Removing the ‘Veil of Ignorance’: Nonlinearities in Education Effects on Gender Wage Inequalities

Sandeep Mohapatra* Bruno Wichmann* Philippe Marcoul*

Abstract

A large literature studies the mean gender wage gap in developing countries and finds mixed evidence about the role of education policies in closing gender earnings inequalities. We contribute to this literature by exploring two types of nonlinearities in wage earning regressions: i) nonlinearities on the effects of education on expected earnings along the distribution of education endowments; and ii) heterogeneities on the contributions of education to the gender wage gap at different quantiles of the wage distribution. Our analyses provide new insights on how these nonlinear effects can be used to set up better targeted gender and development policies.

Keywords: Education; Work Experience; Nonlinearities; Regression Splines; Unconditional Quantile Regression; Multinomial Selection; India.

JEL codes: I26; C14.

*Department of Resource Economics and Environmental Sociology, University of Alberta.

1 Introduction

Empirical work shows that, on average, women are paid less than men in labor markets around the world (e.g. Blau and Kahn, 2000; Manning and Robinson, 2004). A world-wide review of the literature which includes 260 published papers, spans 63 country contexts, and comprises 788 individual analyses shows that the average female-male wage ratio, which was around 65% in the 1960s, persisted at around 30% in the 1990s (Weichselbaumer and Winter-Ebmer, 2005). In the U.S, for example, after decades of stagnation at about 60%, the female-male ratio of annual wage earnings for full-time workers increased dramatically in the 1980s but by 2014 women still earned only 80% of what men earned (Blau and Kahn, 2016). Bayer and Rouse (2016) report gender gaps in salaries and tenure and promotion rates in their study about the diversity in the economics profession in the U.S.

In developing countries, where off-farm wage earnings are an increasingly important source of income and insurance, and where women constitute on average 43% of the agricultural labor force, these inequalities are even more apparent and their implications more stark (United States Agency for International Development, 2012). In India, for instance, the average female-male wage ratio has hovered around 50% throughout the 1980s, the 1990s, and the early 2000s (Menon and Van der Meulen Rodgers, 2009).[1] Relatedly, scholars have found that key indicators of women's welfare such as female mortality rates are linked to female wage earnings (Rosenzweig and Schultz, 1982; Kishor, 1993).

While wage differentials represent one of the many dimensions of gender inequality, they continue to be the focus of an extensive and growing literature on the economics of gender. The magnitude of the gender wage gap and its underlying reasons can vary across regions and within regions over time.[2] Understanding these underlying reasons is especially important for developing countries where gender wage differentials are not only important determinants of women's welfare, poverty alleviation and overall economic growth but also key for assigning status and bargaining power within households (World Bank, 2011).

Studies seeking to inform gender wage policies commonly use Mincerian wage regressions to estimate the labor market returns to education and experience for men and women or, alternatively, to decompose the male-female expected wage difference into its contributing sources. Two methodological considerations, however, have been overlooked in much of the

[1] Although these estimates vary across studies and regions of India.

[2] For example, in the U.S, differences in human capital variables (education and experience) explain little of the gender wage gap observed in the last decade, while in Canada changes in human capital characteristics continue to play a relatively large role (Blau and Kahn, 2016; Baker and Drolet, 2010).

literature. First, wage regressions estimated in the literature fail to capture the full range of education's effects on average wage earnings, across the distribution of education. That is, not enough attention is paid to non-linearities in the effects of education and experience, and their interactions, on earnings. This can be especially problematic in developing countries with gender inequalities in educational attainment (Zeng et al., 2014), as well as highly uneven distributions of education and experience within each gender. The standard method of including linear and squared terms of education and experience has been criticized in a recent contribution due to its lack of flexibility in capturing the partial non-linear effects of individual covariates (Royston and Sauerbrei, 2008). The challenge is compounded when there are multiple non-linear predictors and, therefore, the functional form for each has to be determined jointly, while accounting for possible linear effects of other control variables (Poirier and Garber, 1974; Royston and Sauerbrei, 2004).

Accounting for non-linearities in Mincerian wage regressions is important for at least two reasons. First, they represent the degree to which education and experience complement each other, the turning point in education and experience levels at which the complementarity emerges or disappears, and where labor market returns are maximized. Understanding gender differences in these patterns can help policy makers to differentially target sub-groups of women in different regions of the education and experience distributions. For instance, if experience substitutes for education in securing labor market returns at the bottom of the distribution of education but not at the top, then policies that invest in vocational training for less educated workers can be more effective at increasing their earnings than those that simply invest in formal education. In contrast, if experience amplifies the return to education then education and vocational training policies will be highly complementary.

Second, existing gender wage gap decompositions fail to capture the full range of education's effects across the distribution of male and female wage earning functions. Most of the work on gender wage gaps in developing countries focus on the mean wage gap. The mean gap may not be a sufficient statistic for describing inequality in the labor market if, for instance, the wage gap or its determinants vary at different points of the wage distribution (Koop and Tobias, 2004; Lemieux, 2006). Going beyond the mean can be especially insightful when the gender wage gap can be decomposed into endowment effects and return effects. Such a decomposition allows policy makers to learn the proportion of the gap that comes from differences in endowments between men and women (e.g. different education attainments) against the proportion that comes from differences in returns to endowments, perhaps due to discrimination (e.g. wage gaps for equally qualified workers of opposite gender). Some studies have used conditional quantile regression based approaches to explore distributional gender wage gaps (e.g. Bhaumik and Chakrabarty, 2008). However, quantile regressions estimate the effect of determinants on the conditional distribution of wages and, therefore, are of limited use from a policy perspective in identifying unconditional effects of education on distributional gender gaps (Koenker and Bassett, 1978).[3]

The paper examines the two distributional issues discussed earlier by exploring education's contribution to women's earning and the gender wage gap using a nationally representative dataset on India's formal wage labor markets. We have two specific objectives. First, we seek to estimate *non-linear* earnings functions for men and women in order to identify gender-differentiated patterns in the shape of education's partial effects, how the partial effects of education and experience interact, and the levels of education and experience that maximize average returns. Specifically, we seek to identify regions on the education and experience distributions where different types of human capital policy interventions are likely to be most effective in increasing women's wage earnings. Second, we search for education's effect on gender gaps along the distributions of male and female outcomes. We investigate the following questions: Is the effect of education on the wage gap constant, or does it vary along wage distributions? What is the share of the wage gap due to gender differentials in education endowments vs returns? What is the net effect of education on the gap? Do the effects of educational endowments and returns on the wage gap offset one another?

To meet our first objective we use a semi-parametric approach based on multivariate regression splines to model non-linearities (Royston and Sauerbrei, 2004, 2008). It is widely recognized that semi-parametric models are more flexible than parametric regression based approaches, although the literature on gender wage differentials continues to rely on the latter. Compared to other non- and semi parametric estimators such as Kernel and local regressions which allow for a single independent variable, the semi-parametric approach we use allows modeling of multivariate relationships. Moreover, compared to standard regression methods where coefficients on covariates are constants in a single sample, the approach we use allows us to specify the coefficients as functions. To the best of our knowledge this is the first application of this approach to studying inequalities in men's and women's earnings.

To meet our second objective we use a recently developed distributional decomposition approach based on unconditional quantile regressions (UQRs) (Firpo et al., 2007, 2009). The approach combines the advantage of the popular Oaxaca-Blinder decomposition with

[3]The supplemental information appendix (A4) discuss the limitations of using quantile regressions to examine gender gaps in India.

methods that allow distributional decompositions. The Oaxaca-Blinder decomposition can break up the gender wage gap into an endowment and a return effect for each covariate such as education. The method, however, cannot be used for distributional analysis since the decomposition is valid only at the mean of mens' and womens' wage distributions.[4] The approach we use, in contrast, allows us to do both – estimate the gender wage gap for each quantile as well as to decompose the unconditional return and endowment impacts of education at each quantile. This is the first application of this approach to studying inequalities in men's and women's earnings in India.

