home production and productivity of female time spent in home production increases with her education.[14] In this case, an additional home production effect will come into play. When marginal return from spending time in home production is greater than marginal return from working in the labor market, women will withdraw from the labor force to be at home.

Mother's education is an important input into the human capital (health and education) investment of her children (Schultz 2002).[15] Due to the home production effect, married and educated females can prefer to engage in domestic work as their education increases.[16] This can happen when the return from labor market is lower than the increased value of their time in household production like looking after children.

[14] A different channel could operate through preferences. If increase in female education, changes the preference from market work to leisure or home production.

[15] Lam and Duryea (1999) find that increase in female schooling by up to 8 years reduces fertility, increases investments in children, increases wage rate but does not increase female LFPR in Brazil.

[16] Married females in rural India cannot choose not to have any children due to prevalent social norms in rural India. The number of children can vary though.

NSS allows us to estimate the proportion of females who primarily engage in domestic work and tutor their children. This is a very narrow definition of child care since child care involves cooking meals for children, cleaning them up, supervising their activities etc. However, this is the only measure of child care activity available in NSS. Table 4.2 shows the percentage rural females who report that they spend time tutoring children among the females who primarily engage in domestic work. The sample of females is restricted to those who are spouses of married children of the household head and to the households which have children aged 4-15. Since the NSS data does not allow us to match each child to her mother, we follow the above approach in arriving at the estimates. The data shows that the overall proportion of females reporting to tutor their children is low but over the years this proportion has been increasing. Figure 4.11 plots the above proportion for females at each education level. It can be clearly seen from the graph in figure 4.11 that as female education increases, the percentage reporting spending their time at home teaching children increases.[17]

In light of the above discussion, we incorporate home production in our theoretical model. The model is based on the time allocation theory in Becker (1965), which shows that a woman allocates her time between home and market production so as to maximize the utility, given a market wage and shadow value of time in home production. We consider a model of household decision making in which parents derive utility from a consumption good and the human capital of their child. Production of the child's human capital requires wife's time. The choice facing the household is to divide the female time towards market work and human capital production of the child.

[17] In the empirical analysis, we include women aged 25-65. While less number of women aged 45-65 will have their children in age group 4-15, but they could be supervising their grandchildren. Also, if they have been out of the labor force in their youth because they were looking after their children, they are unlikely to enter the labor market later. Even if we restrict our sample to women aged 25-45 and having children in the household, the conclusions from the empirical analysis do not change.

We consider a household with one child and a couple- a husband(M) and a wife(F). We assume a household utility function(U) which abstracts away from intra-household bargaining. The utility function of the household is defined over the human capital(H) of the child and the household's current consumption(c).

$$U = U(H, c)$$

$U_H(H, c) \geq 0, U_{HH}(H, c) \leq 0, U_c(H, c) \geq 0, U_{cc}(H, c) \leq 0$

Husband and wife are each endowed with one unit of time. Production of human capital of the child requires only wife's time(l). So the wife divides up her time between work($1 - l$) and household production(l) of child's human capital. The husband devotes his entire time to working in the labor market. We take market good (c) as the numeraire and denote the market wage rate of husband and wife by W_M and W_F respectively. We further assume that the wage rate received by the husband and the wife are a function of their education levels E_M and E_F respectively.

$$W_M = {E_M}^{\varphi_1}$$

$$W_F = {E_F}^{\varphi_2}$$

$E_M \geq 1, E_F \geq 1, \varphi_1 > 0, \varphi_2 > 0$

The parents maximize their utility subject to a budget constraint and a human capital production function of the child. The production function of the child's human capital is given by

$$H = h(Bl)$$

$H_l(Bl) \geq 0, H_{ll}(Bl) \leq 0, H_B(Bl) \geq 0$

where 'B' is the productivity of the time devoted by wife to human capital production of the child. The productivity of wife's time in production of the child's human capital is considered to be a function of her education level E_F. We posit the below functional form for B

$$B = E_F{}^{\varphi_3}$$

The budget constraint faced by the household is given by

$$c \leq W_M + W_F(1 - l) + \vartheta$$

Where, the first and the second terms on the right hand side are the labor incomes of the husband and the wife respectively, and ϑ is non-labor income of the household. The objective function of the parents can now be written as

$$Max_{l,c}\, U = U(H, c)$$

$$subject\ to: \quad c \leq W_M + W_F(1 - l) + \vartheta$$

$$H = h(Bl)$$

To solve the above model we assume a structure on the utility and the human capital production function. The utility function is quasilinear. It is concave in H and linear in c.

$$U(H, c) = H^g + c$$

where 'g' is the weight placed by the household on human capital(H) of the child. In the above utility function, there will be no income effect of increase in female wage on female labor supply. In reality, there may be some effect present, but as mentioned earlier, it cannot explain the decision of married females to completely opt out of the labor market. The utility function is hence assumed to be quasilinear in consumption in order to abstract away from presence of income effects on female labor supply and to consider other mechanisms which can explain the fall in female labor supply with increase in her education.

We consider two models below. These models differ in the functional forms assumed for human capital production function. In the first model, the only input in the human capital production of the child is wife's time. The second model incorporates another input which is combined with wife's time to produce the child's human capital.

4.3.1 Model 1: Wife's Time is the Only Input in Home Production

Production function of the child's human capital is assumed to utilize only the wife's time (l). It is given by

$$H = Bl$$

The objective function of the parents can now be written as

$$Max_{l,c}\ H^g + c$$

$$subject\ to: \quad c \leq W_M + W_F(1-l) + \vartheta$$

$$H = Bl$$

The Lagrangian for the maximization problem is written below

$$\mathcal{L} = B^g l^g + c - \lambda(c - W_M - W_F(1-l) - \vartheta)$$

The first order conditions for maximization assuming interior solutions

$$l: \quad gB^g l^{g-1} = \lambda W_F \tag{4.1}$$

$$c: \quad 1 = \lambda \tag{4.2}$$

$$\lambda: \quad c - W_M - W_F(1-l) - \vartheta = 0 \tag{4.3}$$

Solving for 'l' from (4.1) and (4.2)

$$l = \left(g \frac{B^g}{W_F}\right)^{\left(\frac{1}{1-g}\right)}$$

Further, substituting for B and W_F in terms of E_F

$$l = \left(g E_F{}^{(\varphi_3 g - \varphi_2)}\right)^{\left(\frac{1}{1-g}\right)} \tag{4.4}$$

We are interested in characterizing the behavior of wife's labor supply with increase in her education. Increase in wife's education increases her wage, which has a substitution effect on wife's time spent in human capital production of the child. The substitution effect of increase in wife's education decreases her time in home production and increases her labor supply. Increase in wife's education also has a home production effect on her labor supply. The home production effect leads to an increase in wife's time spent at home towards the human capital production of her child or a decrease in her labor supply, with increase in her education. The overall effect of wife's education on her labor supply depends on whether the substitution effect or the home production effect dominates. The derivative of wife’s labor supply with respect to her education is

$$\frac{d(1-l)}{dE_F} = \frac{\varphi_2 - \varphi_3 g}{1-g} E_F{}^{\left(\frac{\varphi_3 g - \varphi_2}{1-g} - 1\right)} g^{\left(\frac{1}{1-g}\right)}$$

The substitution effect of wife’s education on her labor supply is given by

$$\frac{\varphi_2}{1-g} E_F{}^{\left(\frac{\varphi_3 g - \varphi_2}{1-g} - 1\right)} g^{\left(\frac{1}{1-g}\right)}$$

The home production effect of wife’s education on her labor supply is given by

$$\frac{-\varphi_3 g}{1-g} E_F^{\left(\frac{\varphi_3 g-\varphi_2}{1-g}-1\right)} g^{\left(\frac{1}{1-g}\right)}$$

Given the assumptions on the utility and the human capital production function, the substitution effect of wife's education on her labor supply is always positive and the home production effect is always negative. Based on parameter values we can characterize the behavior of wife's labor supply with increase in her education.

Case (a): $\varphi_2 > \varphi_3 g$ i.e. female wage rate is greater than the productivity of wife's time spent at home raised by the exponent of the weight placed on human capital production by the household. In this case, the substitution effect of wife's education is greater than its home production effect. Consequently, as wife's education increases, her time spent in home production decreases and her labor supply increases.

$$\frac{d(1-l)}{dE_F} > 0$$

Case (b): $\varphi_2 < \varphi_3 g$ i.e. female wage rate is lower than the productivity of wife's time spent at home raised by the exponent of the weight placed on human capital production by the household. In this case, the home production effect of wife's education is greater than its substitution effect. Hence, as wife's education increases, her time spent in home production increases and her labor supply decreases.

$$\frac{d(1-l)}{dE_F} < 0$$

Intuitively, when $\varphi_2 < \varphi_3 g$, increase in female education results in a larger increase in marginal utility from wife's time spent in home production than its marginal cost. This leads to an increase in female time spent in home production or a decrease in female labor supply in equilibrium.

4.3.2 Model 2: Wife's Time is Combined with another Input in Home Production

In this model we assume a different structure on the human capital production function of the child. We allow wife's time(l) to be used with another input(Z) in the production of child's human capital(H). Other inputs can be considered to be goods purchased by the household like child-care services, tuition etc. The human capital production function is assumed to be constant elasticity of substitution (CES) between other input(Z) and mother's time(l). It is given by

$$H = \left[\pi Z^{\delta} + (1-\pi)B^{\delta} l^{\delta}\right]^{1/\delta}$$

where π is the share parameter of other input ($\pi \in (0,1)$) and δ determines the degree of substitutability ($\delta \in (-\infty,1)$) between other input and wife's time in home production. The elasticity of substitution is $1/(1-\delta)$. When δ=1 the inputs are perfect substitutes, when δ=-∞ the inputs are perfect complements and the production function takes the Cobb-Douglas form when δ=0. Similar to the previous model, 'B' is the productivity of time devoted by wife to the child's human capital and is a function of her education level (E_F).

The objective function of the parents can now be written as

$$Max_{l,c,Z}\ H^{g} + c$$

$$subject\ to: \quad P_Z Z + c \leq W_M + W_F(1-l) + \vartheta$$

$$H = \left[\pi Z^{\delta} + (1-\pi)B^{\delta} l^{\delta}\right]^{1/\delta}$$

where P_Z is price of other input.

The Lagrangian for the maximization problem is written below

$$\mathcal{L} = \left[\pi Z^{\delta} + (1-\pi)B^{\delta} l^{\delta}\right]^{g/\delta} + c - \lambda(P_Z Z + c - W_M - W_F(1-l) - \vartheta)$$

The first order conditions for maximization assuming interior solutions

$$Z: \quad g\pi\left[\pi Z^{\delta} + (1-\pi)B^{\delta} l^{\delta}\right]^{\frac{g-\delta}{\delta}} Z^{\delta-1} = \lambda P_Z \tag{4.5}$$

$$l: \quad g(1-\pi)B^{\delta}\left[\pi Z^{\delta} + (1-\pi)B^{\delta} l^{\delta}\right]^{\frac{g-\delta}{\delta}} l^{\delta-1} = \lambda W_F \tag{4.6}$$

$$c: \quad 1 = \lambda \tag{4.7}$$

$$\lambda: \quad P_Z Z + c - W_M - W_F(1-l) - \vartheta = 0 \tag{4.8}$$

Solving the above equations we obtain the optimal value of wife’s time spent in home production and other input as

$$l = \left[g(1-\pi)^{\frac{g}{\delta}} \left[\left(\frac{\pi}{(1-\pi)(P_Z)^{\delta}}\right)^{\frac{1}{1-\delta}} E_F^{\left(\frac{\delta(1-g)(\varphi_2-\varphi_3\delta)}{(1-\delta)(\delta-g)}\right)} + E_F^{\left(\frac{\delta(\varphi_2-\varphi_3 g)}{(\delta-g)}\right)}\right]^{\frac{g-\delta}{\delta}}\right]^{\frac{1}{1-g}} \tag{4.9}$$

$$Z = \left(\frac{\pi}{(1-\pi)P_Z}\right)^{\frac{1}{1-\delta}} \left[g(1-\pi)^{\frac{g}{\delta}} \left[\left(\frac{\pi}{(1-\pi)(P_Z)^{\delta}}\right)^{\frac{1}{1-\delta}} + E_F^{\left(\frac{\delta(\varphi_3-\varphi_2)}{(1-\delta)}\right)}\right]^{\frac{g-\delta}{\delta}}\right]^{\frac{1}{1-g}} \tag{4.10}$$

As in the previous model, we are interested in characterizing the behavior of wife's labor supply with increase in her education. Increase in wife's education increases her wage, which has a substitution effect on wife's time spent in human capital production of the child. The substitution effect of increase in wife's education decreases her time in home production and increases her labor supply. Increase in wife's education also has a home production effect on her labor supply. The home production effect has an ambiguous sign in this model. We discuss the cases below.

4.3.2.1 Case I: When $\delta > 0$

In this case, the home production effect of increase in wife's education on her labor supply is negative. Based on parameter values we can characterize the behavior of wife's labor supply with her education.

Case I(a): $\varphi_2 > \varphi_3\delta, \varphi_2 > \varphi_3 g$ i.e. female wage rate is greater than the productivity of wife's time spent at home raised by the exponent of the substitutability parameter and female wage rate is greater than the productivity of wife's time spent at home raised by the exponent of the weight placed on human capital production by the household.

In this case as wife's education increases, her time spent in home production decreases and her labor supply increases. Thus, the substitution effect of increase in wife's education is greater than its home production effect.

$$\frac{d(1-l)}{dE_F} > 0$$

Case I(b): $\varphi_2 < \varphi_3\delta, \varphi_2 < \varphi_3 g$ i.e. female wage rate is lower than the productivity of wife's time spent at home raised by the exponent of the substitutability parameter and female wage rate is lower than the productivity of wife's time spent at home raised by the exponent of the weight placed on human capital production by the household.

In this case as wife's education increases, her time spent in home production increases and her labor supply decreases. Thus, the home production effect of increase in wife's education is greater than its substitution effect.

$$\frac{d(1-l)}{dE_F} < 0$$

4.3.2.2 Case II: When $\delta < 0$

The home production effect has an ambiguous sign in this case. Also, when $\delta < 0$, female wage rate is always greater than the productivity of wife's time spent at home raised by the exponent of the substitutability parameter ($\varphi_2 > \varphi_3\delta$). The further sub-cases are discussed below.

Case II(a): $\varphi_2 > \varphi_3 g$ i.e. female wage rate is greater than the productivity of wife's time spent at home raised by the exponent of the weight placed on human capital production by the household. In this case as wife's education increases, her time spent in home production decreases and her labor supply increases.

$$\frac{d(1-l)}{dE_F} > 0$$

Case II(b): $\varphi_2 < \varphi_3 g$ i.e. female wage rate is lower than the productivity of wife's time spent at home raised by the exponent of the weight placed on human capital production by the household. In this case the behavior of wife's time spent in home production and her labor supply, as her education increases, is ambiguous.

$$\frac{d(1-l)}{dE_F} \rightarrow Ambiguous$$

The above discussion shows that in model 2, the additional parameter which plays a role in determining whether female labor supply increases or decreases with her education is the degree of substitutability between female time and other input in human capital production of the child. To understand the mechanism through which the degree of substitutability acts as a pivot, we take the ratio of the first order conditions for Z and l:

$$\frac{U_Z}{U_l} = MRS_{Z,l} = \left(\frac{\pi}{1-\pi}\right)\left(\frac{l}{Z}\right)^{1-\delta}\frac{1}{B^{\delta}} = \frac{P_Z}{{E_F}^{\varphi_2}}$$

When $\delta > 0$, the relative marginal utility from wife's time(l) in comparison to marginal utility from other input(Z) increases as wife's education increases. Thus, the marginal rate of substitution($MRS_{Z,l}$) between other input and wife's time (L.H.S) decreases with increase in wife's education. The price ratio of other input and wife's time (R.H.S) also decreases with increase in wife's education. Whether (l/Z) decreases or increases in equilibrium depends on whether the fall in L.H.S is smaller or greater than the fall in R.H.S due to increase in wife's education. When $\varphi_2 > \varphi_3\delta$, the L.H.S decreases less than the R.H.S. with increase in wife's education. This results in a decrease in the ratio of wife's time to other input (l/Z) in equilibrium. On the other hand, when $\varphi_2 < \varphi_3\delta$, the L.H.S decreases more than the R.H.S. when wife's education increases. This results in an increase in the ratio of wife's time to other input (l/Z) in equilibrium.

As wife's education increases, $MRS_{Z,l}$ decreases. The magnitude of the decline depends on the substitutability parameter (δ). For larger values of δ, the fall in $MRS_{Z,l}$ is larger. Thus, as the degree of substitutability between other inputs and wife's time increases, the decrease in $MRS_{Z,l}$ becomes greater as wife's education increases. The household is then more likely substitute other input with wife's time. This increases wife's time spent at home relative to other input.

When $\delta < 0$, the relative marginal utility from wife's time(l) in comparison to marginal utility from other input(Z) always decreases as wife's education increases. This increases the $MRS_{Z,l}$ (L.H.S) in the above equation. The price ratio of other input and wife's time in production of child's human capital (R.H.S) however decreases with increase in wife's education. Thus, when $\delta < 0$ the ratio of wife's time to other input (l/Z) always falls as wife's education increases. The mechanism is easier to see in the case when other input and wife's time are perfect complements in production of the child's human capital. The production function for $H = \min({E_F}^{\varphi_3} l, Z)$. As wife's education increases, there is a decline in 'l'. At a given level of other input, smaller level of wife's time is required in equilibrium as education of wife increases.

Thus, the behavior of wife's labor supply with her education and how parameters affect this relationship depends on the degree of substitutability between other inputs and wife's time in production of human capital of the child. In this model the substitution effect and the home production effect drive the relationship between wife's education and her labor supply. While the substitution effect always leads to an increase in female labor supply with increase in her education, the home production effect is ambiguous. The home production effect leads to a decline in female labor supply as her education increases when $\delta > 0$. However, the direction of the home production effect is ambiguous when $\delta < 0$.

The above model offers a potential explanation for fall in female labor supply with female education. It presents a theoretical framework under which certain parametric restrictions can lead to a fall in female labor supply with her education. The posited model of female labor supply which incorporates the role of home production thus has the flexibility to generate differential behavior of female labor supply with her education.

4.4 Reduced Form Analysis

The negative relationship between female labor force participation and female schooling observed in figure 4.5 could be mediated by other factors. First, higher education of females could be highly correlated with education of male members in the household. The negative relationship between female labor force participation and female schooling could simply reflect spousal income effect, since educated females tend to get married in more educated and wealthier households. Second, there could be location specific un-observables which could be driving the negative relationship between female labor force participation and female schooling. For example, a greater number of primary and secondary educated females are present in more developed districts of India. These developed districts could have a lower demand for female labor due to differences in availability of jobs across sectors. In this section, we control for these other factors and examine the reduced form relationship between female education and labor supply in India. Similarly, we also estimate the reduced form relationship between female education and domestic work.

4.4.1 Methodology

We estimate the below regression for each year of data in our analysis.

$$Y_{ihd} = \alpha + \beta X_{ihd} + \gamma Z_{hd} + D_d + \varepsilon_{ihd} \qquad (4E.1)$$

where, 'i' refers to an individual, 'h' refers to a household and 'd' refers to a district. The dependent variable varies with the definition of LFPR used. It is a continuous variable when measured using the daily status and is a dichotomous variable when measured using the UPSS definition. 'X' is a vector of individual characteristics like age and education; 'Z' is a vector of household characteristics like land owned, per capita consumption, maximum education of married males aged 18-65 in the household, religion, caste, share of male members in household, children under age five and household size.[18] We control for district fixed effects (D_d) in the above specification to capture other district level un-observables which could have an impact on labor force participation by women.[19] Table 4.3 shows the summary statistics of individual and household characteristics for each year of data.

4.4.2 Results

Table 4.4 and table 4.5 show the results from a reduced form regression of female labor market participation. The dependent variable is constructed using the Daily Status definition in table 4.4 and using the UPSS definition in table 4.5. The base category for own female education is illiterate females. The results in table 4.4 and table 4.5 show that for females with less than primary to higher secondary education there is a significant decline in probability of being in labor force as compared to the illiterate females (decline in probability increases from primary to higher secondary). At graduate and above schooling there is either a significant increase or no change in the probability of working. This indicates a U shaped relationship between female education and participation in labor force. Thus, even in the presence of controls for household income, male education and labor demand at district level, the negative relationship between female labor force participation and female schooling up till secondary education is robust for all years. Over the years, there is a slight increase in probability of working for females with less than primary to middle education when compared to illiterate women.

[18] Education of married males aged 18-65 in the household is the best indicator for income effect from earnings of male members in the household since in NSS data one cannot match the wife to the husband.

[19] Das and Desai (2003) also estimate reduced form regression for female labor market participation but do not control for male education or local labor market effects in a district.

We discuss the effect of other variables on female labor force participation. Increase in male education reduces female labor supply. This may be due to the income effect on female labor supply as male education increases. Female labor market participation is affected positively when a household owns land. Ownership of land can have two effects on female labor supply - a negative effect on female labor supply, when higher land ownership reflects higher wealth, and a positive effect when land is a productive asset in agriculture which generates a source of self-employment. Increase in household income, captured through the monthly expenditure per capita, also decreases female labor supply in 1987 and 1999. This negative relationship does not exist in year 2009. This result may not however reflect the true income effect on female labor supply since the measure of income used can itself be influenced by female participation in the labor market. Among the included variables, male education is one of the best proxies for income effect of increased household earnings on female LFPR. Increase in household size and greater presence of male members reduces female labor supply. Increase in share of children under age five reduces female labor supply but this relationship is insignificant in year 2009.

We now estimate equation $(4E.1)$ with the dependent variable as the number of days spent in domestic work in a reference week by a female.[20] Table 4.6 shows the results from a reduced form regression of female domestic work (Daily Status). The results show that number of days spent in domestic work increase with increase in female schooling up till secondary education, when the reference category is illiterate females. Females with graduate and above schooling spend lower number of days engaged in domestic work as compared to illiterate females. These results indicate an inverted U-shaped relationship between female education and engagement in domestic work. This is a mirror image of the relationship between female education and labor market participation. The relationship between female

[20] Since domestic work is not captured as a subsidiary activity in NSS, we do not use the UPSS definition of engagement in domestic work.

domestic work and education in figure 4.6 is hence robust to the inclusion of several individual and household controls.

Other variables like male education and household income have a positive effect on female domestic work. Since both male education and household income increase household wealth from other sources, this result is expected. The effect of household landholding is not consistent across years. Increase in household landholding has an insignificant effect on female domestic work in 1987 but in 1999, females in households having a larger landholding spend less time in domestic work as compared to landless households. Increase in household size increases domestic work by females in 1987. It does not have a significant effect on female domestic work in 1999 and 2009. A greater presence of children under age five in households increases domestic work by females but this relationship is insignificant in 2009. The only household demographic variable which has a consistent effect on female domestic work is greater share of male members. More male members increase female domestic work and this result is consistent across years.

The reduced form estimates discussed above, provide strong evidence for a negative (positive) association between female labor supply (domestic work) and female education in India up till secondary education. These results are consistent with the theoretical model posited in section 3, where female labor supply falls and her time spent in home production increases with increase in her education, if the marginal return from spending time in labor market is lower than the marginal return from spending time in home production.

4.5 Decomposition Analysis

We undertake a decomposition analysis to estimate the magnitude of decline in female LFPR which is explained by the changing characteristics of working age females (including female education). We first discuss the decomposition methodology and then present the decomposition results.

4.5.1 Methodology

There are parametric and non-parametric approaches for decomposition analyses. We use both the approaches in this chapter to estimate the proportion of decline which can be attributed to changing female labor force characteristics over time. The non-parametric decomposition technique is used to ascertain the credibility of the explained proportions estimated using the parametric decomposition. The advantage of the parametric approach is that it allows us to estimate the contribution of each characteristic to the explained proportion. We describe the decomposition techniques below.

4.5.1.1 Parametric decomposition: Blinder-Oaxaca

The Blinder (1973) and Oaxaca (1973) decomposition technique has been very popular in labor economics for wage decomposition by gender, race and ethnicity. In our analysis we apply this technique to get the proportion of change in female employment over time which can be explained by changing characteristics of the female population.

4.5.1.1.1 Parametric decomposition: Linear

When the outcome variable is linear, the reduced form specification for days worked by each female 'i' in year 't' can be written as:

$$\hat{Y}_i^t = X_i^t \hat{\beta}^t$$

where Y is the outcome variable of 'days spent in labor force during a week', X is a vector of personal and household characteristics and β are the parameter estimates. Since the regression line passes through the means of the variables, the predicted employment differentials between any two years can be decomposed as

$$\bar{Y}^{1987} - \bar{Y}^{2009} = \left[(\bar{X}^{1987} - \bar{X}^{2009})\hat{\beta}^{1987}\right] + \left[\bar{X}^{2009}\left(\hat{\beta}^{1987} - \hat{\beta}^{2009}\right)\right] \qquad (4E.2)$$

$$\bar{Y}^{1987} - \bar{Y}^{2009} = \left[(\bar{X}^{1987} - \bar{X}^{2009})\hat{\beta}^{2009}\right] + \left[\bar{X}^{1987}\left(\hat{\beta}^{1987} - \hat{\beta}^{2009}\right)\right] \qquad (4E.3)$$

The first term on the right hand side of above equations represents female employment changes attributable to changing female characteristics. The second term on the right hand side represents female employment changes attributable to changing employment rates of females of given characteristics. Equation $(4E.2)$ shows the decomposition when coefficients for the year 1987 ($\hat{\beta}^{1987}$) are assumed to represent the true relationship between female employment and the included characteristics. Equation $(4E.3)$ shows the decomposition when coefficients for the year 2009 ($\hat{\beta}^{2009}$) are assumed to represent the true relationship between female employment and the included characteristics. In general, equations $(4E.2)$ and $(4E.3)$ will give different results. This dilemma is usually encountered in decomposition techniques since they are often saddled with the common index number problem – the coefficients of which year should be used for calculating the explained proportion?

We undertake the decomposition analysis for change in female labor force participation rate (FLFPR) between 1987-2009 and 1999-2009. The change in employment rate over time could also reflect changing demand conditions e.g. a fall in female employment for low income groups in 2009 could reflect supply side changes as well as demand side changes in jobs where low income households are likely to work. Hence the coefficients for the years 1987 and 1999 will reflect the employment-characteristic relationship with supply and demand conditions as given in 1987 and 1999 respectively. The coefficients for 2009 can reflect both supply and demand side changes over time. In our analysis we present the decomposition results with the explained proportion evaluated at the regression coefficients for both the years.

4.5.1.1.1 Parametric decomposition: Non-Linear

When female LFPR is estimated using the principal and subsidiary status in NSS, the dependent variable is binary, indicating whether a female is currently in labor force or not. In this case we use the non-linear Blinder-Oaxaca decomposition to obtain the explained proportion. The below reduced form can be estimated using a logit or a probit regression.

$$\hat{Y}_i^t = F(X_i^t \hat{\beta}^t)$$

The term in the first bracket below is the explained proportion calculated at the employment regression coefficients for the year 1987.

$$\bar{Y}^{1987} - \bar{Y}^{2009}$$

$$= \left[\sum_{i=1}^{N^{1987}} \frac{F\left(X_i^{1987}\hat{\beta}^{1987}\right)}{N^{1987}} - \sum_{i=1}^{N^{2009}} \frac{F\left(X_i^{2009}\hat{\beta}^{1987}\right)}{N^{2009}}\right]$$

$$+ \left[\sum_{i=1}^{N^{2009}} \frac{F\left(X_i^{2009}\hat{\beta}^{1987}\right)}{N^{2009}} - \sum_{i=1}^{N^{2009}} \frac{F\left(X_i^{2009}\hat{\beta}^{2009}\right)}{N^{2009}}\right]$$

As in the linear decomposition, we also present results for the estimate of the explained proportion using non-linear decomposition evaluated at the employment regression coefficients for the year 2009. The term in the first bracket below is the explained proportion calculated at the employment regression coefficients for the year 2009.

$$\bar{Y}^{1987} - \bar{Y}^{2009}$$

$$= \left[\sum_{i=1}^{N^{1987}} \frac{F\left(X_i^{1987}\hat{\beta}^{2009}\right)}{N^{1987}} - \sum_{i=1}^{N^{2009}} \frac{F\left(X_i^{2009}\hat{\beta}^{2009}\right)}{N^{2009}}\right]$$

$$+ \left[\sum_{i=1}^{N^{1987}} \frac{F\left(X_i^{1987}\hat{\beta}^{1987}\right)}{N^{1987}} - \sum_{i=1}^{N^{1987}} \frac{F\left(X_i^{1987}\hat{\beta}^{2009}\right)}{N^{1987}}\right]$$

4.5.1.2 Semi-Parametric decomposition

In recent years a generalization of the Blinder-Oaxaca decomposition technique has become popular to decompose differentials.[21] It is a semi-parametric decomposition technique which uses propensity score re-weighting approach (DiNardo, Fortin, Lemieux 1996 (DFL); DiNardo 2002). This technique has been used to decompose wage and earnings differentials across population groups and within the same population groups over time (Leibbrandt,

[21] It is identical to Blinder-Oaxaca decomposition when variable of interest is mean of the outcome variable and there is a single categorical explanatory variable.

Levinsohn, and McCrary 2010; Biewen 2001; Butcher and DiNardo 2002; Hyslop and Mare 2005; Daly and Valletta 2006).

However, it has not been used much in decomposing employment differentials. We are only aware of a paper by Black, Tseng and Wilkins (2011) who use DFL technique for decomposing the role played by demographic characteristics in explaining the decline in male employment rates in Australia. We elucidate the decomposition method below using similar notations as in Black, Tseng and Wilkins (2011). Additionally, we construct counterfactuals using the relationship between characteristics and female labor force participation rate for different years.

Let $E(e|\, x, t)$ denote the mean employment rate for a vector of characteristics 'x' at time 't' and let $f(x|t)$ denote the distribution of characteristics at time 't'. The aggregate employment rate at time 't' can then be expressed as

$$E(e|t) = \int E(e|\, x, t)\ f(x|t)\, dx$$

The notation highlights the fact that both the employment characteristic relationship and the distribution of characteristics can vary over time.

To arrive at the explained proportion, we need to obtain the change in aggregate employment rate due to changing distribution of characteristics over time, with the employment characteristic relationship held constant. Counterfactual employment rates need to be constructed to carry out the above. This is convenient to do with the following notation that distinguishes between the time-period of the employment characteristic relationship from the time-period of the distribution of characteristics. We denote the time from which the set of employment rates for each characteristic is drawn by 't_e', and denote the time from which the distribution of characteristics is drawn by 't_x'. The employment rate at time 't' can then be alternatively expressed as below

$$E(e\colon t_e = t, t_x = t) = \int E(e|\, x, t_e = t)\ f(x|t_x = t)\, dx$$

Rewriting the above, the employment rate in year 1987 can be expressed as:

$$E(e\colon t_e = 1987, t_x = 1987) = \int E(e|\, x, t_e = 1987)\ f(x|t_x = 1987)\, dx$$

$E(e\colon t_e = 1987, t_x = 1987)$, thus denotes the observed employment rate in year 1987, while $E(e\colon t_e = 1987, t_x = 2009)$ denotes the counterfactual employment rate in 1987 i.e. the employment rate that would have been in 1987 had the distribution of individual characteristics been that of 2009. Similarly, $E(e\colon t_e = 2009, t_x = 1987)$ denotes the counterfactual employment rate in 2009 i.e. the employment rate that would have been in 2009 had the distribution of individual characteristics been that of 1987. The explained proportion can be obtained by both the counterfactuals. The difference between the explained proportions calculated by the two counterfactuals is that the former evaluates it at the employment rate for each characteristic in 1987 and the latter evaluates it at the employment rate for each characteristic in 2009. This is in spirit similar to using the employment regression coefficients for the years 1987 and 2009 in Blinder-Oaxaca decomposition.

Case I: Holding constant, the base year (1987) employment characteristic relationship constant over time, the change in aggregate employment between 1987 and 2009 can be decomposed as

Total change= Explained change + Unexplained change

$$E(e{:}\, t_e = 1987, t_x = 1987) - E(e{:}\, t_e = 2009, t_x = 2009)$$

$$= [E(e{:}\, t_e = 1987, t_x = 1987) - E(e{:}\, t_e = 1987, t_x = 2009)]$$

$$+ [E(e{:}\, t_e = 1987, t_x = 2009) - E(e{:}\, t_e = 2009, t_x = 2009)]$$

The proportion of change in employment which can be explained by changing characteristics over time is given by

$$Explained\ proportion = \frac{E(e{:}\, t_e = 1987, t_x = 1987) - E(e{:}\, t_e = 1987, t_x = 2009)}{E(e{:}\, t_e = 1987, t_x = 1987) - E(e{:}\, t_e = 2009, t_x = 2009)}$$

Case II: Holding constant, the base year (2009) employment characteristic employment relationship constant over time, the change in aggregate employment between 1987 and 2009 can be decomposed and the explained proportion obtained as

$$Explained\ proportion = \frac{E(e{:}\, t_e = 2009, t_x = 1987) - E(e{:}\, t_e = 2009, t_x = 2009)}{E(e{:}\, t_e = 1987, t_x = 1987) - E(e{:}\, t_e = 2009, t_x = 2009)}$$

To arrive at the estimate of the explained proportion in each case, we need to estimate the counterfactual employment rate. This is explained below.

In case I, the counterfactual employment rate is

$$E(e{:}\, t_e = 1987, t_x = 2009) = \int E(e|\, x, t_e = 1987)\ f(x|t_x = 2009)\, dx$$

$$= \int E(e|\, x, t_e = 1987)\, \varphi(x)\, f(x|t_x = 1987)\, dx$$

Where, $\varphi(x) = \frac{f(x|t_x=2009)}{f(x|t_x=1987)}$

$\varphi(x)$ is also known as the re-weighting function. We need to get an estimate of $\varphi(x)$ in order to arrive at the counterfactual employment rate. Applying Bayes rule to get an expression for $\varphi(x)$

$$\varphi(x) = \frac{\Pr\,(t_x = 1987)}{\Pr\,(t_x = 2009)} \frac{\Pr\,(t_x = 2009|x)}{\Pr\,(t_x = 1987|x)}$$

The estimate of $\Pr\,(t_x = t)$ is the proportion of observations that belong to the year 't'. The estimate of $\Pr\,(t_x = 1987|x)$ is obtained by a discrete choice model where the dependent variable is a dichotomous variable for the observations belonging to the year 1987 and 'x' are the explanatory characteristics. Therefore, $\varphi(x)$ is estimated for each individual in year 1987.

Given an estimate of the re-weighting function ($\hat{\varphi}_i(x)$), the counterfactual employment rate can be estimated as

$$\hat{E}(e\colon t_e = 1987, t_x = 2009) = \sum_{i=1}^{N_{1987}} \frac{\hat{\varphi}_i(x)\; e_i}{N_{1987}}$$

where, e_i is whether or not a woman is in labor force or the number of days in the labor force last week by a woman in year 1987. $\hat{\varphi}_i(x)$ is the weight given to each individual in arriving at the mean counterfactual employment rate. Note, that in the absence of $\hat{\varphi}_i(x)$, the actual aggregate employment rate in 1987 would have been obtained. The estimated re-weights ($\hat{\varphi}_i(x)$) are such that the distribution of observed characteristics across years is the same. For example, in the above case, the observations in year 1987 are re-weighted such that the distribution of observed characteristics in year 1987 is identical to that in year 2009. If real income is higher in 2009, then a greater re-weight is given to the individuals belonging to households having a higher income in 1987 so that the proportion of individuals in each income decile after re-weighting is identical across years.