India is an appropriate case study because a fairly large literature devoted to understanding the role of education policies in closing gender Indian wage gaps has produced inconsistent findings (e.g. Menon and Van der Meulen Rodgers, 2009). Moreover, most of the studies focus on the mean wage gap and ignore distributional issues. The few studies that consider distributional issues have failed to identify the role of individual covariates such as education in explaining wage gaps along male and female distributions (e.g. Bhaumik and Chakrabarty, 2008). Despite the mixed findings, reducing the wage differentials continues to be a policy priority and represents an important dimension of women's empowerment and household welfare in the country (e.g. Duraisamy, 2002).

Our semi-parametric estimates reveal distinct non-linearities between men's and women's human capital earnings functions. We find increasing effects of education and experience on men's wages, with the education effect being convex and the experience effect being concave. The effects of education and experience on women's earnings are generally smaller than those of men's. The education effect on women's wages is U-shaped, while the experience effect fluctuates around its mean. Thus, at a low education level, an increase in education will not enhance women's income. The effect of the interaction of education and experience on wages is generally flat for men and increasing for women. This behavior suggests that while experience and education are complements for women, they are substitute inputs for men.

Unlike previous studies we find little evidence that education's return and endowment effects on the gender wage gap are offsetting. Rather, our results suggest that differences in education endowments and returns between men and women contribute to the gender wage gap with varying intensities along the wage distribution. Depending on the education

[4] Alternative methods have been proposed for distributional decompositions, namely, the methods by: Juhn et al. (1993) based on ranking distributions; DiNardo et al. (1996) based on re-weighting techniques; Machado and Mata (2005) and Melly (2005) based on conditional quantile regressions. These distributional methods, however, have the common shortcoming that they cannot break up the gender wage gap into an endowment and return effect corresponding to each covariate.

level and the level of wages, we find cases where: 1) lower endowments among women are overridden by higher returns to education, creating net reductions in the wage gap (such as for primary education at most wage levels and post-primary education at some higher wage levels); 2) lower endowments among women are combined with lower returns to education, creating increases in the wage gap (such as for post-primary education at lower wage levels); 3) higher endowments among women are combined with varying returns to education creating increases and decreases in wage gaps (such as for university education at lower wage levels); and 4) higher endowments among women are combined with higher returns to education creating decreases in wage gaps (such as for university education at higher wage levels). The varying contributions of education endowments and returns to the wage gap suggest an important role for education policy in recognizing heterogeneity in policies targeted towards reducing the gender wage gap.

The rest of the paper is organized as follows. In the next section, we provide a brief description of labor markets in India's economy. Section 3 introduces the data. In section 4 we estimate non-linear effects of education on men's and women's earnings. In section 5 we use UQRs to decompose gender wage gaps along the wage distribution. Section 6 examines the robustness of the UQR decompositions to selection into labor markets. We conclude with an emphasis on policy implications of our results.

2 Gender Gaps and the Indian Context

Summarizing existing empirical results, while the world's gender wage gap is decreasing over time, progress has been slow in India (Duraisamy and Duraisamy, 1996; Kingdon and Unni, 2001). Both differential returns and differential endowments have been used to explain the India's gender wage gap. According to the literature, average education levels contribute little to the gender wage gap. Kingdon and Unni (2001) use data from two states in India and find that the average gender wage gap is bigger due to men's higher education levels, relative to women. They also show that it is smaller by almost the exact same amount due to women's higher returns to education, relative to men. This surprising offsetting result has secured a loose consensus among scholars, and is also supported qualitatively by the findings of other studies using data from multiple years (e.g. Bhaumik and Chakrabarty, 2008).[5] However, the study by Kingdon (1998) of one state in India finds that on average men receive higher wages than women not only because they have more education but because

[5]The signs on the education endowment and return effect are opposing for most education levels.

they also receive a much higher return to their education compared to women.

One clear motivation for our work is the lack of empirical evidence on education as a specific policy lever for closing gender wage gaps in India. We choose to focus on distributional issues which are understudied. By doing so, our results also produce insights into education's contribution to closing gender gaps in developing economies, more generally. Educating girls is regarded as a silver bullet for economic development. Surprisingly, however, while education's value in terms of benefiting girls and women is well recognized[6], evidence on its ability to reduce gender inequalities in developing countries is less clearly established (Sperling and Herz, 2004). Scholars argue that policy-makers often assume that education can reduce wage inequality between men and women (e.g. Gates, 2014).[7] Various reasons have been put forth for the apparent lack of leverage of education policies in closing different types of gender gaps in developing countries (Aslam et al., 2012). We explore an additional and under-emphasized reason, related to empirical methodology, which may contribute to the weak evidence on education's role in closing gender gaps.

Our two methods (splines and decompositions) have different strengths and weaknesses. Spline regressions do not decompose the wage gap but they represent a type of regression model that has not been applied to estimate human capital functions. Specifically, we use spline regressions to construct partial predicted wages from education, experience, and their interaction, holding all other covariates constant. These predictions reflect the non-linear relationships between earnings and human capital and allow us to examine differences between men and women. Decompositions based on UQRs offers a deeper investigation by examining gender inequalities along the wage distribution, however, they abandon the non-linearity explicitly; non-linearities are subsumed in heterogeneous distribution effects. Thus, the decomposition allows us to contrast the education effect on gender gaps faced by the most vulnerable females in the left tail of women's outcome distribution, relative to the effect on the gap faced by relatively less vulnerable women in the middle and right tails of the distribution. Our approach is consistent with Heckman and Vytlacil (2005) who emphasize

[6]Extensive empirical evidence shows that girls' education benefits not only the girls themselves, but also their households and communities, and contributes to overall economic growth (see Hanushek and Kimko, 2000; World Bank, 2011). An additional year of education among women leads to an increase in their earnings (Levin, 2009), which can translate into a greater influence of women over household decisions. This has been shown to have salutary effects on maternal and child health, human capital accumulation, and productivity (World Bank, 2011).

[7]Education may have individual level effects on human capital which can increase the value of women's skills in the labor market. Moreover, if women have low skills and, hence, lower wages relative to men, then increasing education among women can be expected to make the low skills group more competitive, thereby, reducing the gender wage gap.

the importance of assessing the impact of public policy on outcome distributions without relying on the *veil of ignorance* assumption of welfare economics. The assumption maintains that only shifts in the overall outcome distribution and its mean matters for assessing the impact of a policy. The position of an individual within the outcome distribution does not matter. We believe that studying wage gaps all along the earnings distribution curve helps to setup *better targeted* policy instruments.

3 Data

The paper uses data from the 62nd round of the National Sample Survey (NSS), collected by the government of India in 2005-2006.[8] The survey is nationally representative and contains information on labor market participation, wage earnings, and the workers' occupations and industries classification. The survey also collects information on socio-economic variables and demographics such as gender, education, age, social group, and location of residence. The dataset contains 377,377 individuals, living in 78,879 households, located across 5,125 urban blocks and 4,798 rural villages.

The general goal of the paper is to study heterogeneities related to the distribution of human capital (education and experience) and the distribution of earnings. Thus, our empirical models include variables that capture earnings, education, experience, and a set of controls to capture differences in the type of wage employment (i.e. industry classification), the skill level of the employment (i.e. occupation classification), social and religious affiliations, and their geographical location. We describe these variables below while summary statistics are presented in table 1.

Our focus is the wage employment portion of the Indian labor market. As a result, our sample consists of 31,541 individuals, aged 15-60 years, who are engaged in wage employment. The sample has 5,156 women and 26,385 men, which reflects India's low female labor market participation. Wage earners are defined as persons who receive a wage on a regular basis through a formal contract (our sample does not include the informal labor market). We measure earnings using information on daily wage (in cash and in-kind), based on a 7-day

[8] The survey was implemented with a two-stage design. In a first stage, primary sampling units (i.e. villages in rural areas and area blocks in urban areas) were selected based on population size. In a second stage, the survey was administered to households within the first stage sampling units, with selection based on income. This two-stage stratified random sampling design offers an accurate representation of varying income and population distributions in India.