Similarly, for case II we can derive the counterfactual employment rate as below:

$$\hat{E}(e: t_e = 2009, t_x = 1987) = \sum_{i=1}^{N_{2009}} \frac{\hat{\varphi}'_i(x)\, e_i}{N_{2009}}$$

Where, $\varphi'_i(x) = \frac{\Pr(t_x=2009)}{\Pr(t_x=1987)} \frac{\Pr(t_x=1987|x)}{\Pr(t_x=2009|x)}$

4.5.2 Decomposition Results

In the decomposition analysis we include age structure of female population, female education, land owned by a household, household income deciles and education of male members in a household. [22] The summary statistics show that there has not been much change in the age structure of female population. There has been a change in the education structure for females with a large reduction in illiteracy. The land ownership structure has changed between 1987 and 1999 with an increase in households with smaller landholdings. But there has been little change between 1999 and 2009. There has been an increase in real household incomes with a larger proportion of females in the upper deciles of real household per capita expenditure in year 2009. Male education is defined as the maximum education of 18-65 year married male members of the household. The education structure of male members in the

[22] Other household characteristics which could possibly explain change in female labor supply like household size, share of children under age five, share of male members, caste and religion have not been included in the main regressions since they do not alter our main results. The decomposition results including these variables are shown in appendix 4.B. These variables do not add much to the explained variation and none of the conclusions change with the inclusion of these variables.

household has altered with a larger proportion of married males with at least middle level of education.

We discuss the results from the parametric and the non-parametric decompositions below. Our first objective is to estimate the proportion of change in participation which can be explained by the changing individual and other household level socio-economic characteristics. Hence, we first discuss the results on proportion of decline in female LFPR which is explained by the changing characteristics and then look at the contribution of each individual characteristic to the explained proportion.

4.5.2.1 Proportion of decline that is explained by observed characteristics

Table 4.7 shows the decomposition results for change in female labor force participation rate (FLFPR) estimated using *Daily Status* definition. We use the linear Blinder-Oaxaca decomposition since the dependent variable is number of days worked in a reference week by a female. For the daily status we show the results for change in FLFPR between 1999 and 2009 only, since daily status employment estimates are not comparable across 1987 and 2009. The specifications across columns differ in the explanatory characteristics included.[23] When all the characteristics are included, 60 percent and 16 percent of the decline in FLFPR between 1999 and 2009 can be explained at the employment regression coefficients for 1999 and 2009 respectively.

Table 4.8 shows the decomposition results for change in FLFPR estimated using the *Usual Principal and Subsidiary Status (UPSS)*. Now, we use the non-linear Blinder-Oaxaca decomposition technique since our dependent variable is binary. Panel A of table 4.8 shows

[23] Specification (4) excludes the income variable since it can be potentially endogenous to female participation in the labor force.

the decomposition results for change in FLFPR between 1987 and 2009. The explained proportion is 70 percent (40 percent) when it is evaluated at the regression coefficients for the year 1987 (2009) and all the characteristics are included. Panel B shows the decomposition results for the change in FLFPR during 1999-2009.[24] In Panel B, the estimates of explained proportion are slightly lower than those in table 4.7 when evaluated at the regression coefficients for 1999. Hence, there are slight differences in the explained proportion depending on the definition of employment.

The results in table 4.7 and table 4.8 show that by evaluating the explained proportion at the regression coefficients for 1987 or 1999 we are able to explain a higher proportion of the fall in FLFPR in comparison to evaluating it at the regressions coefficients for 2009. This is due to changing structural conditions which have led to an alteration in the relationship between individual characteristics and labor force participation by women over time. As discussed earlier, the employment regression coefficients for 1987 and 1999 are likely to reflect the relationship between employment and characteristics, given the demand conditions in 1987 and 1999 respectively. Changing demand conditions over time could have altered the observed relationship between characteristics and employment in 2009.

We now compare the above estimates of explained proportion with the estimates of explained proportion obtained using the non-parametric decomposition technique. Table 4.9 shows the proportion of fall in daily status female employment rates explained by changing characteristics, estimated using the DFL decomposition. The results show that 62 percent (11 percent) of the fall in daily status FLFPR between 1999 and 2009 can be explained by the included characteristics at the regression coefficients for 1999 (2009). These explained proportions are very similar to those obtained in table 4.7 using the Blinder-Oaxaca

[24] The results in table 4.8 show that the proportion of decline in female labor force participation explained by the included characteristics is less during 1999-2009 when compared to 1987-2009.

decomposition. Similarly, the estimates in table 4.10 for the proportion of explained change in female employment rate (UPSS), using the DFL decomposition technique, are similar to the estimates in table 4.8.

To check the validity of semi-parametric results, Black, Tseng and Wilkins (2011) suggest conducting t-tests for individual variable means, across the re-weighted observations and the actual observations, when using the DFL decomposition. Since our aim is to match the distribution of characteristics across years to estimate the explained proportion, we must statistically test if we have achieved it. For example, in panel A of table 4.10, for specification 3, we must test if the reweighted observations in 1987 have the same means for age group, education, land owned by the household and household income as the observations in year 2009. These tests are conducted for specifications (1)-(5) in table 4.9 and table 4.10. The null for equality of each characteristic is not rejected for any specification when evaluating the explained proportion between 1999 and 2009. This shows that we are able to statistically match the means of the characteristics across the observations for years 1999 and 2009. However, when decomposing the change in FLFPR between 1987 and 2009, the equality of means could not be achieved for all characteristics.

Black, Tseng and Wilkins (2011) advocate inclusion of interaction terms among characteristics to achieve a more precise re-weighting function. In our case, for the change in FLFPR between 1987 and 2009, the re-weighting function achieved is not adequate. As a robustness check, we show the decomposition results when the interaction terms among the characteristics are included as explanatory variables while estimating the re-weighting function. The results are presented in column (6) of table 4.9 and table 4.10. The specification in column (6) includes two-way interactions among all the characteristics. The null for equality of each characteristic is not rejected when the interaction terms are included. There are only minor differences in explained proportion across column (6) and column (5) in table 4.9 and table 4.10.

To conclude, at the maximum (minimum) around 60 percent (16 percent) of the fall in female daily employment between 1999 and 2009 is explained by the included characteristics, depending on the base year used to evaluate the explained proportion. Thus a large proportion of the decline in female employment between 1999 and 2009 remains unexplained. What is the unexplained proportion likely to constitute? Firstly, there are supply side factors which have not been included in the above decomposition analysis like changing female and male wage rates. Increase in real male wage is likely to have a negative impact on female labor supply while the impact of an increase in real female wage on female labor supply is ambiguous. These variables are however endogenous to female employment and have been excluded to prevent simultaneity issues in the absence of suitable instruments. Inclusion of education as an explanatory variable captures rise in wage rates due to increasing human capital over time. But it does not capture any shift in the distribution of wage rates. We use per capita household expenditure as a proxy for income effect (because of data limitations) due to increasing wage rates but it may not be a perfect control and is also endogenous to female employment. Second, the unexplained proportion is also likely to reflect changing demand conditions over time.

4.5.2.2 Contribution of characteristics to the estimated explained proportion

The above results show that the explained proportions estimated using both the parametric and the non-parametric techniques are quite similar. This gives us confidence in the Binder-Oaxaca decomposition results and therefore we use its findings to estimate the contribution of each characteristic to the explained proportion for the parametric decomposition. This will inform us about the characteristics which play the most important role in explaining the decline in FLFPR over time.

Consider the contribution of the individual characteristics in explaining the decline using the daily status definition of employment (specification 3 of table 4.7). This specification includes the characteristics of age group, female education, land owned by the household per capita and household income. Results show that an increase in female education and household income are the major drivers for fall in FLFPR when the explained proportion is evaluated at the year 1999 employment regression coefficients. When the explained proportion is evaluated at the 2009 regression coefficients, increase in female education is the primary driver behind the fall in FLFPR. Increase in female education is able to explain almost 28 percent (16 percent) of the fall in female LFPR between 1999 and 2009 when the explained proportion is evaluated at the regression coefficients for the year 1999 (2009).[25] Adding controls for education of male members in the household, in specification (4) and (5), reduces the contribution of female education to the explained proportion by almost half. This is due to a high correlation between female and male education in the same household since educated females are likely to get married in households having educated males.[26] Increasing male education explains 22 percent (15 percent) of the total fall in female LFPR between 1999 and 2009 at the regression coefficients for the year 1999 (2009).

Now we discuss the contribution of individual characteristics in explaining the decline using the UPSS definition of employment (table 4.8). We use the method proposed by Fairlie

[25] This proportion is arrived at by multiplying the contribution of education to the explained variation and the explained proportion. For e.g. in specification (3) of table 4.7, contribution of education to the explained variation is 50% and the explained proportion is 55%., the contribution of education in explaining the decline in female LFPR is then 28%.

[26] Assortative mating has been on a rise in India. The correlation between education of 18-35 year old daughters-in-law in the household with the highest education of married males who are sons of the household head has increased from 0.54 in 1987 to 0.64 in 2009. We reach the same conclusion of a rise in assortative mating on education if we use the average level of education of males in the household.

(2005) for non-linear Blinder-Oaxaca decomposition to calculate the proportion contributed by each set of independent variables to the explained proportion. The results in panel A and panel B (specification 3) show that increasing education levels of females is the driving factor behind the explained proportions. Similar to table 4.7, the contribution of own female education falls by almost half when male education is controlled for in specification (4) and (5). Thus, the main conclusions do not vary by the definition of female employment in NSS. The above results show that own female education and education of male members of the household are the primary contributors to the explained proportion. The other variables do not show consistent results.

Increase in household income is important in explaining the fall in FLFPR between 1987-2009 and 1999-2009 only when the explained proportion is evaluated at the regression coefficients for the year 1987 and 1999 respectively. This is because of the weak negative relationship between FLFPR and household income in 2009. As discussed earlier, male education is a much better indicator of income effect from increased household earnings.

Change in per capita landholding of the household contributes to the explained proportion for the fall in FLFPR between 1987 and 2009 only. Change in FLFPR is affected by change in household landholding per capita since females belonging to households having more land have a larger employment probability (primarily due to self-employment on own land). Greater fragmentation of land holdings over time in India as household land is divided up among sons of a household has led to a fall in female self-employment on own land. However, change in landholding size does not explain the fall in FLFPR between 1999 and 2009. This is because of limited fragmentation of landholdings during 1999-2009.

If female education is an important driving force behind the decline in female labor force participation, we should also observe an increase in domestic work by females with increase in female schooling. We undertake a decomposition exercise to explain the increase in domestic work by females over the years. Table 4.11 presents the Blinder-Oaxaca decomposition results for change in domestic work by daily status between 1999 and 2009. Female education explains 25 percent and 14 percent of the increase in domestic work using the year 1999 and the year 2009 regressions coefficients respectively (specification 1 of table 4.11). This contribution reduces to half when male education is controlled for. These results are similar to that for decomposition of decrease in labor force participation rates, except that now increase in female education explains the increase in domestic work. Increase in household income also explains a substantial part of increase in domestic work (20 percent) when the explained proportion is evaluated at the coefficients for 1999 (specification 2 of table 4.11). But it is not able to explain the increase in domestic work when the coefficients for 2009 are used. The only other variable which is able to consistently explain the increase in domestic work is increase in male education.

To summarize the decomposition results, female and male education are able to consistently explain the decrease in female LFPR and increase in female domestic work across different specifications. When the maximum education of males in a family is included as a control in the decomposition analysis, there is a fall in the proportion explained by own education of females. However, the contribution of increase in own female education to the explained proportion continues to be significant and large, even when male education is controlled for.

4.6 Conclusion

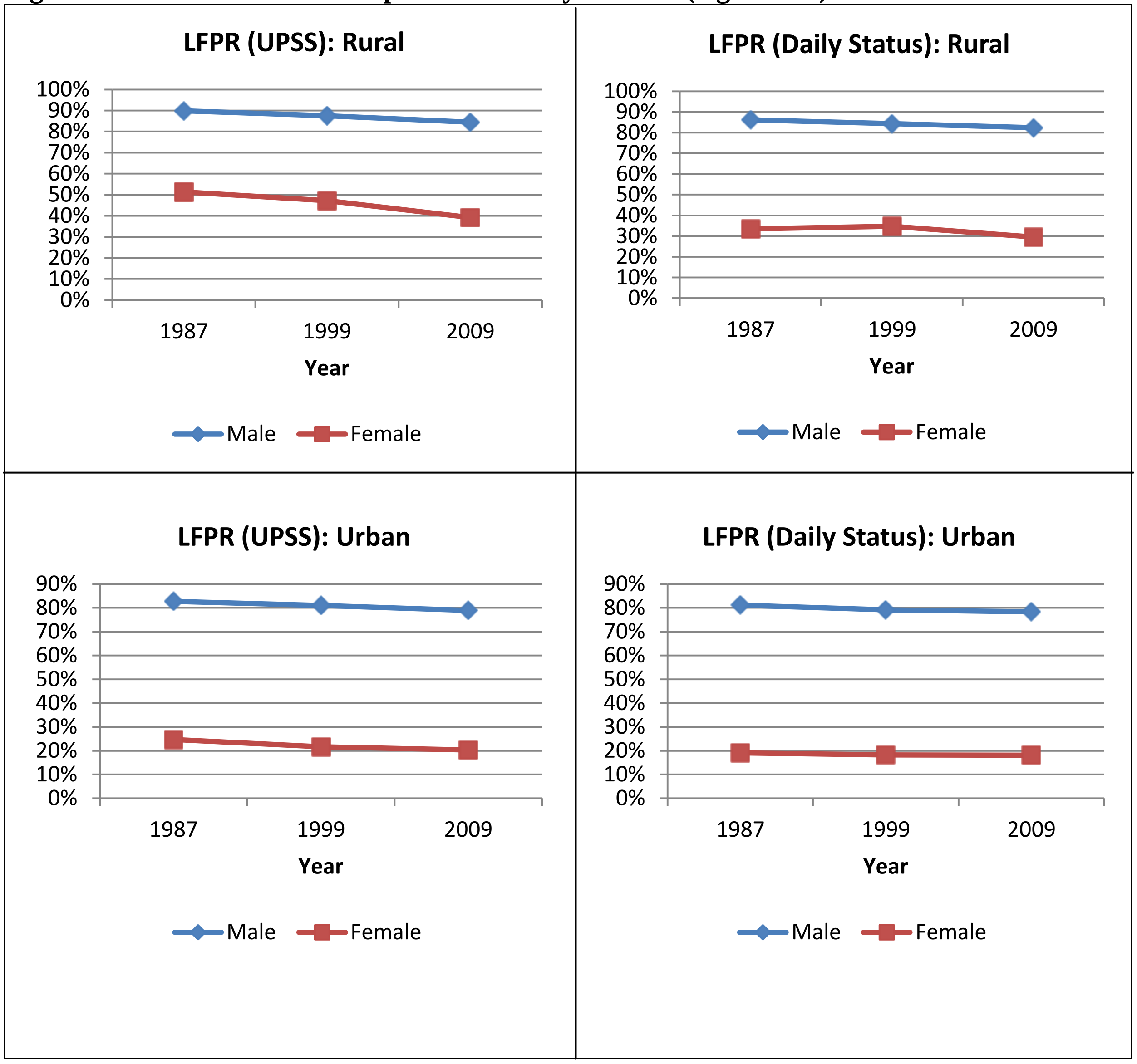
Cross-regional and over time variation in female employment has been studied extensively in economics due to its association with access to economic opportunities for women. In India, female employment in rural areas has fallen over the last three decades, which has also been a period of rapid growth in incomes and educational attainment. This fall has occurred among the married females and has been accompanied by an increase in domestic work by them.

While growth in income can have a negative effect on female labor supply through the wealth effect, the effect of education needs an explanation. We construct a theoretical model of household decision making which incorporates home production to show that increase in female education can reduce female labor supply. Increase in female schooling increases female wage but it also increases her productivity in home production of her child's human capital. This can lead to a withdrawal of married females from the labor force. The gain to the household in terms of increased human capital of the child comes at the expense of female time spent in remunerative activities.

We empirically examine the relationship between female labor supply and education for India including controls for income, male education and local demand. We find that female labor supply falls with increase in female education up till secondary level. On the other hand, female domestic work increases with increase in female education up till secondary level. We then use parametric and non-parametric decomposition techniques to estimate the proportion of fall in female LFPR explained by the changing characteristics of females in working age and identify the important characteristics. The magnitude of the explained proportion depends on the year whose employment- characteristic relationship is used for estimating it. At the maximum 60 percent (16 percent) of the fall in female daily employment during 1999-2009 is explained by the changing characteristics of females in working age, when evaluated at the characteristic employment relationship for the year 1999 (2009). Further, the decomposition exercise shows that increase in female and male education in India play the most important role in explaining the decline. The chapter also estimates the contribution of changing female characteristics to the increase in female domestic work. The decomposition analysis here is a mirror image to the decomposition analysis for female LFPR. It turns out that increases in female and male education are the most important factors associated with the increase in female domestic work.

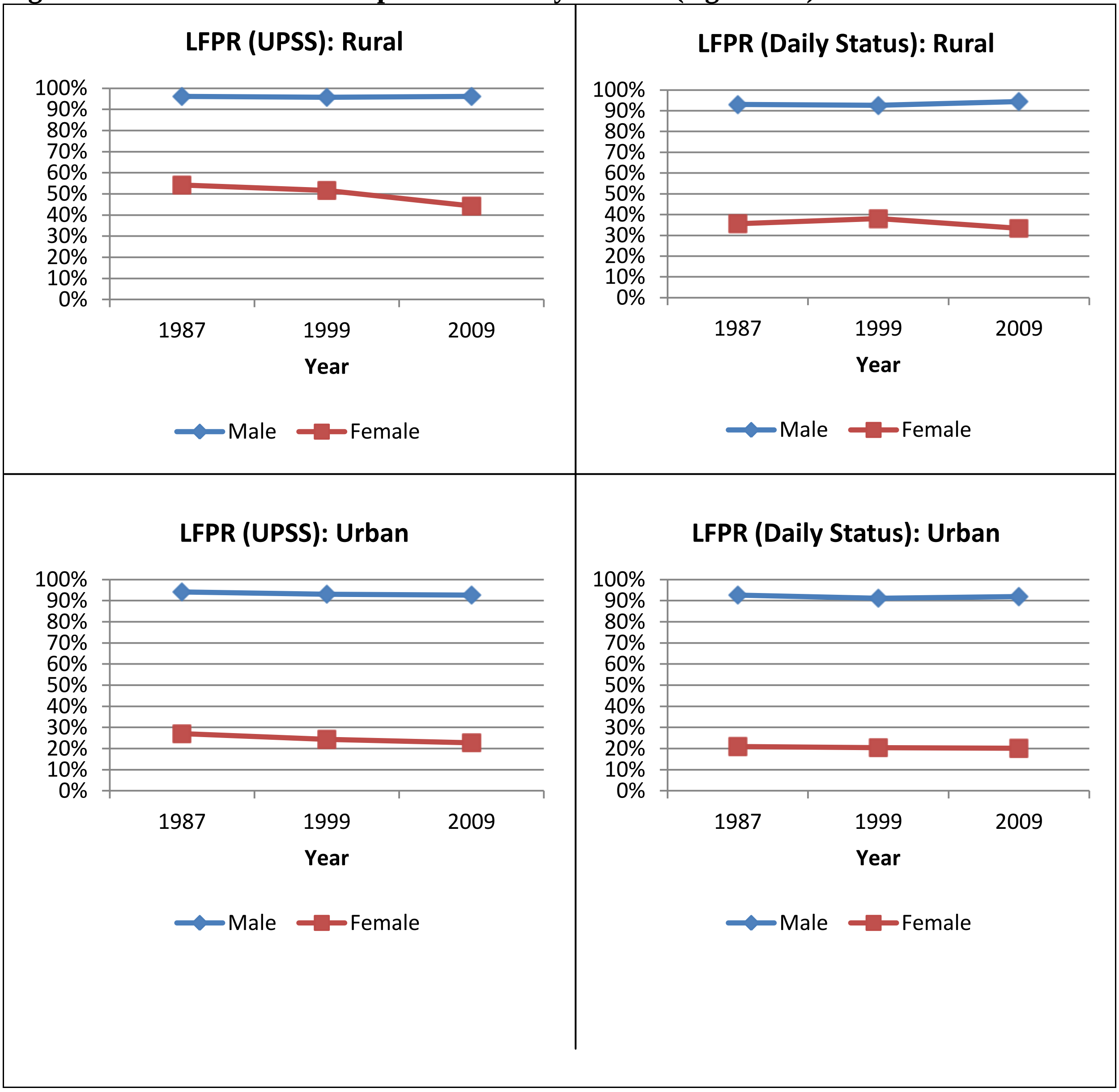
The empirical findings provide robust evidence that home production and female education matter to female labor supply. Male education is also an important factor. The mechanism by which male education matters is an interesting question. One possibility is that it is merely capturing the income effect. Another possibility is that male education alters household preferences. This is not captured in our theoretical model and warrants further exploration.

Figure 4.1 Labor Force Participation Rates by Gender (Age 15-65)



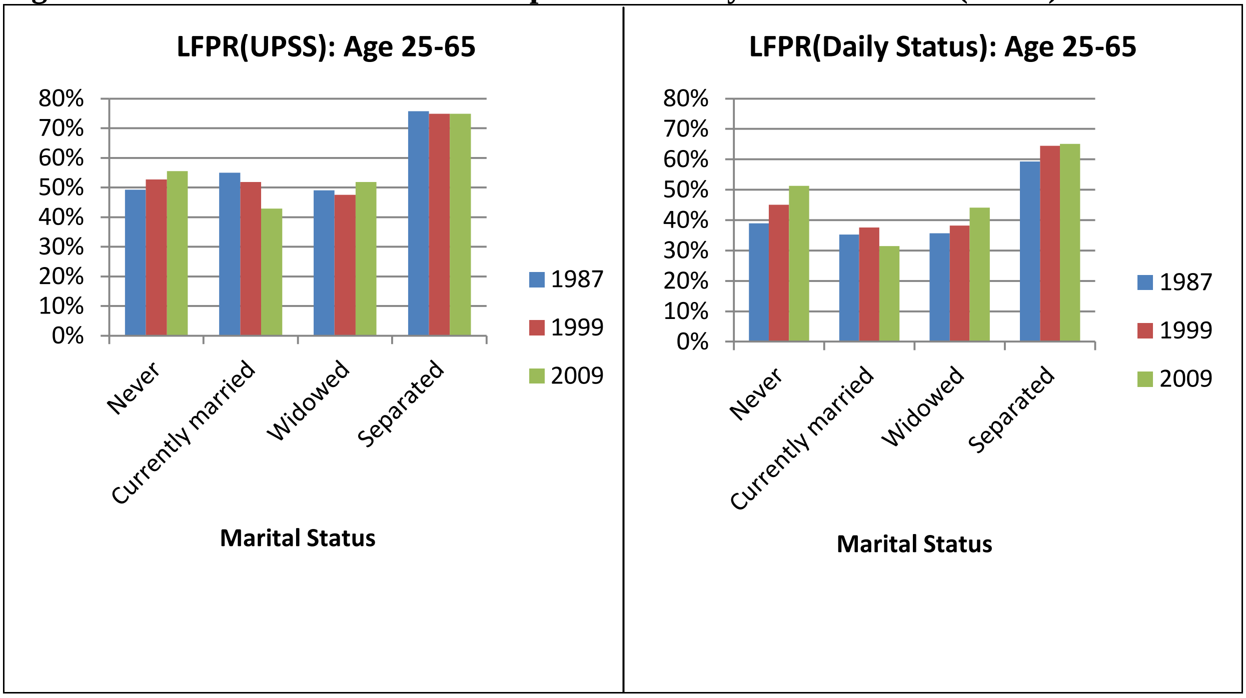
Source: NSS (1987, 1999, 2009) Employment and Unemployment Schedule (Author's calculations).
Note: The definitions of the Usual Principal and Subsidiary Status (UPSS) and the Daily Status labor force participation rates are discussed in appendix 4.A.

Figure 4.2 Labor Force Participation Rates by Gender (Age 25-65)



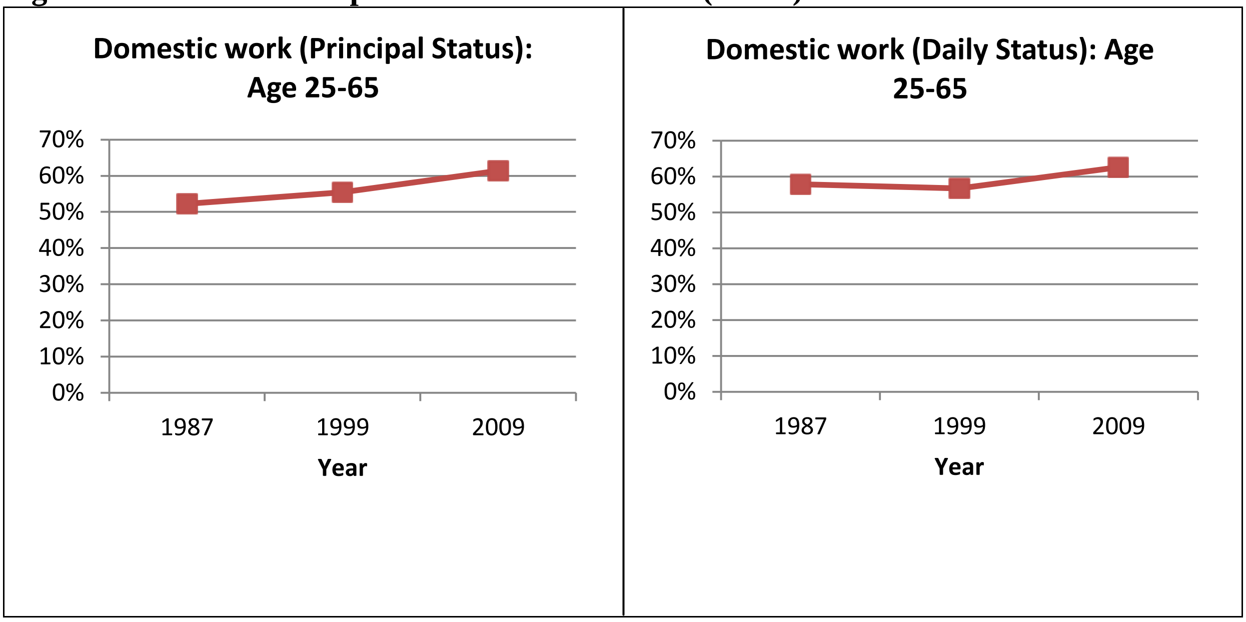
Source: NSS (1987, 1999, 2009) Employment and Unemployment Schedule (Author's calculations).
Note: The definitions of the Usual Principal and Subsidiary Status (UPSS) and the Daily Status labor force participation rates are discussed in appendix 4.A.

Figure 4.3 Female Labor Force Participation Rates by Marital Status (Rural)



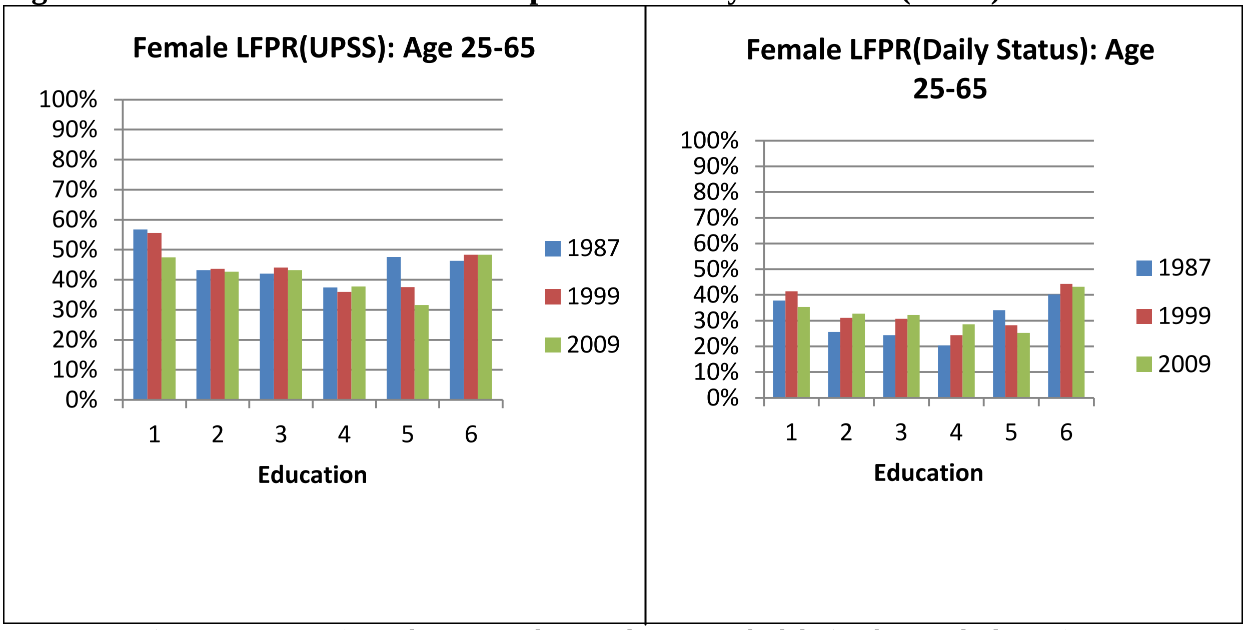
Source: NSS (1987, 1999, 2009) Employment and Unemployment Schedule (Author's calculations).
Note: The definitions of the Usual Principal and Subsidiary Status (UPSS) and the Daily Status labor force participation rates are discussed in appendix 4.A. The LFPR is calculated within each marital category.

Figure 4.4 Female Participation in Domestic Work (Rural)



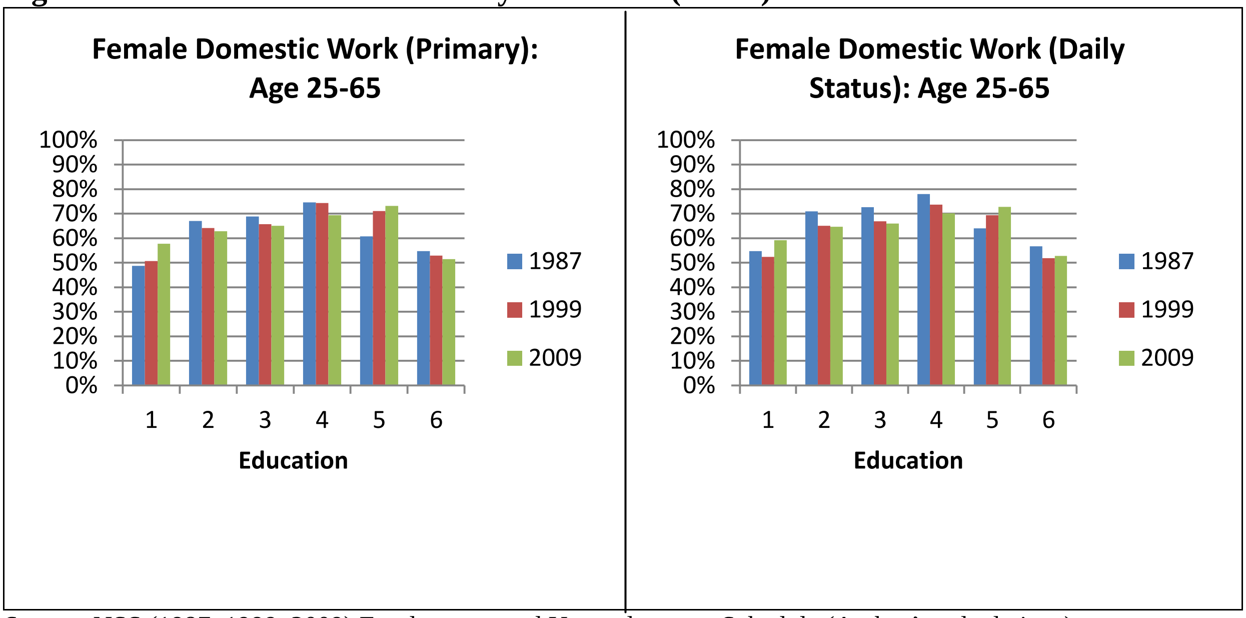
Source: NSS (1987, 1999, 2009) Employment and Unemployment Schedule (Author's calculations).
Note: The definitions of the Usual Principal Activity Status and the Daily Status labor force participation rates are discussed in appendix 4.A.

Figure 4.5 Female Labor Force Participation Rates by Education (Rural)



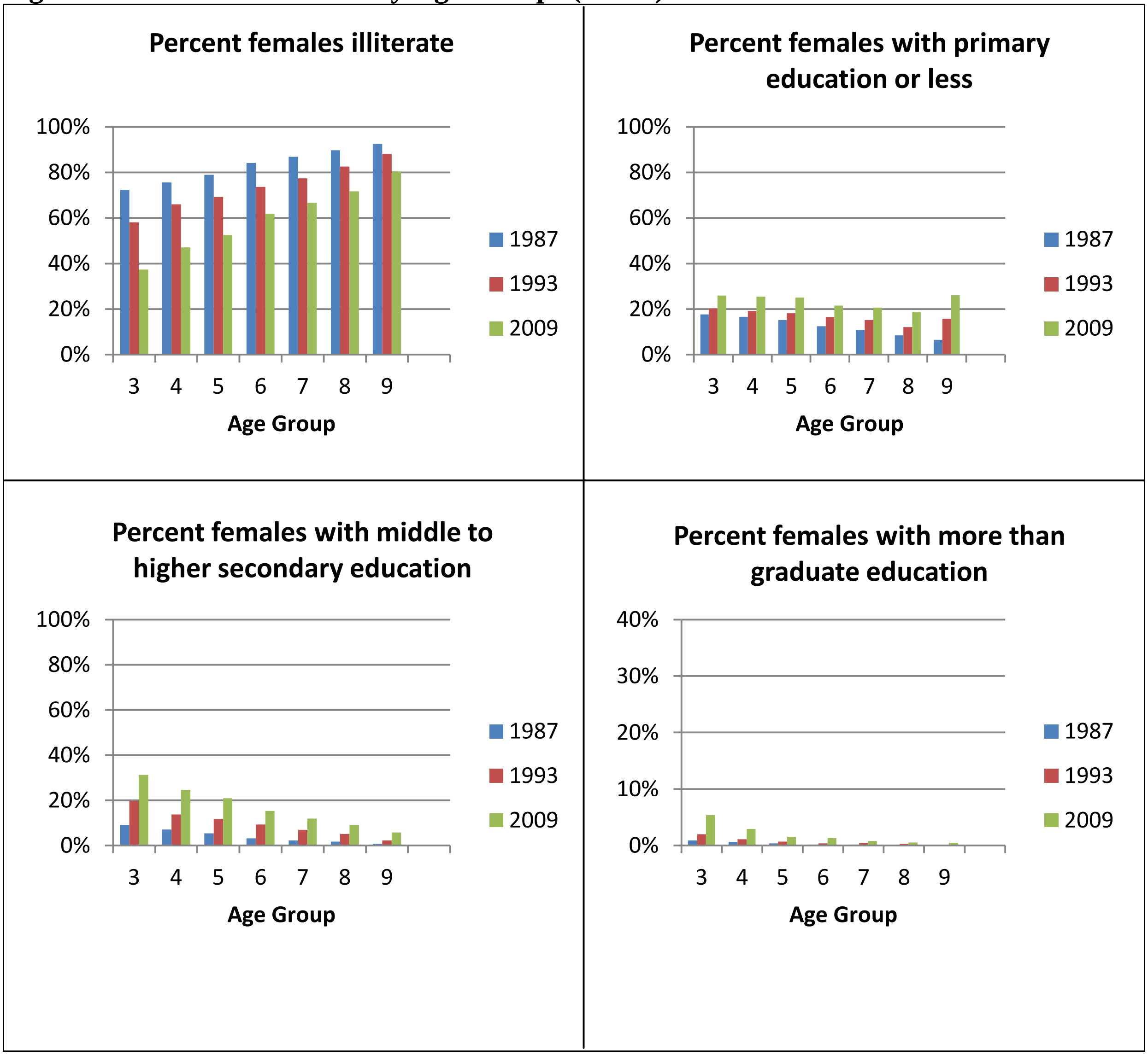
Source: NSS (1987, 1999, 2009) Employment and Unemployment Schedule (Author's calculations).
Note: The education categories are defined as: 1=(Illiterate) 2=(Literate less than primary) 3=(Primary) 4=(Middle) 5=(Higher Secondary) 6=(Graduate and above). The corresponding grade levels are: Primary=Grade 5, Middle=Grade 6-8, Higher Secondary=Grade 9-12, Graduate and above= College and above. The LFPR is calculated within each education category.