Table 1: Summary Statistics

	Men		Women	
	Mean	Std. Dev.	Mean	Std. Dev.
Earnings:				
Ln Wage	4.907	0.788	4.569	0.959
Human Capital:				
Primary Education	0.104	0.306	0.079	0.269
Middle Education	0.198	0.399	0.105	0.306
Secondary Education	0.188	0.390	0.119	0.323
Higher Secondary Education	0.122	0.327	0.103	0.304
Graduate/University Education	0.285	0.451	0.391	0.488
Technical Degree	0.025	0.155	0.041	0.197
Years of Education	9.544	4.688	9.348	5.781
Experience	21.41	11.92	20.80	12.86
Occupation:				
High Skilled	0.193	0.394	0.414	0.493
Medium Skilled	0.291	0.454	0.201	0.401
Low Skilled	0.377	0.485	0.132	0.339
Industry:				
Construction	0.040	0.195	0.014	0.117
Low Tech Mfg.	0.180	0.385	0.119	0.324
High Tech Mfg.	0.079	0.269	0.016	0.124
Modern Services	0.482	0.500	0.624	0.484
Computer & IT	0.008	0.088	0.012	0.108
Social/Religion:				
Scheduled Caste/Tribe	0.219	0.413	0.271	0.444
Other Backward Classes	0.314	0.464	0.303	0.459
Muslim	0.128	0.334	0.082	0.274
Region:				
Urban	0.732	0.443	0.724	0.447
North	0.168	0.374	0.130	0.336
South	0.167	0.373	0.266	0.442
East	0.099	0.299	0.098	0.298
West	0.190	0.392	0.154	0.361
North-East	0.097	0.296	0.131	0.338

Notes: All variables are binary indicators except for Ln Wage (weekly Ruppes), Years of Education, and Experience.

reference week.[9] It is important to incorporate in-kind earnings because wage workers in India typically receive bonuses and other amenities such as free lodging, medical treatment,

[9]The activity on which a person spent the longest time during the 365 days preceding the survey is regarded as the principal activity of the person. To assign this status, a two-stage binary categorization was first used to determine if the person was or not in the labor force (depending on the time spent in the previous year in each category). In a second stage, for persons belonging to the labor force, their status of employed or unemployed was determined (also based on the same longest time spent time criteria). Our wage variable excludes amounts received or receivable as over-time beyond normal daily work hours. Each day of a reference week designated by the survey was assigned discrete intensity levels of work performed and these intensities factored into daily wage rate calculations. Thus, individuals in the sample are subject to the same reference of what they earn in a normal day. This approach allows us to study wage gaps (intensive margin of earnings) as opposed to earning gaps (that account for both differences in wages or differences in hours). While this approach can have some measurement error it is quite well received in the literature.

telephones, etc. These amenities are evaluated at retail prices and included in our measure of wage earnings. The data in table 1 shows that, on average, men's earnings are higher than women's ($p < 0.01$, two-sided t-test).

The information in the NSS allows us to quantify human capital with different measures of education and experience. First, we construct five education dummy variables: primary education, middle education, secondary education, higher secondary education, and graduate/university education. The base category is below primary school (which includes illiterates). Individuals in India may also obtain an additional graduate level technical certificate or diploma in a specialized discipline (e.g. engineering, business management). As a result, we also construct a dummy variable indicating technical degree. Second, we follow Duraisamy (2002) and construct a measure using years of education. According to the education system in India, the number of schooling years increments required to complete primary, middle, secondary, higher secondary and university education are 5, 3, 2, 2 and 3, respectively. Degrees in technical subjects require about three years. Our years of education measure uses this information to assign education years over the previous level of schooling. Experience is also an important component of human capital. We follow the literature and measure labor market experience in terms of potential employable years following education (e.g. Ito, 2009). Experience is calculated as an individual's age minus number of years of schooling minus six, i.e. the number of years before schooling. Women in our sample have, on average, 0.2 years of education less than men ($p < 0.01$, two-sided t-test), and 0.6 years of experience less than men ($p < 0.01$, two-sided t-test).

We estimate wage earning functions to examine returns to human capital controlling for several covariates. Specifically, we use a series of dummy variables to capture workers' differences in occupation, industry, social stratification, religion, and spacial characteristics (i.e., urban/rural worker and region effects). The supplemental information appendix (A1) presents a detailed discussion of these variables.

4 Non-linear Effects of Education and Experience on Earnings

We specify the following expected earnings function with the standard assumption that wages are distributed log-normally with constant variance. For each individual i, we have

$$E(W_i|X_{ij}, Z_{il}) = \alpha_0 + \sum_{j=1}^{3} f_j(X_{ij}) + \sum_{l=1}^{L} \beta_l Z_{il} \tag{1}$$

where W_i denotes log wages, X_{ij} denotes measures of human capital (years of education, years of experience and their interaction), Z_{il} denotes a set of L control variables representing shifters of the earnings function and α_0 is an intercept term. The f_j are smooth functions that respectively link the measures of education and experience to the expected value of wages. They are non-linear counterparts of coefficients in a linear regression.[10]

Estimating separate models for men and women, we examine the contributions of three predictors to expected wages: education, experience and their interaction. All predictions are statistically significant. The top panel of figure 1 shows a plot of the partial predicted wages from education, holding all other covariates constant, along with 90% confidence bands. We refer to these predictions as the *pure education effect*. There is a striking gender difference in the contribution of education. For men, the pure educational effect is increasing and convex. For higher years of education, men get higher wages and at an accelerated rate. In contrast, women's pure education effect is U-shaped. At low levels of education, women are predicted to get higher wages on average than men but the overlapping confidence bands suggest there is no statistical difference. Women's predicted wages actually decline with higher education until they achieve about 8 years of education. Beyond this threshold the effect of education starts to rise. However, with the exception of earnings at low education levels, women are predicted to receive significantly lower wages than men for 3.5 years of education and higher; this is true in particular for the highest levels of education (i.e., between 10 and 15 years of education).

The theory of human capital predicts diminishing marginal returns to education accumulation. However, recent literature focusing on developing countries reports large returns to education at the upper-end of the education distribution (e.g. Söderbom et al., 2006). This convexity may be a result of imbalances between the demand and supply of low and high skill workers, short term frictions at the supply level, and structural changes in the Indian economy.[11]

The middle panel of figure 1 shows similar estimates of the effects of experience holding education and other covariates constant, along with 90% confidence bands. Men face a familiar concave earning experience function. Men with several years of experience receive large

[10] The regression splines are further discussed in the supplemental information appendix (A2).

[11] Refer to the supplement information appendix (A3) for a deeper discussion.