Figure 4.6 Female Domestic Work by Education (Rural)



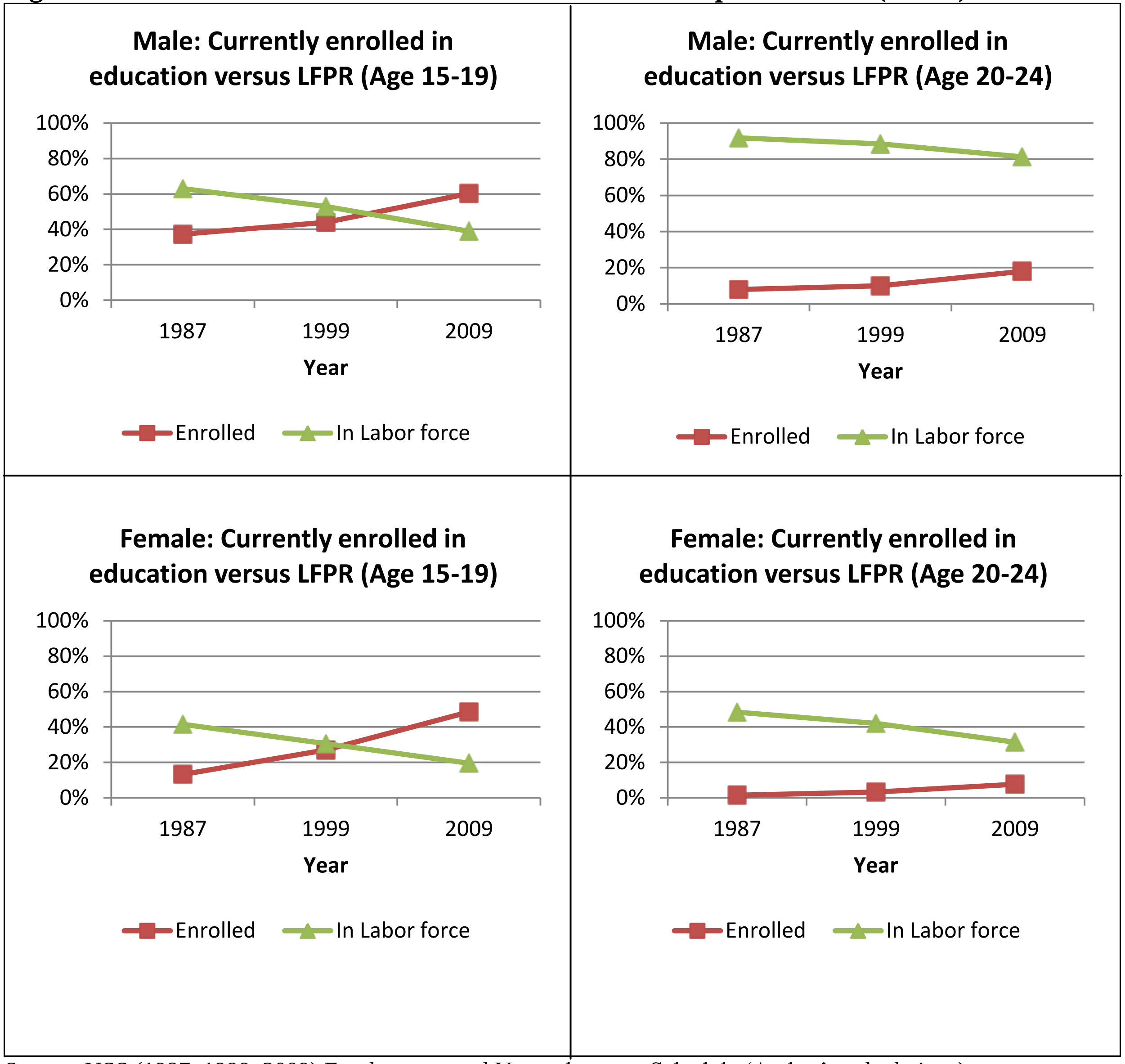
Source: NSS (1987, 1999, 2009) Employment and Unemployment Schedule (Author's calculations).
Note: The education categories are defined as: 1=(Illiterate) 2=(Literate less than primary) 3=(Primary) 4=(Middle) 5=(Higher Secondary) 6=(Graduate and above). The corresponding grade levels are: Primary=Grade 5, Middle=Grade 6-8, Higher Secondary=Grade 9-12, Graduate and above= College and above. The percentage females engaged in domestic work is calculated within each education category.

Figure 4.7 Female Education by Age Groups (Rural)



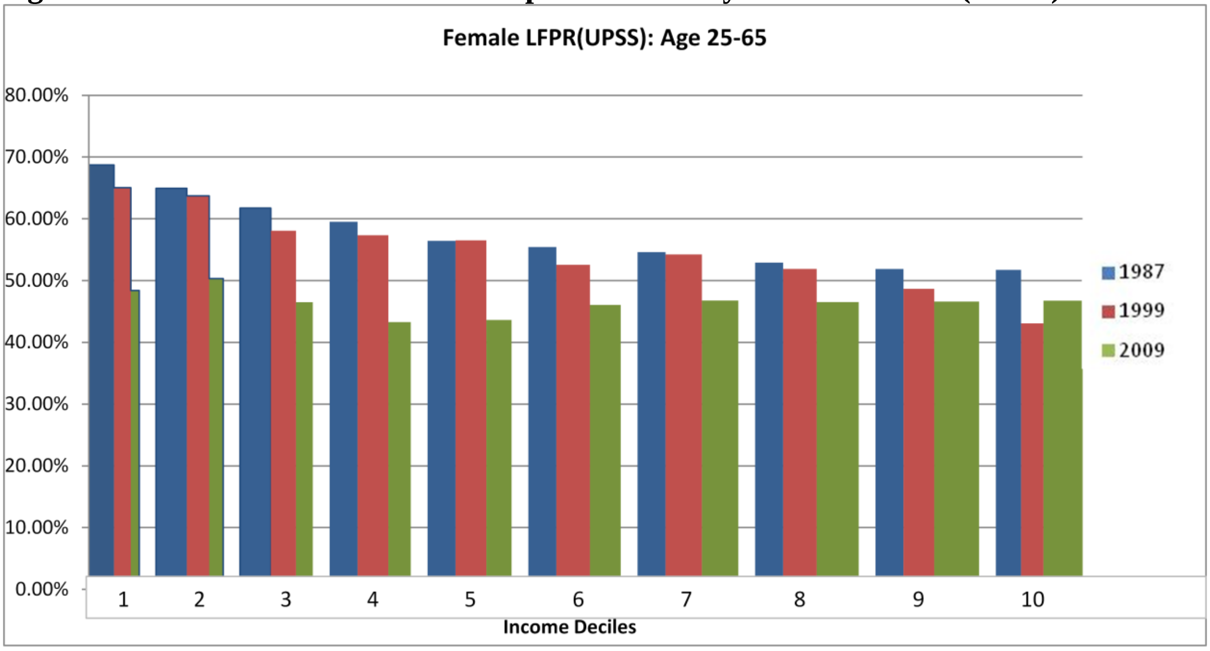
Source: NSS (1987, 1999, 2009) Employment and Unemployment Schedule (Author's calculations).
Note: The age groups are defined as: 1=(15-19) 2=(20-24) 3=(25-29) 4=(30-34) 5=(35-39) 6=(40-44) 7=(45-49) 8=(50-54) 9=(55-65). The percentages are calculated within each age category. The grade levels for education are: Primary=Grade 5, Middle=Grade 6-8, Higher Secondary=Grade 9-12, Graduate and above= College and above. The LFPR is calculated within each age category.

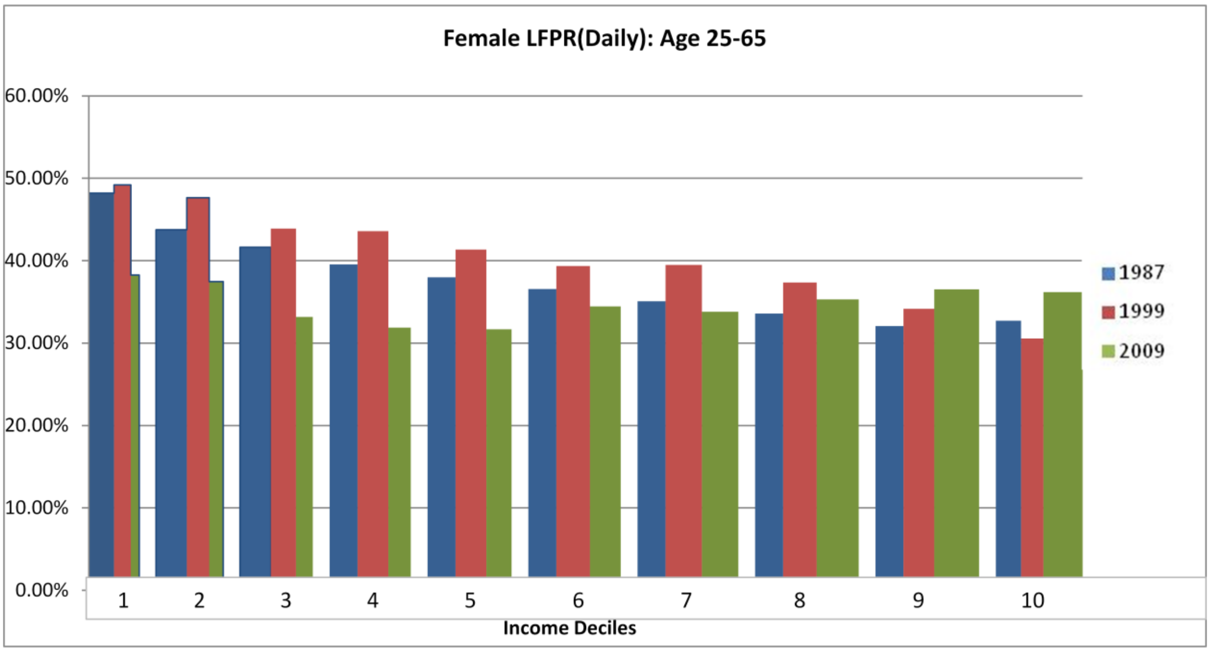
Figure 4.8 Education Enrollment and Labor Force Participation Rates (Rural)



Source: NSS (1987, 1999, 2009) Employment and Unemployment Schedule (Author's calculations).
Note: The labor force participation rates are estimated using the UPSS definition. The detailed definition is discussed in appendix 4.A.

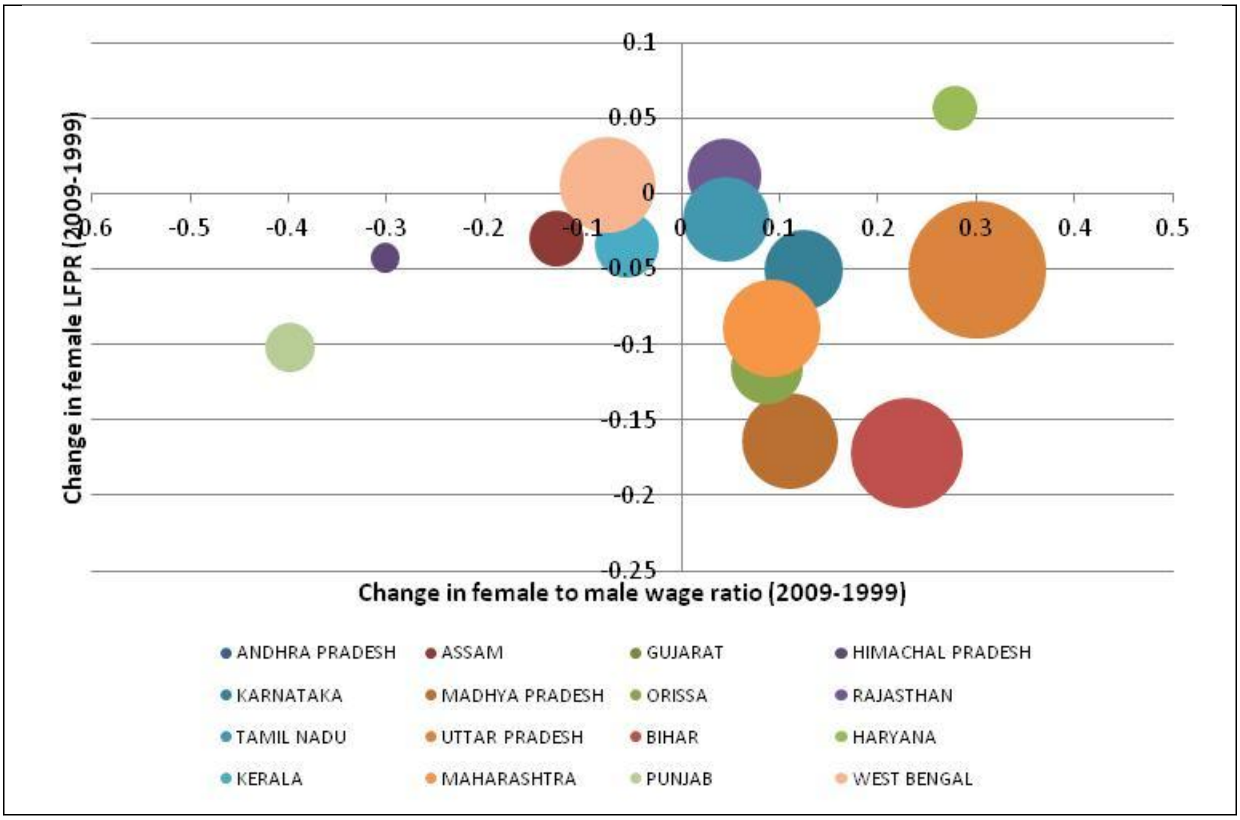
Figure 4.9 Female Labor Force Participation Rates by Income Deciles (Rural)





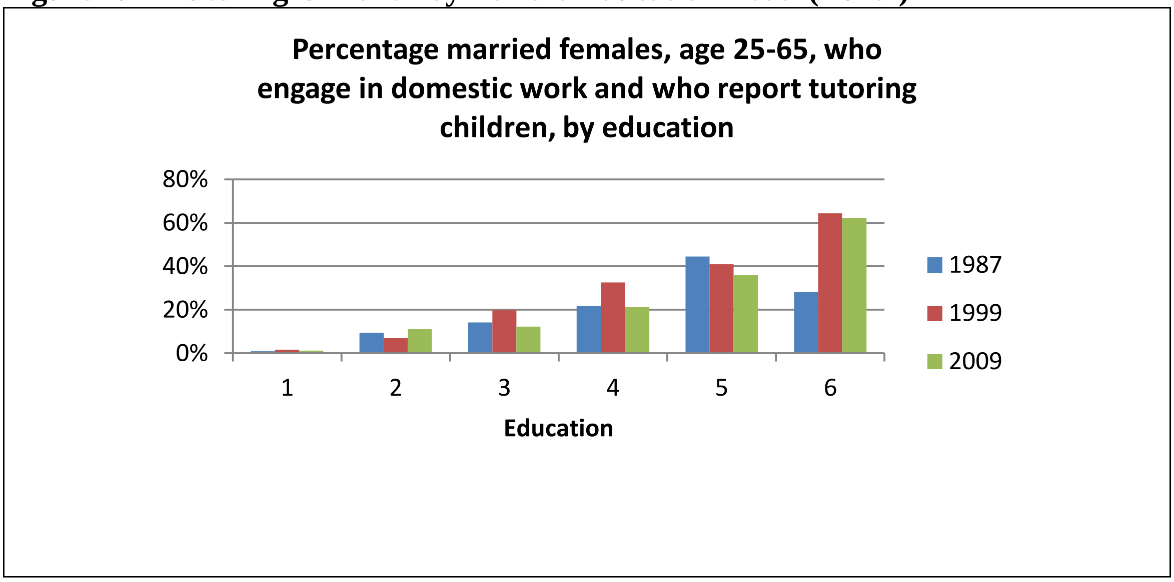
Source: NSS (1987, 1999, 2009) Employment and Unemployment Schedule (Author's calculations).
Note: The cut-offs of income deciles for the year 1987 are used and adjusted for cost of living for the years 1999 and 2009. The LFPR is calculated within each income decile. The width of the bars shows the proportion of females in the income group in that year.

Figure 4.10 Change in Female Labor Force Participation Rates and Relative Female Wage across States of India (2009-1999)



Source: NSS (1999, 2009) Employment and Unemployment Schedule (Author's calculations).
Note: The sample includes men and women in age group 25 to 65 in rural areas.

Figure 4.11 Tutoring Children by Female Education Level (Rural)



Source: NSS (1987, 1999, 2009) Employment and Unemployment Schedule (Author's calculations).
Note: The education categories are defined as: 1=(Illiterate) 2=(Literate less than primary) 3=(Primary) 4=(Middle) 5=(Higher Secondary) 6=(Graduate and above). The corresponding grade levels are: Primary=Grade 5, Middle=Grade 6-8, Higher Secondary=Grade 9-12, Graduate and above= College and above. The percentage is calculated within each education category. The sample of women is restricted to those who are spouses of married children of the household head and to the households which have children aged 4-15.

Table 4.1 Growth in Real Wage Rate between 1999-2009 (Rural)

	Overall (1)	Casual (2)	Salaried (3)
Male	1.9%	2.7%	1.1%
Female	3.3%	3.4%	1.3%

Source: NSS (1999, 2009) Employment and Unemployment Schedule (Author's calculations).

Table 4.2 Percentage Married Women Who Report Tutoring Children Among Those Who Primarily Engage in Domestic Work (Age 25-65)

Year		
1987	1999	2009
7.43	14.29	17.81

Source: NSS (1987, 1999, 2009) Employment and Unemployment Schedule (Author's calculations).
Note: The sample of women is restricted to those who are spouses of married children of the household head and to the households which have children aged 4-15.