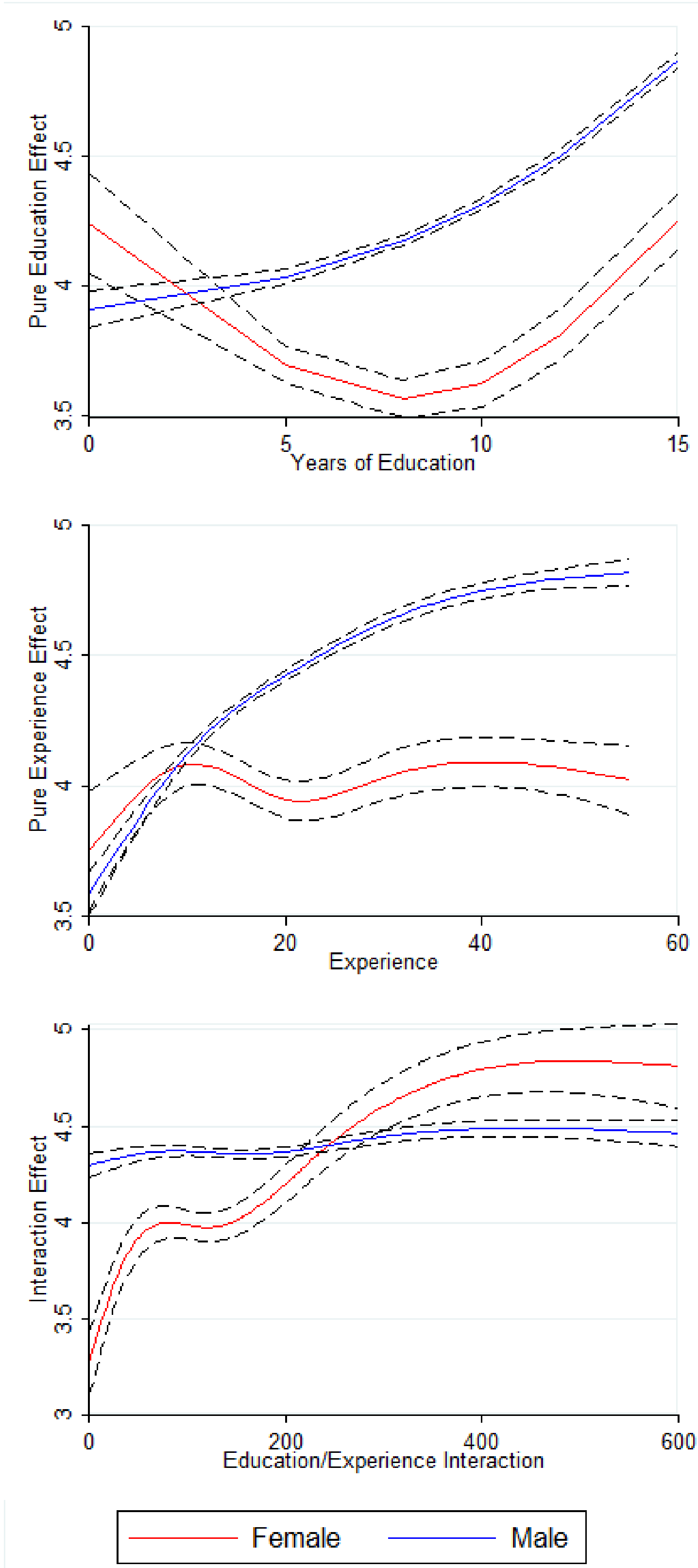

Figure 1: Non-linear effects of education, experience, and their interaction

wage premiums due to the pure experience effect. Experience, however, has a very different prediction for women's wages. While the effect of experience is statistically significant, it does not exhibit the increasing pattern we estimate for men. In fact, wages at high levels of experience are not much higher than those at low experience levels.

Women's wages, however, are highly responsive to the interacted or joint effects of education and experience. Figure 1 (bottom panel) shows that women's interaction effect is lower than that of men's at low interaction levels. Nevertheless, women's joint effect is increasing over a wide range of education and experience, and reaches levels above those of men at higher interaction levels after which it starts to level off. Men face an almost flat pattern of interaction effects. These results suggest that experience and education are complements for women and substitutes for men.

It is important to acknowledge that our analysis, like many in the literature (e.g. Duraisamy, 2002), faces some data limitations. Our experience measure can be thought of as potential experience. While widely used in the literature due to a lack of data on actual experience, the measure reflects the joint evolution of education and age rather than experience per se (Machado and Mata, 2002). Specifically, the experience measure is likely to overstate the market work experience of women since women are more likely to face interruptions in work than men and therefore the experience measure could understate gender differences in experience. Our results regarding women's experience should thus be interpreted with this caveat in mind.

5 Decomposing Earnings Inequalities

Thus far, our results suggest that Indian labor markets reward men's and women's education differently. The previous section used spline regressions to estimate non-linear effects and found important inequalities between men's and women's returns to human capital. As these results highlight gender discrepancies in education and experience effects, a natural next step is to obtain a deeper understanding of why Indian labor markets generate such inequalities. This section focuses on decomposing the effect of education and experience on the gender wage gap into two components: gender differences in endowments and gender differences in returns to those endowments.

The standard Oaxaca-Blinder method for decomposing the gender wage gap concentrates on mean wages earned by an individual of gender $g = m, f$ using the conditional mean functions:

$$E(W_g|x_{gk}) = \sum_{k=1}^{K} \beta_k^g x_{gk} \tag{2}$$

where $W_{g=m}$ and $W_{g=f}$ are the (log) wages respectively received by the men and the women of our sample and where x_k are vectors of endowment variables that explain wages. We can compute the difference between $E(W_m|x_{mk})$ and $E(W_f|x_{fk})$ and introduce the counterfactual $\beta_k^m \overline{x}_{fk}$ in the new expression. Then, replacing expectations with sample averages, a mean wage gap is defined as

$$E(W_m) - E(W_f) = \underbrace{(\overline{W}_m - \overline{W}_f)}_{\substack{\text{Mean} \\ \text{Gender Wage Gap}}} = \underbrace{\sum_{k=1}^{K} \beta_k^m (\overline{x}_{mk} - \overline{x}_{fk})}_{\text{Endowment Effect}} + \underbrace{\sum_{k=1}^{K} \overline{x}_{fk} (\beta_k^m - \beta_k^f)}_{\text{Return Effect}}, \tag{3}$$

where $\overline{W}_m$ and $\overline{W}_f$ are the mean wages for men and women respectively, and $\overline{x}_{mk}$ and $\overline{x}_{fk}$ denotes the corresponding average values of the explanatory variables. The two terms on the right hand side represent the decomposition of the mean wage gap due to gender differences in mean endowments (e.g. differences in education) and gender differences in returns to education that can be interpreted as resulting from discrimination. A positive (resp. negative) mean endowment effect contributes to increase (resp. decrease) the mean gender wage gap and represents the portion of the gap that comes from differences in endowments between men and women, holding the marginal effect of endowments constant. A positive (resp. negative) return effect also contributes to increase (resp. decrease) the gender wage gap and captures the portion of the gap that comes from differences in endowments' returns between men and women, holding endowments constant. The additive components of the return effects and endowment effects corresponding to each of the K sources are typically computed using OLS methods (Jann, 2008).

Notice that the results calculated from equation (3) oversimplify the analysis of the gender wage gap as they ignore possible heterogeneities of the gap along the wage distribution. For instance, suppose that the effect of education on the average gender wage gap is close to zero or null (a situation similar to that of India, where positive average endowment effects are somewhat offset by negative average return effects). This result would lead policy makers to conclude that education plays no role in determining the gender wage gap. However, this zero average effect could arise, for instance, from a negative effect of education on the gap in the upper tail of the wage distribution (i.e., for highly paid workers, education contributes to reduce the gap) paired with a positive education effect from workers in the lower tail of the wage distribution (i.e., the education effect for low-paid workers increases the gender wage gap). Clearly, average education effects may hide a significant heterogeneity along the wage distribution. In India, this heterogeneity is reflected on differences in wage gaps along the wage distribution, but may also be reflected on differences in the wage gap *components* along the distribution (i.e., varying endowments and return effects at different wage quantiles).

We examine gender wage gaps in India using a detailed distributional decomposition procedure developed by Firpo et al. (2007). This approach starts by defining an *influence function* (IF) which captures the effect of a small disturbance in the wage distribution on the τ^{th} unconditional wage quantile, θ_g^τ, for each gender $g = m,f$. Formally, we have $IF(\theta_g^\tau)$ $= [\tau - \mathbf{1}_{(W_g \leq \theta_g^\tau)}]/f(\theta_g^\tau)$, where f is the PDF of wages. Subsequently, a recentered influence function (RIF) is obtained by adding the IF to the quantile, it is defined as $RIF(\theta_g^\tau) = IF(\theta_g^\tau) + \theta_g^\tau$ and is used as dependent variables for estimating the conditional mean function for an individual of gender g

$$E(RIF[W_g;\theta_g^\tau|x_{gk}]) = \sum_{k=1}^{K}\beta_k^{g\tau}x_{gk} \tag{4}$$

where β_k^τ are the RIF-regression coefficients on endowment variables at quantile τ. Firpo et al. (2007) show that a detailed decomposition, similar to the Oaxaca-Blinder decomposition, can be obtained for the gender wage gap at any quantile τ as

$$E(RIF[W_m;\theta_m^\tau|x_{mk}]) - E(RIF[W_f;\theta_f^\tau|x_{fk}]) = \tag{5}$$

$$\underbrace{(\theta_m^\tau - \theta_f^\tau)}_{\substack{\text{Quantile}\\ \text{Gender Wage Gap}}} = \underbrace{\sum_{k=1}^{K}\beta_k^{m\tau}(\overline{x}_{mk} - \overline{x}_{fk})}_{\text{Endowment Effect}} + \underbrace{\sum_{k=1}^{K}\overline{x}_{fk}(\beta_k^{m\tau} - \beta_k^{f\tau})}_{\text{Return Effect}},$$

The above expression is derived using the property that the expectation of the RIF of an unconditional quantile equals the quantile itself and the LIE, that is, $E(RIF[W_g;\theta_g^\tau]) = \theta_g^\tau = E_x(E(RIF[W_i;\theta_g^\tau|x_{gk}]) = E_x(\sum_{k=1}^{K}\beta_k^{g\tau}x_{gk})$. A method for computing standard errors to test the statistical significance of each contributing factor is available from Jann (2008). We use the standard errors on each marginal effect to test for their statistical significance.

Another estimation challenge could arise due to the potential endogeneity of the education variables. We do not control for this potential problem for a number of reasons. First, there is no information in the dataset to construct commonly used instrumental variables, such as parental characteristics, to deal with the possible bias. In fact, the absence of this information does not allow other studies in India using a nationally representative dataset to control for endogeneity, including the benchmark study of Duraisamy (2002). The same issue arises in the papers by Chi and Li (2008) and Ahmed and Maitra (2015), which to date, are the only studies in the literature that applies a UQR-based approach to developing countries. Therefore, our study of education's distributional impacts is thus comparable to these previous studies.[12] Moreover, Harmon et al. (2003) point out that if educational returns are heterogeneous, instrumental variable estimation is not valid. Harmon et al. (2003) also show that the biases due to endogeneity and sample selection could work in opposite directions, thereby canceling each other out and leaving the OLS estimates to be the best approximation.

Table 2 presents the results of the Firpo et al. (2007) decomposition of the gender wage gap into its driving factors estimated at 10th, 50th and 90th quantiles. Note that the UQR-based decomposition relies on linear earnings functions and, as a result, does not explicitly accommodate non-linear effects of education. To address this, instead of incorporating education in the model through our variable "years of education" (see discussion above), we control for different education levels by specifying five dummy variables for the different education attainments, namely: primary, middle, secondary, higher-secondary, and university educations. This allow us to use the approach of Firpo et al. and decompose each education level effect along the wage distribution. We provide graphical representations of the marginal effects of education and experience on the gender wage gap for the 5th to the 95th quantiles of the wage distributions, computed at 5-unit intervals (figures 2-4). Further to previous studies, our results also suggest the existence of a wage gap between men and women. Unlike the previous literature, our results also suggest that the gender wage gap differs across the wage distribution.

For each education level, we discuss how endowment differentials and return differentials among men and women affect the gender wage gap. The decomposition results for primary education on the gender wage gap at 10th, 50th and 90th quantiles are reported in table 2, and are shown at 5-unit intervals in the connected scatterplot in figure 2 (top row). We estimate that lower participation in primary education by women, relative to men, has a statistically significant widening effect (of about 0.3%) on the male-female wage gap (columns

[12]The Indian study by Duraisamy (2002) focuses on mean returns to education while the studies by Chi and Li (2008) and Ahmed and Maitra (2015) decompose wage gaps along the wage distribution in China and Bangladesh, respectively.

Table 2: Wage decomposition between women and men at selected quantiles

	10th Quantile		50th Quantile		90th Quantile	
	Endowment Effect	Return Effect	Endowment Effect	Return Effect	Endowment Effect	Return Effect
	(1)	(2)	(3)	(4)	(5)	(6)
Human Capital:						
Primary	0.0028***	-0.0003	0.0030***	-0.0307***	0.0027***	0.0107*
	(0.001)	(0.006)	(0.001)	(0.007)	(0.001)	(0.006)
Middle	0.0267***	0.0461***	0.0303***	0.0071	0.0160***	0.0052
	(0.003)	(0.007)	(0.002)	(0.008)	(0.002)	(0.008)
Secondary	0.0211***	0.0349***	0.0270***	0.0094	0.0119***	0.0003
	(0.002)	(0.007)	(0.003)	(0.008)	(0.002)	(0.008)
Higher Secondary	0.0119***	0.0411***	0.0181***	0.0020	0.0070***	-0.0096
	(0.002)	(0.006)	(0.003)	(0.007)	(0.001)	(0.007)
Graduate/University	-0.0384***	0.1579***	-0.0640***	-0.1091***	-0.0569***	0.0027
	(0.005)	(0.021)	(0.008)	(0.025)	(0.007)	(0.024)
Technical Degree	-0.0033***	0.0060*	-0.0021***	-0.0024	-0.0120***	-0.0028
	(0.001)	(0.004)	(0.001)	(0.004)	(0.002)	(0.004)
Experience	0.0111***	0.4924***	0.0016	0.1947***	0.0016	-0.1865***
	(0.004)	(0.046)	(0.006)	(0.056)	(0.004)	(0.053)
Occupation:						
High Skilled	0.0165***	-0.0685***	-0.0087***	-0.1308***	-0.0920***	0.0835***
	(0.004)	(0.021)	(0.003)	(0.025)	(0.006)	(0.024)
Medium Skilled	-0.0002	-0.0370***	0.0012	-0.1057***	0.0029	-0.0236**
	(0.002)	(0.010)	(0.001)	(0.013)	(0.002)	(0.012)
Low Skilled	0.0227***	0.0713***	-0.0200***	-0.0033	-0.0017	0.0058
	(0.005)	(0.014)	(0.004)	(0.017)	(0.006)	(0.016)
Industry:						
Construction	0.0169***	-0.0009	0.0205***	-0.0065***	0.0119***	0.0004
	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)
Low Tech Mfg.	0.0127***	-0.0579***	0.0112***	-0.0629***	0.0029***	-0.0365**
	(0.002)	(0.013)	(0.002)	(0.016)	(0.001)	(0.015)
High Tech Mfg.	0.0260***	-0.0036**	0.0258***	-0.0005	0.0180***	-0.0003
	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
Modern Services	-0.0357***	0.0856***	-0.0527***	0.0808***	-0.0110***	0.0230
	(0.003)	(0.025)	(0.005)	(0.031)	(0.002)	(0.029)
Computer & IT	-0.0044***	0.0040*	-0.0059***	-0.0005	-0.0029***	-0.0010
	(0.001)	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)
Social/Religion:						
Social/Religion	0.0072***	0.0166	0.0046***	0.0281	0.0007	0.0804***
	(0.001)	(0.022)	(0.001)	(0.027)	(0.002)	(0.025)
Region/Constant:						
Region	0.0025	0.0696**	-0.0017	0.0385	0.0091***	-0.0825**
	(0.003)	(0.032)	(0.003)	(0.040)	(0.003)	(0.038)
Constant		-0.3873***		0.6040***		0.3304***
		(0.081)		(0.098)		(0.094)
Total Effect	0.0961***	0.4698***	-0.0120	0.5120***	-0.0918***	0.1996***
	(0.008)	(0.016)	(0.011)	(0.019)	(0.012)	(0.018)

Standard errors are reported in parenthesis.
$*p < 0.10, **p < 0.05, ***p < 0.01$.