Table 4.3 Variable Definition and Summary Statistics

Variable	Definition	Indicators	Mean		
Year -->			1987	1999	2009
Age group	Own Age	3. Age 25-29	23%	21%	19%
		4. Age 30-34	19%	19%	18%
		5. Age 35-39	16%	17%	18%
		6. Age 40-44	13%	13%	14%
		3. Age 45-49	11%	10%	11%
		3. Age 50-54	8%	8%	8%
		3. Age 55-65	10%	12%	13%
Education	Own Education	1. Illiterate	80%	71%	56%
		2. Less than primary	7%	9%	11%
		3. Primary	7%	8%	12%
		4. Middle	3%	7%	11%
		5. Secondary - High Secondary	2%	4%	8%
		6. Graduate and above	0%	1%	2%
Land Owned Household	Land owned by the household divided by the household size	1. less than 0.1 hectare	51%	60%	61%
		2. 0.1-0.2 hectare	16%	15%	14%
		3. 0.2-0.4 hectare	16%	13%	13%
		4. 0.4-0.6 hectare	7%	5%	6%
		5. More than 0.6 hectare	10%	6%	6%
Household Income	Monthly per capita consumption expenditure of the household	1. Income decile 1 (less than Rs 76)	9%	6%	3%
		2. Income decile 2 (Rs 76 Rs 93)	11%	8%	5%
		3. Income decile 3 (Rs 93 - Rs 107)	11%	9%	6%
		4. Income decile 4 (Rs 107 - Rs 121)	10%	11%	8%
		5. Income decile 5 (Rs 121 - Rs 135)	10%	11%	9%
		6. Income decile 6 (Rs 135 - Rs 153)	10%	12%	11%
		7. Income decile 7 (Rs 153 - Rs 177)	10%	13%	13%
		8. Income decile 8 (Rs 177 - Rs 212)	10%	12%	15%
		9. Income decile 9 (Rs 212 - Rs 281)	10%	11%	16%
		10. Income decile 10 (More than Rs 281)	9%	8%	14%
Max Male Education	Maximum Education of males in the household belonging to the age group 18-65	1. Illiterate	45%	37%	27%
		2. Less than primary	16%	14%	11%
		3. Primary	15%	13%	15%
		4. Middle	12%	16%	19%
		5. Secondary - High Secondary	10%	16%	21%
		6. Graduate and above	3%	5%	7%
Household size	Log of total number of members in household	Log household size	1.72	1.69	1.59
Male members	Number of adult males/Household size	Percentage male adults	0.48	0.48	0.49
Children	Number of children under age 5/Household size	Share of children under 5	0.15	0.13	0.10
Social group	Social group of household	Scheduled Caste (SC)	18%	21%	21%
		Scheduled Tribe (ST)	10%	10%	10%
		Others	71%	69%	69%
Religion	Religion of household	Hindu	85%	85%	85%
		Muslim	10%	10%	10%
		Christian	2%	2%	2%
		Others	3%	3%	3%

Source: NSS (1987, 1999, 2009) Employment and Unemployment Schedule (Author's calculations).
Note: The sample includes rural married women aged 25-65. The descriptive statistics have been estimated using sampling weights provided in the NSS. The cut-offs for income deciles for the year 1987 are used and adjusted for cost of living for the years 1999 and 2009.

Table 4.4 Determinants of Female Labor Force Participation (Daily Status)

Year -->	1987		1999		2009	
	Coefficient	S.E.	Coefficient	S.E.	Coefficient	S.E.
Age 30-34	0.171***	(0.042)	0.222***	(0.044)	0.196***	(0.067)
Age 35-39	0.260***	(0.046)	0.391***	(0.048)	0.480***	(0.075)
Age 40-44	0.211***	(0.051)	0.405***	(0.054)	0.416***	(0.078)
Age 45-49	0.069	(0.052)	0.249***	(0.055)	0.348***	(0.085)
Age 50-54	-0.262***	(0.057)	-0.012	(0.061)	0.202**	(0.092)
Age 55-65	-0.754***	(0.053)	-0.612***	(0.054)	-0.375***	(0.084)
Less than Primary Education	-0.489***	(0.052)	-0.431***	(0.051)	-0.211***	(0.076)
Primary Education	-0.616***	(0.054)	-0.561***	(0.057)	-0.204***	(0.073)
Middle Education	-0.757***	(0.069)	-0.765***	(0.059)	-0.502***	(0.075)
Higher Secondary Education	0.021	(0.109)	-0.504***	(0.081)	-0.619***	(0.089)
Graduate and above Education	0.566**	(0.226)	0.796***	(0.190)	0.405**	(0.184)
Male: Less than Primary Education	-0.273***	(0.043)	-0.310***	(0.047)	-0.144*	(0.081)
Male: Primary Education	-0.380***	(0.042)	-0.253***	(0.049)	-0.147**	(0.073)
Male: Middle Education	-0.578***	(0.046)	-0.467***	(0.045)	-0.239***	(0.070)
Male: Higher Secondary Education	-0.795***	(0.051)	-0.641***	(0.050)	-0.503***	(0.071)
Male: Graduate and above Education	-0.939***	(0.074)	-0.852***	(0.075)	-0.786***	(0.103)
Income decile 2	-0.005	(0.066)	-0.032	(0.089)	0.064	(0.150)
Income decile 3	-0.089	(0.066)	-0.205**	(0.085)	0.019	(0.149)
Income decile 4	-0.161**	(0.065)	-0.199**	(0.085)	-0.126	(0.139)
Income decile 5	-0.159**	(0.067)	-0.288***	(0.086)	-0.194	(0.141)
Income decile 6	-0.287***	(0.067)	-0.400***	(0.087)	-0.166	(0.139)
Income decile 7	-0.323***	(0.067)	-0.445***	(0.085)	-0.263*	(0.139)
Income decile 8	-0.364***	(0.069)	-0.528***	(0.088)	-0.242*	(0.140)
Income decile 9	-0.461***	(0.071)	-0.696***	(0.093)	-0.199	(0.143)
Income decile 10	-0.503***	(0.074)	-0.853***	(0.099)	-0.350**	(0.153)
Land per capita (0.1-0.2 hectare)	-0.017	(0.038)	0.210***	(0.040)	0.200***	(0.062)
Land per capita (0.2-0.4 hectare)	0.102**	(0.040)	0.399***	(0.044)	0.319***	(0.069)
Land per capita (0.4-0.6 hectare)	0.070	(0.053)	0.177***	(0.062)	0.157	(0.098)
Land per capita (More than 0.6 hectare)	0.107**	(0.051)	0.262***	(0.061)	0.222**	(0.099)
Log household size	-0.291***	(0.034)	-0.132***	(0.036)	-0.036	(0.056)
Percentage male adults	-1.346***	(0.117)	-1.116***	(0.122)	-1.427***	(0.182)
Share of children under 5	-0.657***	(0.105)	-0.669***	(0.113)	-0.149	(0.167)
Caste: ST	0.133**	(0.065)	0.353***	(0.064)	0.343***	(0.100)
Caste: Others	-0.444***	(0.040)	-0.338***	(0.039)	-0.150***	(0.057)
Religion: Muslim	-0.567***	(0.048)	-0.643***	(0.053)	-0.604***	(0.074)
Religion: Christian	0.122	(0.111)	-0.129	(0.123)	0.043	(0.158)
Religion: Others	0.196**	(0.094)	-0.089	(0.104)	-0.045	(0.160)
Constant	4.727***	(0.283)	6.205***	(0.300)	1.537***	(0.208)
District fixed effects	Yes		Yes		Yes	
Observations	66,944		62,713		54,257	
R-squared	0.272		0.314		0.328	

Note: The dependent variable is the number of days women work in labor force in a reference week. Coefficients from a linear regression are shown and robust standard errors are reported in the parenthesis. The sample includes rural married women aged 25-65. The regressions have been weighted using the sampling weights in the NSS. Robust standard errors are in parenthesis; ***, ** and * indicate significance at the 1, 5 and 10% levels respectively.

Table 4.5 Determinants of Female Labor Force Participation (UPSS)

Year -->	1987		1999		2009	
	Coefficient	S.E.	Coefficient	S.E.	Coefficient	S.E.
Age 30-34	0.0359***	(0.006)	0.0406***	(0.007)	0.0313***	(0.011)
Age 35-39	0.0391***	(0.007)	0.0529***	(0.007)	0.0631***	(0.013)
Age 40-44	0.0423***	(0.007)	0.0470***	(0.008)	0.0471***	(0.013)
Age 45-49	0.0117	(0.007)	0.0254***	(0.009)	0.0224	(0.014)
Age 50-54	-0.0458***	(0.008)	-0.0250***	(0.009)	6.99e-05	(0.015)
Age 55-65	-0.154***	(0.008)	-0.141***	(0.009)	-0.107***	(0.014)
Less than Primary Education	-0.0781***	(0.008)	-0.0661***	(0.008)	-0.0464***	(0.012)
Primary Education	-0.108***	(0.008)	-0.0845***	(0.008)	-0.0429***	(0.012)
Middle Education	-0.126***	(0.012)	-0.125***	(0.011)	-0.0994***	(0.012)
Higher Secondary Education	-0.0372**	(0.016)	-0.107***	(0.012)	-0.139***	(0.014)
Graduate and above Education	-0.0237	(0.032)	0.0390	(0.025)	0.000933	(0.027)
Male: Less than Primary Education	-0.0411***	(0.006)	-0.0461***	(0.007)	-0.0254*	(0.013)
Male: Primary Education	-0.0529***	(0.006)	-0.0484***	(0.007)	-0.0215*	(0.012)
Male: Middle Education	-0.0937***	(0.007)	-0.0924***	(0.007)	-0.0453***	(0.012)
Male: Higher Secondary Education	-0.136***	(0.008)	-0.120***	(0.008)	-0.0885***	(0.012)
Male: Graduate and above	-0.187***	(0.014)	-0.168***	(0.013)	-0.167***	(0.016)
Income decile 2	0.00283	(0.009)	0.00599	(0.014)	0.0394	(0.026)
Income decile 3	-0.0159*	(0.009)	-0.0295**	(0.013)	0.0354	(0.025)
Income decile 4	-0.0231***	(0.009)	-0.0299**	(0.013)	-0.0122	(0.024)
Income decile 5	-0.0425***	(0.009)	-0.0392***	(0.013)	-0.0164	(0.024)
Income decile 6	-0.0547***	(0.009)	-0.0743***	(0.013)	-0.0275	(0.023)
Income decile 7	-0.0590***	(0.009)	-0.0660***	(0.013)	-0.0269	(0.023)
Income decile 8	-0.0701***	(0.01)	-0.0792***	(0.013)	-0.0328	(0.023)
Income decile 9	-0.0852***	(0.01)	-0.101***	(0.014)	-0.0349	(0.023)
Income decile 10	-0.0966***	(0.011)	-0.135***	(0.015)	-0.0378	(0.025)
Land per capita (0.1-0.2 hectare)	0.0196***	(0.006)	0.0427***	(0.006)	0.0326***	(0.01)
Land per capita (0.2-0.4 hectare)	0.0211***	(0.006)	0.0604***	(0.007)	0.0562***	(0.011)
Land per capita (0.4-0.6 hectare)	0.00941	(0.007)	0.0440***	(0.009)	0.0316**	(0.015)
Land per capita (More than 0.6 hectare)	0.0216***	(0.007)	0.0438***	(0.009)	0.0409***	(0.015)
Log household size	-0.0649***	(0.005)	-0.0433***	(0.006)	-0.0216**	(0.009)
Percentage male adults	-0.157***	(0.017)	-0.147***	(0.019)	-0.196***	(0.029)
Share of children under 5	-0.0419***	(0.015)	-0.0595***	(0.017)	-0.0196	(0.029)
Caste: ST	0.0174**	(0.009)	0.0530***	(0.01)	0.0468***	(0.015)
Caste: Others	-0.0786***	(0.006)	-0.0745***	(0.006)	-0.0370***	(0.009)
Religion: Muslim	-0.103***	(0.008)	-0.115***	(0.01)	-0.127***	(0.014)
Religion: Christian	0.0292**	(0.014)	-0.00798	(0.015)	0.0401*	(0.021)
Religion: Others	0.0225*	(0.014)	-0.0125	(0.016)	-0.00785	(0.025)
District fixed effects	Yes		Yes		Yes	
Observations	67,062		62,480		53,311	

Note: The dependent variable is dichotomous, taking a value equal to one when a female is in the labor force. The coefficients show the average marginal effects from the logistic regression and robust standard errors are reported in the parenthesis. The sample includes rural married women aged 25-65. The regressions have been weighted using the sampling weights in the NSS. Robust standard errors are in parenthesis; ***, ** and * indicate significance at the 1, 5 and 10% levels respectively.

Table 4.6 Determinants of Female Participation in Domestic Work (Daily Status)

Year -->	1987		1999		2009	
	Coefficient	S.E.	Coefficient	S.E.	Coefficient	S.E.
Age 30-34	-0.145***	(0.042)	-0.197***	(0.044)	-0.187***	(0.068)
Age 35-39	-0.254***	(0.046)	-0.367***	(0.048)	-0.471***	(0.077)
Age 40-44	-0.221***	(0.051)	-0.394***	(0.055)	-0.437***	(0.080)
Age 45-49	-0.098*	(0.052)	-0.272***	(0.056)	-0.378***	(0.086)
Age 50-54	0.073	(0.057)	-0.114*	(0.062)	-0.285***	(0.095)
Age 55-65	-0.174***	(0.057)	-0.184***	(0.059)	-0.269***	(0.092)
Less than Primary Education	0.529***	(0.053)	0.471***	(0.053)	0.291***	(0.079)
Primary Education	0.671***	(0.055)	0.624***	(0.058)	0.278***	(0.074)
Middle Education	0.822***	(0.071)	0.837***	(0.060)	0.568***	(0.078)
Higher Secondary Education	0.020	(0.110)	0.545***	(0.083)	0.686***	(0.093)
Graduate and above Education	-0.639***	(0.235)	-0.801***	(0.193)	-0.373**	(0.189)
Male: Less than Primary Education	0.272***	(0.044)	0.278***	(0.048)	0.095	(0.083)
Male: Primary Education	0.348***	(0.042)	0.224***	(0.050)	0.065	(0.075)
Male: Middle Education	0.577***	(0.047)	0.448***	(0.046)	0.188***	(0.071)
Male: Higher Secondary Education	0.767***	(0.053)	0.591***	(0.051)	0.423***	(0.075)
Male: Graduate and above Education	0.924***	(0.079)	0.788***	(0.078)	0.700***	(0.111)
Income decile 2	0.063	(0.066)	0.045	(0.090)	0.061	(0.156)
Income decile 3	0.157**	(0.066)	0.238***	(0.086)	0.054	(0.155)
Income decile 4	0.203***	(0.065)	0.233***	(0.086)	0.201	(0.144)
Income decile 5	0.253***	(0.066)	0.352***	(0.087)	0.304**	(0.147)
Income decile 6	0.356***	(0.067)	0.462***	(0.087)	0.262*	(0.144)
Income decile 7	0.390***	(0.068)	0.499***	(0.086)	0.401***	(0.144)
Income decile 8	0.448***	(0.069)	0.601***	(0.089)	0.357**	(0.145)
Income decile 9	0.556***	(0.071)	0.757***	(0.095)	0.316**	(0.150)
Income decile 10	0.614***	(0.074)	0.919***	(0.101)	0.416***	(0.159)
Land per capita (0.1-0.2 hectare)	0.030	(0.039)	-0.195***	(0.041)	-0.193***	(0.064)
Land per capita (0.2-0.4 hectare)	-0.058	(0.040)	-0.357***	(0.045)	-0.286***	(0.071)
Land per capita (0.4-0.6 hectare)	-0.032	(0.053)	-0.149**	(0.063)	-0.116	(0.101)
Land per capita (More than 0.6 hectare)	-0.027	(0.051)	-0.195***	(0.062)	-0.160	(0.100)
Log household size	0.198***	(0.034)	0.035	(0.037)	-0.094	(0.058)
Percentage male adults	1.755***	(0.118)	1.580***	(0.126)	1.613***	(0.188)
Share of children under 5	0.657***	(0.106)	0.639***	(0.115)	0.136	(0.173)
Caste: ST	-0.113*	(0.065)	-0.305***	(0.065)	-0.345***	(0.102)
Caste: Others	0.471***	(0.040)	0.361***	(0.040)	0.140**	(0.059)
Religion: Muslim	0.582***	(0.049)	0.630***	(0.056)	0.612***	(0.076)
Religion: Christian	-0.126	(0.111)	0.139	(0.126)	-0.134	(0.162)
Religion: Others	-0.177*	(0.094)	0.062	(0.104)	0.037	(0.164)
Constant	1.938***	(0.271)	0.495*	(0.293)	5.509***	(0.212)
District fixed effects	Yes		Yes		Yes	
Observations	66,944		62,713		54,257	
R-squared	0.274		0.304		0.315	

Note: The dependent variable is the number of days women are engaged in domestic work in a reference week. Coefficients from a linear regression are shown and robust standard errors are reported in the parenthesis. The sample includes rural married women aged 25-65. The regressions have been weighted using the sampling weights in the NSS data. Robust standard errors are in parenthesis; ***, ** and * indicate significance at the 1, 5 and 10% levels respectively.

Table 4.7 Blinder-Oaxaca Decomposition for Change in Female LFPR (Daily Status)

Change in FLFPR (1999-2009)	(1)	(2)	(3)	(4)	(5)
Difference in predicted LFPR	0.43	0.43	0.43	0.45	0.45
Explained proportion using 1999 returns	35.8%	38.2%	55.0%	45.0%	58.1%
Explained proportion using 2009 returns	11.8%	14.4%	12.9%	18.6%	15.8%
Characteristics Included					
Age group	Yes	Yes	Yes	Yes	Yes
Education	Yes	Yes	Yes	Yes	Yes
Land Owned Household		Yes	Yes	Yes	Yes
Household Income			Yes		Yes
Max Male Education				Yes	Yes
Variable contribution to the Proportion of explained variation (1999 returns)					
Age group	3%	4%	1%	1%	0%
Education	97%	95%	50%	41%	23%
Land Owned Household		2%	1%	2%	1%
Household Income			48%		37%
Max Male Education				56%	39%
Variable contribution to the Proportion of explained variation (2009 returns)					
Age group	-11%	-6%	-6%	-8%	-8%
Education	111%	103%	122%	23%	34%
Land Owned Household		3%	3%	3%	4%
Household Income			-20%		-28%
Max Male Education				82%	98%

Note: The dependent variable is a continuous variable defined as the number of days working or seeking work in the past week. The sample includes rural married women aged 25-65. The analysis incorporates the sampling weights in NSS.