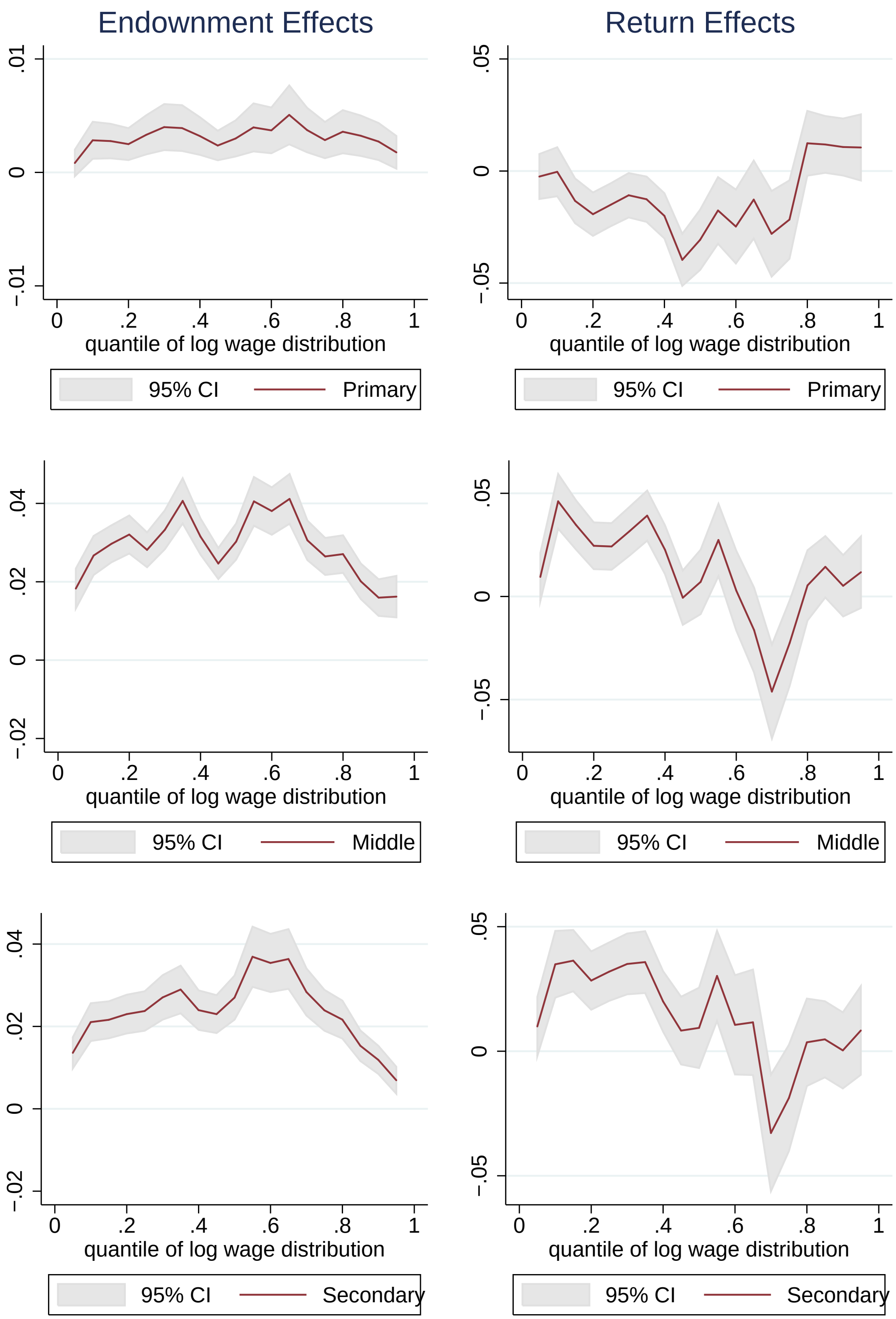

Figure 2: Endowment and return effects of education on the log gender wage gap (Y-axis) along the wage distribution (X-axis)

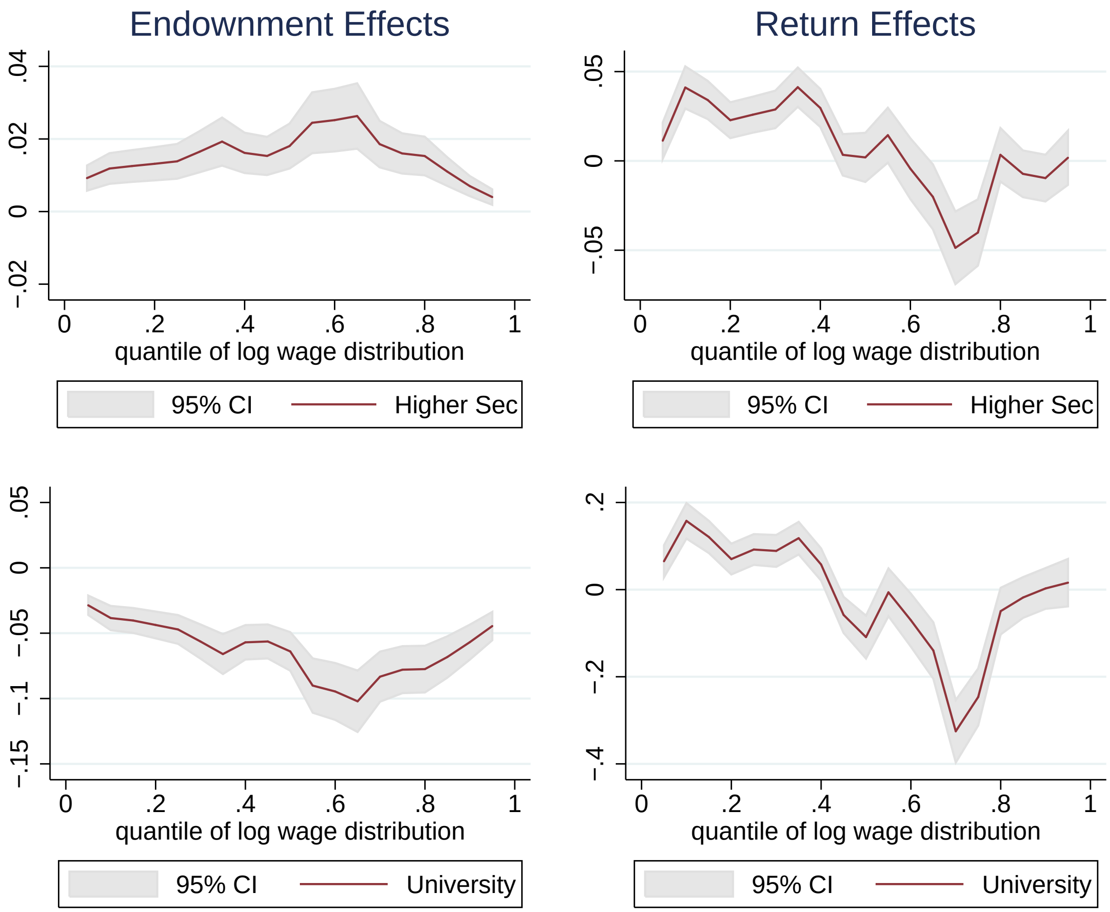

Figure 2: cont'd

1, 3, and 5 of table 2). The endowment effect of education is fairly constant across the wage distribution (figure 2, top left). In contrast, a higher return to primary education among women, relative to men, has a statistically significant narrowing effect on the male-female wage gap. The effect is largest around the median where higher returns to female primary education closes the gender gap by about 3% (figure 2, top right). Taken together the endowment and return effects of primary education do not appear to be offsetting. In fact, for most of the wage distribution, primary education among women has the potential to narrow the gender wage gap, thanks to its higher returns.

We also estimate that a lower prevalence of post-primary education (middle, secondary and higher-secondary school levels) among women, relative to men, has a statistically significant widening effect on the male-female wage gap (table 2 columns 1, 3, and 5; figure 2, middle and bottom left; figure 2 (cont'd), top left). Although this effect disadvantages women, its magnitude decreases at higher levels of post-primary education levels, and follows, roughly, an inverted u-shaped pattern across the wage distribution. The effect of gender differentials in returns to post-primary education on the wage gap is more complex (table 2 columns 2, 4, and 6; figure 2, middle and bottom right; figure 2 (cont'd), top right). Approximately below the median, lower returns to post-primary education among women widens the gender gap (by over 4%). Above the median, differential returns do not have a significant effect, except around the 70th quantile where they help reduce the wage gap by approximately 5%. In sum, for this middle group of education, a lack of female endowments and lower returns to education combine to disadvantage women at low wage levels, but higher returns to education for female at some higher levels of wages may, again, override this endowment effect.

We estimate that the greater prevalence of college/university among women, relative to men, has a statistically significant narrowing effect on the male-female wage gap (table 2, columns 1, 3, and 5; figure 2 (cont'd), bottom left). Greater graduate/university achievement among women reduces the gap in the left tail of the wage distribution by approximately 4% (table 2, column 1); the effect is magnified towards the median (table 2, column 2) and reaches approximately 11% at the 65th quantile, after which the effect decreases (see figure 2 (cont'd), bottom left). While graduate/university education as an endowment helps women at all points of the wage distribution, the returns to university education generally widens the gap below the 40th quantile and narrows the gap above the 40th quantile (figure 2 (cont'd), bottom right). Lower returns to university education works against women on the left tail of the distribution, making the gender gap wider by 16 percent (table 2, column 2). Above the 40th quantile the effect reduces the gender wage gap by almost 40% (at the 75th quantile), although the effect varies substantially along the wage distribution. Beyond the 40th quantile, endowment and return effects of higher education reinforce one another in helping women close the gender wage gap.

We present a novel way to examine the effect of education on the wage gap along the wage distribution. We use our estimates of the Firpo et al. (2007) decomposition to construct a surface for the endowment and return effects across education levels and quantiles of the log-wage distribution. This three-dimensional illustration of the education-quantile-gap relationship is depicted in figure 3. While the various panels in figure 2 represent cuts of figure 3 orthogonal to the education axis (i.e., holding education constant), figure 3 allows us to observe, at each quantile, the effect of increasing education on the components of the gender wage gap.

Holding the quantile of the log-wage distribution constant, generally we find a concave relationship between education and the endowment contribution to the gap. While endowments of university education contribute to reduce the gap, the endowment effect sharply increases and becomes positive as the education level decreases from university to higher secondary-middle education, and slightly decreases to just above zero for primary education. This suggests that policymakers should focus on policies that increase the female endowments of post-primary (but less than university) education, as women's insufficient endowments are contributing to the wage gap in these education levels.

While this inverted U-shape between education and endowment effect is relatively homogeneous across quantiles, the shape of the relationship between education and return effects is generally increasing for low quantiles, but decreasing for high quantiles. Holding quantiles below the 40th quantile fixed, return effects tend to increase (widening the gap) as education increases from primary to university. For the region between the 40th and the 85th quantiles, we observe an inverted-U shape relationship between education and the return effect, with a sharp decrease in the return effect at university level education. For the top quantiles (above 85th), the return effect is generally orthogonal to education and hovers around zero. In other words, our results suggest that, for women at the very top of the wage distribution, education does not have a "structural" effect on the gender wage gap, and the lack of endowments at the primary - secondary education level should be the target of education endowment policies for this group.

In general, we do not find a statistically significant experience endowment effect on the gender wage gap, with exception of the bottom of the wage distribution, where experience endowments widen the gap (see column 1 of table 2 and figure 4, left). On the other hand, we find high levels of heterogeneity in how returns to experience affect the wage gap. Below the 60th quantile, experience returns significantly contribute to increasing the gap (see columns 2 and 4 of table 2 and figure 4, right). However, beyond the 60th quantile, the return effect of experience makes the gap narrower (see column 6 of table 2 and figure 4, right). This pattern is generally the same of that displayed by higher secondary and university educations.[13]

Among the occupation categories (table 2), women face a wage advantage at the median and 90th quantile due to higher participation rates in high skilled occupations (columns 3 and 5). They face higher returns in these occupations, relative to men, in all parts of

[13] As discussed above, it is important to acknowledge that our experience variable captures potential experience. In fact, this is a limitation faced by most papers in the literature. It is possible that our experience variable overstates the market work experience of women since women may face more work interruptions than men. Hence, the results of this section should be interpreted keeping this caveat in mind. The next section incorporates the effect of labor market selection by gender into our unconditional quantile decompositions.

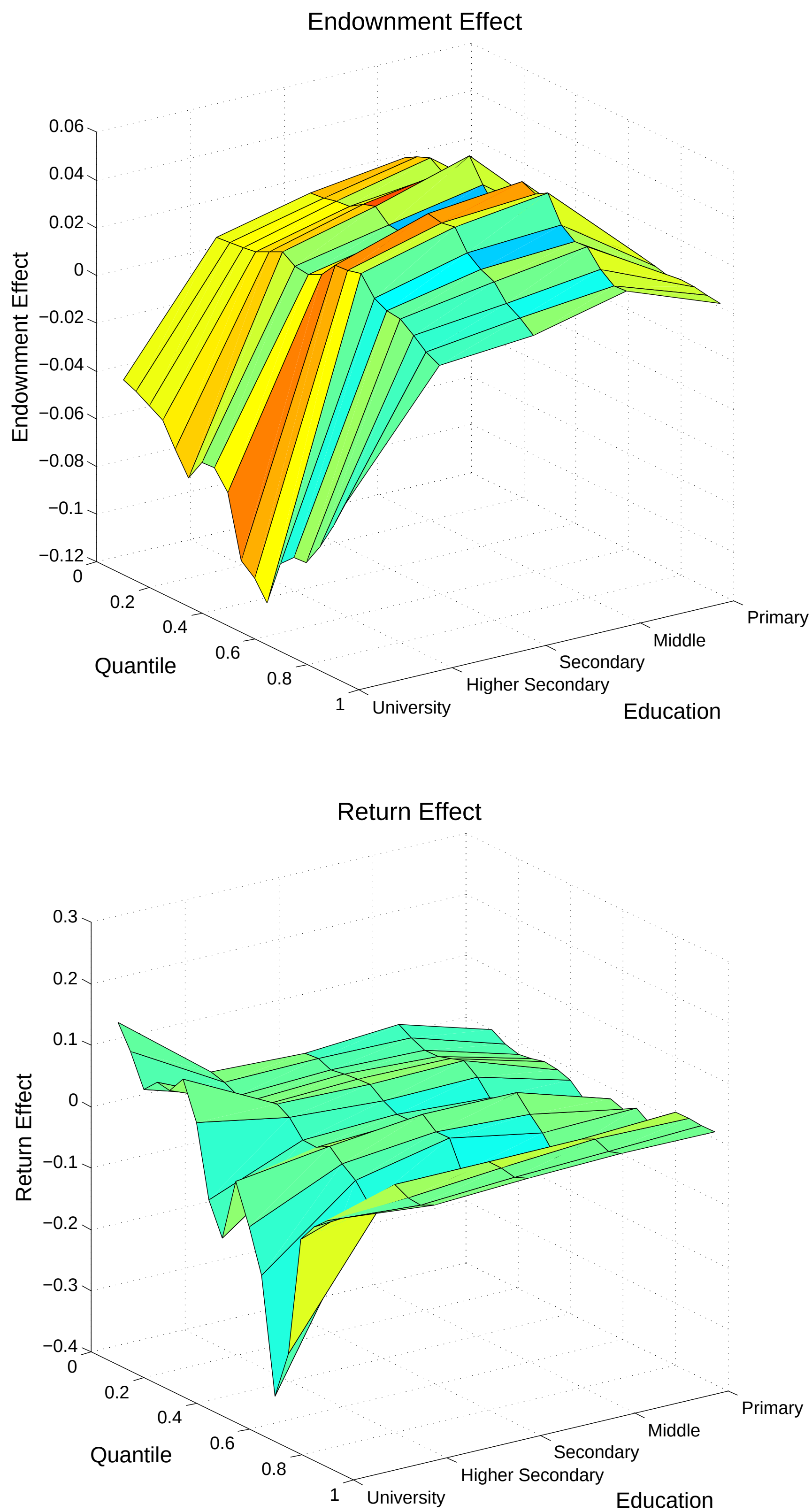

Figure 3: Three-dimensional representation of the effect of education on the log gender wage gap along the wage distribution