Table 4.8 Blinder-Oaxaca Decomposition for Change in Female LFPR (UPSS)

Panel A					
Change in FLFPR (1987-2009)	(1)	(2)	(3)	(4)	(5)
Difference in predicted LFPR	0.12	0.12	0.12	0.12	0.13
Explained proportion using 1987 returns	42.1%	47.4%	56.1%	64.5%	69.5%
Explained proportion using 2009 returns	22.0%	33.0%	31.8%	40.2%	38.3%
Characteristics Included					
Age group	Yes	Yes	Yes	Yes	Yes
Education	Yes	Yes	Yes	Yes	Yes
Land Owned Household		Yes	Yes	Yes	Yes
Household Income			Yes		Yes
Max Male Education				Yes	Yes
Variable contribution to the Proportion of explained variation (1987 returns)					
Age group	2%	3%	2%	1%	1%
Education	98%	89%	62%	26%	19%
Land Owned Household		9%	10%	9%	10%
Household Income			26%		15%
Max Male Education				65%	56%
Variable contribution to the Proportion of explained variation (2009 returns)					
Age group	-1%	0%	1%	-1%	-1%
Education	101%	71%	76%	33%	37%
Land Owned Household		29%	29%	23%	23%
Household Income			-5%		-9%
Max Male Education				46%	49%

	Panel B				
Change in FLFPR (1999-2009)	(1)	(2)	(3)	(4)	(5)
Difference in predicted LFPR	0.09	0.09	0.09	0.09	0.09
Explained proportion using 1999 returns	33.4%	35.5%	46.4%	41.7%	49.9%
Explained proportion using 2009 returns	16.7%	19.2%	17.0%	21.9%	18.6%
Characteristics Included					
Age group	Yes	Yes	Yes	Yes	Yes
Education	Yes	Yes	Yes	Yes	Yes
Land Owned Household		Yes	Yes	Yes	Yes
Household Income			Yes		Yes
Max Male Education				Yes	Yes
Variable contribution to the Proportion of explained variation (1999 returns)					
Age group	-1%	-1%	0%	-1%	-2%
Education	101%	99%	62%	47%	32%
Land Owned Household		2%	2%	1%	2%
Household Income			36%		27%
Max Male Education				53%	42%
Variable contribution to the Proportion of explained variation (2009 returns)					
Age group	-8%	-7%	-5%	-6%	-7%
Education	108%	100%	115%	45%	59%
Land Owned Household		7%	9%	2%	3%
Household Income			-18%		-28%
Max Male Education				60%	73%

Note: The dependent variable is a dichotomous variable. A female is defined to be in labor force if she is working or seeking work as her primary or subsidiary activity. The sample includes rural married women aged 25-65. The analysis incorporates the sampling weights in NSS.

Table 4.9 DFL Decomposition for Change in Female LFPR (Daily Status)

Change in FLFPR (1999-2009)	(1)	(2)	(3)	(4)	(5)	(6)
Difference in LFPR	0.43	0.43	0.43	0.43	0.43	0.43
Explained proportion using 1999 returns	35.8%	38.3%	55.3%	47.9%	62.3%	64.4%
Explained proportion using 2009 returns	10.9%	14.1%	12.2%	14.0%	10.5%	11.1%
Characteristics Included						
Age group	Yes	Yes	Yes	Yes	Yes	Yes
Education	Yes	Yes	Yes	Yes	Yes	Yes
Land Owned Household		Yes	Yes	Yes	Yes	Yes
Household Income			Yes		Yes	Yes
Max Male Education				Yes	Yes	Yes

Note: Employment is captured as a continuous variable defined as the number of days working or seeking work in the past week. The sample includes rural married women aged 25-65. The analysis incorporates sampling weights in NSS. The specification in column (6) includes two-way interactions between all the variables when estimating the re-weighting function e.g. interaction between age-group and education, age-group and land owned, age-group and income, age-group and male education, education and land owned and so on.

Table 4.10 DFL Decomposition for Change in Female LFPR (UPSS)

Panel A

Change in FLFPR (1987-2009)	(1)	(2)	(3)	(4)	(5)	(6)
Difference in LFPR	0.12	0.12	0.12	0.12	0.12	0.12
Explained proportion using 1987 returns	40.4%	43.9%	53.0%	65.2%	71.8%	65.8%
Explained proportion using 2009 returns	20.9%	33.8%	32.2%	38.0%	36.4%	37.7%

Panel B

Change in FLFPR (1999-2009)	(1)	(2)	(3)	(4)	(5)	(6)
Difference in LFPR	0.09	0.09	0.09	0.09	0.09	0.09
Explained proportion using 1999 returns	33.5%	35.8%	46.7%	44.2%	53.2%	53.4%
Explained proportion using 2009 returns	16.1%	18.8%	15.8%	17.9%	13.7%	12.8%
Characteristics Included						
Age group	Yes	Yes	Yes	Yes	Yes	Yes
Education	Yes	Yes	Yes	Yes	Yes	Yes
Land Owned Household		Yes	Yes	Yes	Yes	Yes
Household Income			Yes		Yes	Yes
Max Male Education				Yes	Yes	Yes

Note: A female is defined to be in the labor force if she is working or seeking work as her primary or subsidiary activity. The sample includes rural married women aged 25-65. The analysis incorporates the sampling weights in NSS. The specification in column (6) includes two-way interactions between all the variables when estimating the re-weighting function e.g. interaction between age-group and education, age-group and land owned, age-group and income, age-group and male education, education and land owned and so on.

Table 4.11 Blinder-Oaxaca Decomposition for Change in Domestic Work (Daily Status)

Change in Domestic Work (1999-2009)	(1)	(2)
Difference in predicted Domestic Work	-0.49	-0.51
Explained proportion using 1999 returns	48.6%	51.6%
Explained proportion using 2009 returns	12.3%	14.0%
Characteristics Included		
Age group	Yes	Yes
Education	Yes	Yes
Land Owned Household	Yes	Yes
Household Income	Yes	Yes
Max Male Education		Yes
Variable contribution to the Proportion of explained variation (1999 returns)		
Age group	-3%	-3%
Education	50%	24%
Land Owned Household	1%	1%
Household Income	52%	42%
Max Male Education		36%
Variable contribution to the Proportion of explained variation (2009 returns)		
Age group	-18%	-18%
Education	116%	52%
Land Owned Household	3%	4%
Household Income	-1%	-13%
Max Male Education		75%

Note: Domestic work is captured as a continuous variable defined as the number of days spent in domestic work in the past week. The sample includes rural married women aged 25-65. The analysis incorporates the sampling weights in NSS.

Appendices for Chapter 4

4.A Definition of Labor Force Participation Rates

The National Sample Survey captures various measures of labor force participation rates in India. There are three reference periods for this survey (i) one year, (ii) one week and (iii) each day of the previous week. This chapter employs two definitions: Usual Principal and Subsidiary Status (UPSS) and Daily Status. Their description is provided below.

Usual Principal and Subsidiary Status (UPSS):

This definition is based on the reference period of a year. It relates to the activity status of a person during the reference period of 365 days preceding the date of survey. The activity status on which a person spent relatively longer time (major time criterion) during the 365 days preceding the date of survey is considered the Usual Principal Activity Status of the person. Persons are first categorized as those in the labor force and those out of the labor force depending on the major time spent during the 365 days preceding the date of survey. For persons belonging to the labor force, the broad activity status of either 'working' (employed) or ‘not working but seeking and/or available for work’ (unemployed) is then ascertained based on the time criterion.

After determining the principal status, the economic activity on which a person spent 30 days or more during the reference period of 365 days preceding the date of survey is recorded as the Subsidiary Economic Activity Status of a person. In case of multiple subsidiary economic activities, the major activity and status based on the relatively longer time spent criterion is considered.

UPSS is defined on the basis of Usual Principal Activity Status and Subsidiary Economic Activity Status of a person. If a person is defined to be in the labor force in either the principal activity status or the subsidiary activity status then she is defined to be in the labor force according to the UPSS.

Daily Status:

This definition is determined on the basis of each day of the previous week (each day of the seven days preceding the date of survey). The different activities are identified and recorded in terms of 'status' and 'industry' codes for persons in urban areas and 'status', 'industry' and 'operation' codes for persons in rural areas. The time intensity is measured in half-day units. Each day of the reference week is comprised of either two 'half days' or a 'full day' for assigning the activity status.

If a person was engaged in more than one of the economic activities for 4 hours or more on a day, she would be assigned two economic activities out of the different economic activities on which she devoted relatively longer time on the reference day. In such cases, one 'half day' work will be considered for each of those two economic activities (for each of those two activities, the intensity will be 0.5). If a person had worked or was seeking work for 4 hours or more during the day then she is considered to be in the labor force for the entire day (1). If a person had worked or was seeking work for 1 hour or more but less than 4 hours, then she is considered to be in labor force for half-day (0.5). The total number of days the person was in the labor force on each day of the reference week is then summed up to arrive at the daily status employment.

4.B Definition of Domestic Work

Domestic work in the NSS includes domestic chores and not-for-wages collection of goods (vegetables, roots, firewood, cattle feed, etc.), sewing, tailoring, weaving, etc. for household use. This chapter employs two definitions: Primary Status and Daily Status. Their description is provided below.

Primary Status:

This definition is based on the reference period of a year. It relates to the activity status of a person during the reference period of 365 days preceding the date of survey. If a person reports spending majority of the time in the previous one year in domestic work activities then she is defined to be engaged in domestic work according to the Primary Status definition.

Daily Status:

The total number of days the person spent engaged in domestic work activities on each day of the reference week is summed up to arrive at the daily status measure of domestic work. As in the case of labor force participation measurement, each day of the reference week is comprised of either two 'half days' or a 'full day' for assigning the domestic work activity status. If a person was engaged in domestic work for 4 hours or more during the day then she is considered to be involved in domestic work for the entire day (1). If a person engaged in domestic work for 1 hour or more but less than 4 hours, then she is considered to be involved in domestic work for half-day (0.5).

4.C Blinder-Oaxaca Decomposition with Additional Controls

Table 4.C.1 Blinder-Oaxaca Decomposition for Change in Female LFPR with Additional Controls

Panel A		
Change in FLFPR (1987-2009)	UPSS	
	(1)	(2)
Difference in predicted LFPR	0.13	0.13
Explained proportion using 1987 returns	69.5%	54.4%
Explained proportion using 2009 returns	38.3%	31.6%
Characteristics Included		
Age group	Yes	Yes
Education	Yes	Yes
Land Owned Household	Yes	Yes
Household Income	Yes	Yes
Max Male Education	Yes	Yes
Social group		Yes
Religion		Yes
Percentage male adults		Yes
Share of children under 5		Yes
Household size		Yes
Variable contribution to the Proportion of explained variation (1987 returns)		
Age group	1%	1%
Education	19%	30%
Land Owned Household	10%	12%
Household Income	15%	21%
Max Male Education	56%	53%
Social group		-4%
Religion		7%
Percentage male adults		1%
Share of children under 5		-3%
Household size		-16%
Variable contribution to the Proportion of explained variation (2009 returns)		
Age group	-1%	-1%
Education	37%	49%
Land Owned Household	23%	24%
Household Income	-9%	-16%
Max Male Education	49%	48%
Social group		-3%
Religion		13%
Percentage male adults		2%
Share of children under 5		-3%
Household size		-11%

Panel B				
Change in FLFPR (1999-2009)	UPSS		Daily	
	(1)	(2)	(3)	(4)
Difference in predicted LFPR	0.09	0.09	0.45	0.45
Explained proportion using 1999 returns	49.9%	39.6%	58.1%	45.2%
Explained proportion using 2009 returns	18.6%	15.7%	15.8%	9.3%
Characteristics Included				
Age group	Yes	Yes	Yes	Yes
Education	Yes	Yes	Yes	Yes
Land Owned Household	Yes	Yes	Yes	Yes
Household Income	Yes	Yes	Yes	Yes
Max Male Education	Yes	Yes	Yes	Yes
Social group		Yes		Yes
Religion		Yes		Yes
Percentage male adults		Yes		Yes
Share of children under 5		Yes		Yes
Household size		Yes		Yes
Variable contribution to the Proportion of explained variation (1999 returns)				
Age group	-2%	-2%	0%	1%
Education	32%	41%	23%	33%
Land Owned Household	2%	4%	1%	1%
Household Income	27%	34%	37%	50%
Max Male Education	42%	34%	39%	39%
Social group		-5%		0%
Religion		17%		4%
Percentage male adults		1%		2%
Share of children under 5		-5%		-14%
Household size		-19%		-17%
Variable contribution to the Proportion of explained variation (2009 returns)				
Age group	-7%	-9%	-8%	-12%
Education	59%	76%	34%	83%
Land Owned Household	3%	4%	4%	6%
Household Income	-28%	-40%	-28%	-55%
Max Male Education	73%	60%	98%	143%
Social group		-2%		0%
Religion		37%		20%
Percentage male adults		0%		9%
Share of children under 5		-5%		-32%
Household size		-22%		-61%

Note: A female is defined to be employed using the UPSS definition in Columns (1)-(2). Employment is defined as the number of days worked or seeking work in past week in Columns (3)-(4). The sample includes rural married women aged 25-65. The analysis incorporates the sampling weights in NSS.

Bibliography

Acemoglu, Daron, Autor, David H., and David Lyle. 2004. “Women, War, and Wages: The Effect of Female Labor Supply on the Wage Structure at Mid-Century.”*Journal of Political Economy* 112, no. 3:497-551.

Adhvaryu, Achyuta, Char, A.V., and Siddharth Sharma. 2013. “Firing Costs and Flexibility: Evidence from Firms’ Employment Responses to Shocks in India.” *Review of Economics and Statistics* 95, no. 3:725-740.

Aggarwal, Bina. 1986. “Women, Poverty and Agricultural Growth in India.” *The Journal of Peasant Studies* 13, no. 4:165-220.

-------------------. 1994. *A Field of One’s Own: Gender and Land rights in South Asia*. Cambridge: Cambridge University Press.

Agrawal, Tushar, and Reeve Vanneman. 2014. “Gender-based Wage Differentials in India: Evidence Using a Matching Comparisons Method.” http://paa2014.princeton.edu/papers/140790 (accessed on April 2, 2015).

Aldrich, Mark. 1989. "The gender gap in earnings during World War II: New evidence." *Industrial and Labor Relations Review* 42, no. 3:415-429.

Aller, Ricardo Aláez, and Miren Ullíbarri Arce. 2001. “Explaining the reduction in the gender wage gap during the 1990-94 economic recession in Spain.” *Atlantic Economic Journal* 29, no. 1:63–74.

Anderson, Siwan. 2011. “Caste as an Impediment to Trade.” *American Economic Journal: Applied Economics* 3, no.1:239-263.

Angrist, Joshua D., and Jorn-Steffen Pischke. 2008. *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press.

Aromolaran, Adebayo B. 2004. “Female Schooling, Non-Market Productivity and Labor Market Participation in Nigeria.” Center Discussion Paper No 879, Yale University, USA.

Badiani, Reena, and Abla Safir. 2008. “Coping with Aggregate Shocks : Temporary Migration and Other Labor Responses to Climatic Shocks in Rural India”. http://www.webmeets.com/files/papers/ESWC/2010/2531/Temporary%20Migration%20India%20Badiani%20Safir.pdf (accessed on December 20, 2011).

Bagchi, Deipica, and Saraswati Raju. 1993. *Women and work in South Asia*. London: Routledge.

Banerjee, Abhijit, and Rohini Somanathan. 2007. “The political economy of public goods: Some evidence from India.” *Journal of Development Economics* 82: 287–314.

Bardhan, Kalpana. 1984. “Work Patterns and Social Differentiation: Rural Women of West Bengal.” In *Contractual Arrangements, Employment and Wages in Rural Labor Markets in Asia*, ed. Hans P. Binswanger and Mark R. Rosenzweig. New Haven: Yale University Press, pp. 184-208.

Bardhan, Pranab. 1974. “On Life and Death Questions.” *Economic and Political Weekly* 9, no. 32/34:1293-1304.

Bardhan, Pranab. 1984. “Determinants of supply and demand for labor in a poor agrarian economy: an analysis of household survey data from rural West Bengal.” In *Contractual Arrangements, Employment and Wages in Rural Labor Markets in Asia*, ed. Hans P. Binswanger and Mark R. Rosenzweig. New Haven: Yale University Press, pp. 242-262.

Basu, Alaka M. 1992. *Culture, the Status of Women, and Demographic Behaviour: Illustrated with the Case of India*. Oxford: Clarendon Press.

Becker, Gary S. 1965. “A Theory of the Allocation of Time.” *Economic Journal* 75, no. 299:493-517.

Beteille, Andre. 1969. *Caste, Class and Power: Changing Patterns of Stratification in a Tanjore Village*. University of California Press.

Bhalotra, Sonia. 2010. “Fatal Fluctuations? Cyclicality in Infant Mortality in India.” *Journal of Development Economics* 93, no. 1:7-19.

Biddle, Jeff, and Daniel S. Hamermesh. 2011. “Cycles of wage discrimination.” NBER Working Paper No. 17326.

Biewen, Martin. 2001. “Measuring the effects of socio-economic variables on the income distribution: an application to the East German transition process.” *Review of Economics and Statistics* 83, no. 1:185–90.

Black, David, Tseng, Yi-Ping, and Roger Wilkins. 2011. "Do changes in demographic characteristics explain declining male employment rates? Examination of the Australian case using a propensity score re-weighting decomposition approach." *Applied Economics* 43, no. 28:4215-4226.

Blau, Francine D., and Lawrence M. Kahn. 2003. “Understanding International Differences in the Gender Pay Gap.” *Journal of Labor Economics* 21, no. 1:106-144.

Blinder, Alan S. 1973. “Wage Discrimination: Reduced form and Structural Estimates.” *The Journal of Human Resources*, 8:436-455.

Boserup, Ester. 1970. *Women's role in economic development*. New York: St. Martin'.

Burgess, Robin, Deschenes, Olivier, Donaldson, Dave, and Michael Greenstone. 2011. "Weather and death in India." Cambridge, United States: Massachusetts Institute of Technology, Department of Economics. Manuscript.

Burton, Michael L., and Douglas R. White. 1984. “Sexual division of labor in agriculture.” *American Anthropologist* 86, no. 3:568–583.

Butcher, Kristin F., and John DiNardo. 2002. “The immigrant and native-born wage distributions: evidence from United States censuses.” *Industrial and Labor Relations Review* 56, no. 1:97–121.

Cameron, Lisa A., Dowling, J. Malcolm, and Christopher Worswick. 2001. “Education and Labor Market Participation of Women in Asia: Evidence from Five Countries.” *Economic Development and Cultural Change* 49, no. 3:459-477.

Chakravorty , Sanjoy, and Somik Vinay Lall. 2005. “Industrial Location and Spatial Inequality: Theory and Evidence from India.” *Review of Development Economics* 9, no. 1:47–68.

Chatterjee, Urmila, Murgai, Rinku, and Martin Rama. 2014. “Job Opportunities along the Rural-Urban Gradation and Female Labor Force Participation in India.” Mimeo, World Bank, Washington DC.

Chen, Martha. 1995. “A Matter of Survival: Women’s Right to Employment in India and Bangladesh.” In *Women, culture and development: a study of human capabilities*, ed. Martha C. Nussbaum and Jonathan Glover. Oxford University Press, pp. 37-59.

Chernozhukov, Victor, and Christian Hansen. 2008. “The reduced form: A simple approach to inference with weak instruments.” *Economics Letters* 100, no. 1: 68-71.

Chin, Yoo-Mi. 2011. "Male backlash, bargaining, or exposure reduction?: women’s working status and physical spousal violence in India." *Journal of Population Economics* 25, no. 1:175-200.

Chowdhury, Subhanil. 2011. “Employment in India: What Does the Latest Data Show?” *Economic and Political Weekly* 34, no. 32:23–26.

Daly, Mary C., and Robert G. Valletta. 2006. “Inequality and poverty in United States: the effects of rising dispersion of men’s earnings and changing family behavior.” *Economica* 73, no. 289:75–98.

Das, Maitreyi Bordia, and Sonalde Desai. 2003. “Are Educated Women Less Likely to be Employed in India?” Social Protection Discussion Paper No. 313, World Bank, Washington DC.

Das, Maitreyi Bordia. 2006. “Do Traditional Axes of Exclusion Affect Labor Market Outcomes in India?” Social Development Papers, South Asia Series, Paper no. 97.

Das, Maitreyi Bordia, and Puja Vasudeva Dutta. 2008. “Does caste matter for wages in the Indian labor market?” Draft paper, World Bank, Washington DC.

Dasgupta, Purnamita, and Bishwanath Goldar. 2006. "Female labour supply in rural India: An econometric analysis." *Indian Journal of Labour Economics* 49, no. 2:293-310.

Datt, Gaurav. 1996. *Bargaining power, wages and employment: an analysis of agricultural labor markets in India*. New Delhi: Sage Publications.

DiNardo, John. 2002. “Propensity score reweighting and changes in wage distributions.” Mimeo, University of Michigan, Ann Arbor.

DiNardo, John, Fortin, Nicole M., and Thomas Lemieux. 1996. "Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach." *Econometrica* 64, no. 5:1001-1044.

Doss, Cheryl R. 1999. “Twenty Five Years of Research on Women Farmers in Africa: Lessons and Implications for Agricultural Research Institutions: with an Annotated Bibliography.” CIMMYT Economics Program Paper No. 99-02, Mexico.

Dyson, Tim, and Mick Moore. 1983. “On kinship structure, female autonomy, and demographic behavior in India.” *Population and Development Review* 9, no. 1:35–60.

Ehsanullah, Ijaz Iqbal, Ahmad, Ashfaq, and Shakeel Ahmad Randhawa. 2000. "Effect of direct seeding and transplanting methods on the yield and quality of fine rice Basmati-370." *International Journal of Agriculture & Biology* 2, no. 3:251-252.

Eswaran, Mukesh, Kotwal, Ashok, Ramaswami, Bharat, and Wilima Wadhwa. 2009. “Sectoral Labor Flows and Agricultural Wages in India, 1982-2004: Has Growth Trickled Down?” *Economic and Political Weekly* 44, no. 2:46–55.

Eswaran, Mukesh, Ramaswami, Bharat, and Wilima Wadhwa. 2013. “Status, Caste, and the Time Allocation of Women in Rural India.” *Economic Development and Cultural Change* 61, no. 2:311-333.

Fairlie, Robert W. 2005. “An extension of the Blinder-Oaxaca decomposition technique to logit and probit models.” *Journal of Economic and Social Measurement* 30, no. 4:305–316.

Foster, Andrew D., and Mark R. Rosenzweig. 2003. “Agricultural development, industrialization and rural inequality.” Mimeo, Cambridge, Massachusetts: Harvard University.

---------------. 2004. “Agricultural Productivity Growth, Rural Economic Diversity, and Economic Reforms: India, 1970–2000.” *Economic Development and Cultural Change* 52, no. 3:509-542.

Gaddis, Isis, and Stephan Klasen. 2014. “Economic development, structural change and women’s labor force participation.” *Journal of Population Economics* 27, no. 3:639-681.

Gajbhiye, K.S., and C. Mandal. 2006. “Agro-Ecological Zones, their Soil Resource and Cropping Systems.” http://agricoop.nic.in/farm%20mech.%20pdf/05024-01.pdf (accessed January 12, 2011).

Goldberg, Jessica. 2011. “Kwacha Gonna Do? Experimental Evidence about Labor Supply in Rural Malawi.” Mimeo, Department of Economics and Ford School of Public Policy, University of Michigan, Ann Arbor, MI.

Goldin, Claudia. 1995. "The U-Shaped Female Labor Force Function in Economic Development and Economic History." In *Investment in Women's Human Capital*, ed. T. Paul Schultz. Chicago: University of Chicago Press, pp. 61-90.

Gronau, Reuben. 1980. "Home production--a forgotten industry." *The Review of Economics and Statistics* 62, no. 3:408-416.

Halliday, Timothy J. 2012. “Intra-Household Labor Supply, Migration, and Subsistence Constraints in a Risky Environment: Evidence from Rural El Salvador”. *European Economic Review* 56, no. 6:1001-1019.

Hertz, Tom, Campos, Ana P., Zezza, Alberto, Azzarri, Carlo, Winters, Paul, Quinones, Esteban J., and Benjamin Davis. 2009. “Wage Inequality in International Perspective: Effects of Location, Sector, and Gender.” Paper presented at the FAO-IFAD-ILO Workshop on Gaps, trends and current research in gender dimensions of agricultural and rural employment: differentiated pathways out of poverty, Rome, March 31- April 2.

Himanshu. 2005. “Wages in Rural India: Sources, Trends and Comparability.*” Indian Journal of Labor Economics* 48, no. 2:375-406.

Himanshu. 2011. “Employment Trends in India: A Re-examination.” *Economic and Political Weekly* 46, no. 37:43–59.

Hyslop, Dean R., and David C. Mare. 2005. “Understanding New Zealand’s changing income distribution, 1983–1998: a semi-parametric analysis.” *Economica* 72, no. 287:469–495.

IRRI. 2000. “Gender, Labor and Rainfed Technology”. In *Rainfed Rice: A Sourcebook of Best Practices and Strategies in Eastern India*, ed. V.P. Singh and R.K. Singh. Philippines: International Rice Research Institute, pp. 31-35.

Ito, Takahiro, and Takashi Kurosaki. 2009. “Weather risk and the off-farm labour supply of agricultural households in India. “ *American Journal of Agricultural Economics* 91, no. 3:697–710.

Jacoby, Hanon G. 1992. “Productivity of men and women and the sexual division of labor in peasant agriculture of the Peruvian Sierra.” *Journal of Development Economics* 37, no. 1-2:265-287.

Jayachandran, Seema. 2006. “Selling labor low: Wage responses to productivity shocks in developing countries.” *Journal of Political Economy* 114, no. 3:538–575.

Jejeebhoy, Shireen J. 2001. “Women’s autonomy in rural India: Its dimensions, determinants, and the influence of context.” In *Women’s Empowerment and Demographic Processes: Moving Beyond Cairo*, ed. Harriet Presser and Gita Sen. Oxford University Press, pp. 205-238.

Kannan, K.P., and G. Raveendran. 2012. "Counting and Profiling the Missing Labor Force." *Economic and Political Weekly* 47, no. 6:77-80.

Kanwar, Sunil. 2004. "Seasonality and Wage Responsiveness in a Developing Agrarian Economy." *Oxford Bulletin of Economics and Statistics* 66, no. 2:189-204.

Kharbanda, Varun. 2012. "Returns to Education in India." Mimeo, The University of Iowa.

Klasen, Stephan, and Janneke Pieters. 2013. "What Explains the Stagnation of Female Labor Force Participation in Urban India?" IZA Discussion Paper No. 7597.

Kumar, Hemanshu, and Rohini Somanathan. 2009. "Mapping Indian Districts across Census years, 1971-2001." *Economic and Polictical Weekly* 44, no. 41/42:69-73.

Kumar, K. Krishna. 2009. "Impact of Climate Change on India's Monsoonal Climate and Development of High Resolution Climate Change Scenarios for India." IITM Report.

Lahoti, Rahul, and Hema Swaminathan. 2013. "Economic Growth and Female Labor Force Participation in India." Working paper no. 414, Indian Institute of Management, Bangalore.

Lam, David, and Suzanne Duryea. 1999. "Effects of schooling on fertility, labor supply, and investments in children, with evidence from Brazil." *The Journal of Human Resources* 34, no. 1:160-192.

Lanjouw, Peter, and Rinku Murgai. 2009. "Poverty decline, agricultural wages, and non-farm employment in rural India: 1983-2004." *Agricultural Economics* 40, no. 2:243-263.

Laufer, Leslie A. 1985. "The Substitution between Male and Female Labor in Rural Indian Agricultural Production." Economic Growth Center Discussion Paper No. 472, Yale University, New Haven.

Leibbrandt, Murray, Levinsohn, James A., and Justin McCrary. 2010. "Incomes in South Africa after the fall of apartheid." *Journal of globalization and development* 1, no. 1.

Lincove, Jane Arnold. 2008. "Growth, Girls' Education, and Female Labor: A Longitudinal Analysis." *The Journal of Developing Areas* 41, no. 2:45-68.

Maitra, Pushkar. 2001. "Is consumption smooth at the cost of volatile leisure? An investigation of rural India."*Applied Economics* 33, no. 6:727-734.

Mammen, Kristin, and Christina Paxson. 2000. "Women's Work and Economic Development." *The Journal of Economic Perspectives* 14, no. 4:141-164.

Marchand, Joseph, and Sara Olfert. 2013. "The US gender gap through the Great Recession using an alternative approach to cyclicality." *Applied Economics Letters* 20, no. 3:276-281.

Mbiti, Isaac M. 2007. “Moving Women: Household Composition, Labor Demand and Crop Choice.” Mimeo, Department of Economics, Southern Methodist University, Dallas, TX, USA.

Mehta, Aashish, Felipe, Jesus, Quising, Pilipinas, and Shiela Camingue. 2013. "Where have All the Educated Workers Gone? Services and Wage Inequality in Three Asian Economies." *Metroeconomica* 64, no. 3:466-497.

Ministry of Home Affairs, Government of India. 2011. Census Data Online. http://www.censusindia.gov.in/2011-common/censusdataonline.html (accessed October 20, 2013).

Mitra, Arup, and Mayumi Murayama. 2008. “Rural to Urban Migration: A District Level Analysis for India.” IDE Discussion Paper 137.

Morduch, Jonathan. 1990. “Risk, Production, and Saving: Theory and Evidence from Indian Households.” Mimeo, Harvard University, Cambridge.

Mueller, Valerie, and Daniel E. Osgood. 2009. “Long-term Impacts of Droughts on Labor Markets in Developing Countries: Evidence from Brazil.” *Journal of Development Studies* 45, no. 10:1651-1662.

Mueller, Valerie, and Agnes R. Quisumbing. 2011. “How Resilient are Labor Markets to Natural Disasters? The Case of the 1998 Bangladesh Flood.” *Journal of Development Studies* 47, no. 12:1954-1971.

Munshi, Kaivan, and Mark R. Rosenzweig. 2009. “Why is mobility in India so low? Social insurance, inequality, and growth.” NBER Working Paper No. 14850.

Nayyar, Rohini. 1987. “Female Participation Rates in Rural India.” *Economic and Political Weekly* 22, no. 51:2207-2216.

Oaxaca, Ronald. 1973. “Male-Female Wage Differentials in Urban Labor Markets.” *International Economic Review* 14, no. 3:693-709.

O’Neill, June. 1985. “The Trend in the Male-Female Wage Gap in the United States.” *Journal of Labor Economics* 3, no. 1:S91-S116.

Palmer-Jones, Richard, and Kunal Sen. 2003. “What has luck got to do with it? A regional analysis of poverty and agricultural growth in rural India.” *Journal of Development Studies* 40, no. 1:1-31.

Pandey, Sushil, Bhandari, Hem Singh, and Bill Hardy. 2007. *Economic costs of drought and rice farmers’ coping mechanisms: a cross-country comparative analysis*. Philippines: International Rice Research Institute.

Parida, Jajati Keshari, and S. Madheswaran. 2010. “Spatial Heterogeneity and Population Mobility in India.” Institute for Social and Economic Change Working Paper 234.

Park, Seonyoung, and Donggyun Shin. 2005. "Explaining procyclical male–female wage gaps." *Economics Letters* 88, no. 2:231-235.

Pathak, H., Tewari, A. N., Sankhyan, S., Dubey, D. S., Mina, U., Singh, V., Jain, N., and A. Bhatial. 2011. "Direct-seeded rice: Potential, performance and problems–A review." *Current Advances in Agricultural Sciences* 3, no. 2:77-88.

Quisumbing, Agnes R. 1996. "Male-Female Differences in Agricultural Productivity: Methodological Issues and Empirical Evidence." *World Development* 24, no. 10:1579-1595.

Rahman, Lupin, and Vijayendra Rao. 2004. "The Determinants of Gender Equity in India: Examining Dyson and Moore's Thesis with New Data." *Population and Development Review* 30, no. 2:239-268.

Rajaraman, Indira. 1986. "Offered wage and recipient attribute: Wage functions for rural labor in India." *Journal of Development Economics* 24, 179-195.

Rao, V. M. 1972. "Agricultural Wages in India – A Reliability Analysis." *Indian Journal of Agricultural Economics* 27, no. 3:39-61.

Ribot, Jesse C., Najam, Adil, and Gabrielle Watson. 1996. "Climate Variation, Vulnerability and Sustainable Development in the Semi-Arid Tropics." In *Climate Variability, Climate Change and Social Vulnerability in the Semi-Arid Tropics,* ed. Jesse C. Ribot, Antonio Rocha Magalhaes and Stahis Panagides. Cambridge: Cambridge University Press, pp. 13-54.

Rose, Elaina. 2001. "Ex Ante and Ex Post Labor Supply Responses to Risk in a Low Income Area." *Journal of Development Economics* 64, no. 2:371-388.

Rosenzweig, Mark R. 1978. "Rural wages, labor supply, and land reform: A theoretical and empirical analysis." *American Economic Review* 68, no. 5:847-861.

---------------. 1984. "Determinants of Wage Rates and Labor Supply Behaviour in the Rural Sector of a Developing Country." *In Contractual Arrangements, Employment and Wages in Rural Labor Markets in Asia,* ed. Hans P. Binswanger and Mark R. Rosenzweig. New Haven: Yale University Press, pp. 211-241.

Schultz, T. Paul. 1994. "Human Capital Investment in Women and Men: Micro and Macro Evidence of Economic Returns." Occasional Papers Number 44, International Center for Economic Growth.

Schultz, T. Paul. 2002. "Why Governments Should Invest More to Educate Girls." *World Development* 30, no. 2:207-225.

Sen, Gita. 1985. "Paddy production, Processing and Women workers in India: The South versus the North East." In *Women in Rice farming, Proceedings of a conference on Women in Rice Farming Systems*. Philippines: International Rice Research Institute, pp. 385-404.

Sen, Kunal, Neff, Daniel, and Veronika Kling. 2012. "The Puzzling Decline in Rural Women's Labor Force Participation in India: A Re-Examination." GIGA working paper no. 196, GIGA Research Unit: Institute of Asian Studies, Hamburg, Germany.

Singh, Inderjit, Squire, Lyn, and John Strauss. 1986. *Agricultural Household Models: Applications, Extensions and Policy.* Baltimore: Johns Hopkins University Press.

Singh, Ram D. 1996. "Female Agricultural Workers' Wages, Male-Female Wage differentials, and Agricultural Growth in a Developing Country, India." *Economic Development and Cultural Change* 45, no. 1:89-123.

Singh, A.K., Choudhury, B.U., and B.A.M. Bouman. 2002. "Effects of rice establishment methods on crop performance." In *Water-wise rice production*, ed. B.A.M. Bouman, H. Hengsdijk, B. Hardy, P.S. Bindraban, T.P. Tuong and J.K. Ladha. Philippines: International Rice Research Institute, pp. 237-246.

Skoufias, Emmanuel. 1993. "Seasonal Labor Utilization in Agriculture: Theory and Evidence from Agrarian Households in India." *American Journal of Agricultural Economics* 75, no. 1:20-32.

Solon, Gary, Barsky, Robert, and Jonathan A. Parker. 1994. "Measuring the cyclicality of real wages: how important is composition bias?" *Quarterly Journal of Economics* 109, no. 1:1–25.

Tam, Henry. 2011. "U-shaped female labor participation with economic development: Some panel data evidence." *Economics Letters* 110, no. 2:140-142.

Tumbe, C. 2014. "The Great Indian Migration Wave, 1870-2010, Persistence & Consequences." Mimeo, The Economic Growth Center, Yale University.

Wadood, Syed Naimul, and Russell L. Lamb. 2006. "Choice of Crops and Employment Uncertainty in the Off-farm Labor Market." MPRA Paper No. 10779.

WEDO. 2008. "Gender, Climate Change and Human Security Lessons from Bangladesh, Ghana and Senegal." http://www.wedo.org/wp-content/uploads/hsn-study-final-may-20-2008.pdf (accessed December 18, 2011).

World Bank. 2008. "World development report 2008: Agriculture for development." http://siteresources.worldbank.org/INTWDR2008/Resources/WDR_00_book.pdf (accessed September 24, 2013).

---------------. 2012. "World Development Report 2012: Gender Equality and Development." http://siteresources.worldbank.org/INTWDR2012/Resources/7778105-1299699968583/7786210-1315936222006/Complete-Report.pdf (accessed November 18, 2013).