the distribution (see columns 2 and 4) except the top (see column 6) reflecting the glass ceiling effect. Overall the return effects are larger than the endowment effects. Among the industry categories greater participation of women in the modern services and computer and IT services has helped narrow the gender wage gap (columns 1, 3, and 5). At the 90th quantile, return differentials among industry categories are mostly insignificant along the wage distribution. Social and religious factors matter to the wage gap, but below and at the median primarily through endowment effects, and above the median primarily through return effects. Regional determinants also contribute to the gender wage gap in India.

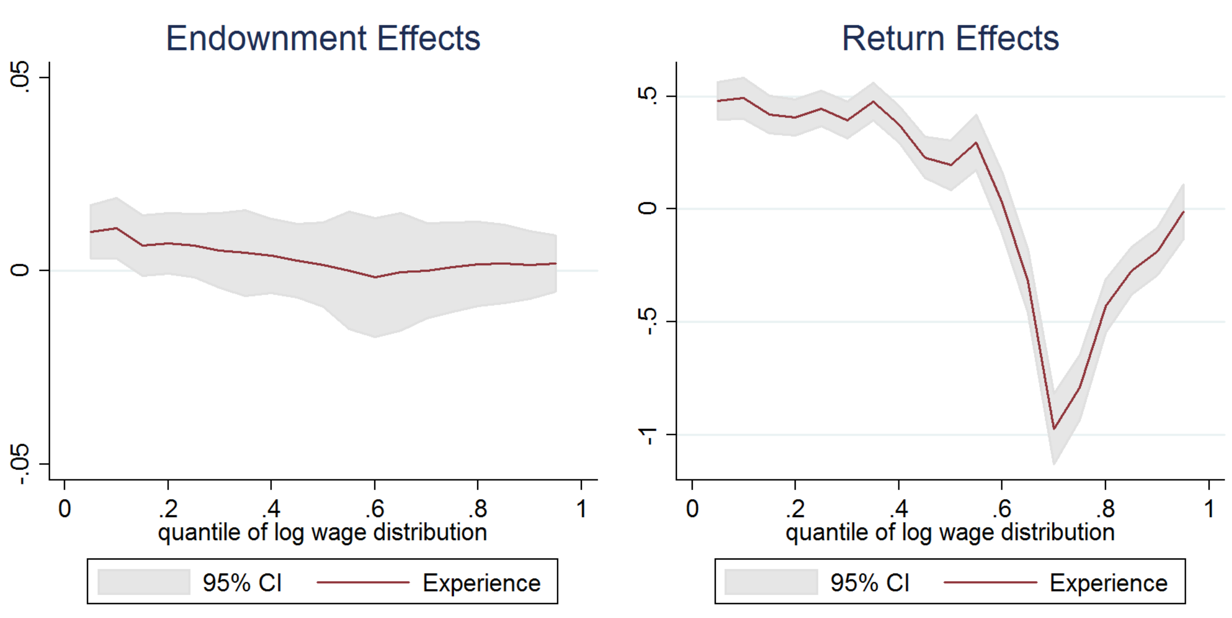

Figure 4: Endowment and return effects of experience on the log gender wage gap (Y-axis) along the wage distribution (X-axis)

6 Robustness of UQR decompositions to selection into labor markets

This section examines the robustness of our results to self-selection into the formal job market. Note that while the NNS data contains information on 377,377 individuals, our wage employment sample has only a fraction of these individuals (31,541 observations). Therefore, individuals and data that we observe can be different than individuals and data that we do not observe due to self-selection. There are two possible selectivity layers: the selection of individuals into the labor market and, among those, the selection into wage employment. Further, the selectivity patterns could manifest differently among men and women.

To address selection we revisit the human capital earnings functions that underlie equation (2). We write the log daily wages earned by an individual of gender $g = m, f$ in the regular wage employment sector, R, as

$$W_{gR} = X_g \alpha_g + e_{gR} \tag{6}$$

where X_g denotes our set of endowment variables.[14] Accounting for selection in equation (6) is important for obtaining consistent estimates of the returns to education and experience.

We address the concerns about multiple layers of selectivity separately for men and women by dividing the population into 3 sectors: i. regular wage employment (R); ii. self-employed and casual wage employed (S); and iii; out of the labor force and unemployed (O). We follow a multinomial selection approach (Lee, 1983) that specifies an individuals' discrete choice of labor market sector j based on a random utility model $U_{gj} = z_g \beta_g + \mu_{gj}$, $j = R, S, O$.[15] Wages are only observed in our analysis when the regular wage employment (R) is chosen. Specifying the choice model as a multinomial logit, selection due to each labor market alternative in equation (6) can be accounted for by augmenting equation (6) with a series of conditional expectations as additional explanatory variables.

These selection correction terms represent expectations of the error term in equation (6), e_{gR}, conditional on each alternative being chosen, $E_j(e_{gR}|\mu_{gj})$. They are computed using a variant of the Dubin and McFadden (1984) model following the approach of Bourguignon et al. (2007). The selectivity corrected wage earning function is thus:

$$W_{gR} = X_g \alpha_g + \sigma \Big(\rho_R \big[E_R(e_{gR}|\mu_{gR}) \big] + \rho_S \big[E_S(e_{gR}|\mu_{gS}) \big] + \rho_O \big[E_O(e_{gR}|\mu_{gO}) \big] \Big) + \varepsilon_{gR} \tag{7}$$

where the terms in the square brackets are the selection correction terms. Selection effects related to each sector j are captured by coefficients, $\sigma\rho_j$ where σ is the standard deviation of the errors in equation (6). The model is identified by excluding at least one of the variables in X from Z, and by the nonlinearity of functional forms used.[16]

[14] We also estimate baseline models (with and without selection) where X_g contains only human capital variables. Results are similar and available upon request.

[15] The supplemental information appendix (A5) offers further discussion about our approach to control for selection.

[16] We identify the selection equation using a dummy variable for landownership, marital status, number of children in the household, number of adults in the household, whether the individual is the head of the household, and whether they are recipients of ration cards that are based on pre-determined poverty status. We expect these variables to affect participation in the j-th labor market but not wages.

In equation (7), the presence of self-selection is indicated by the significance of ρ_R, the coefficient on the first selection term. The interpretation of the selection effect is similar to that in the Heckman model; that is, unobserved productivity of individuals increases their probability of entering the regular wage sector. Further, a positive ρ_R, indicates that the unobservable productivity is valued in the labor market and that the selection of individuals based on the unobservable has led to higher than expected wages in the sector. The role of the non-wage labor market sectors (self-employment and casual wage) in selectivity is captured by ρ_S. The significance of the coefficients ρ_S suggests that the selected sample in the regular wage market is linked to people reallocating to self-employment/casual work out of regular wage work. A positive ρ_S indicates that the reallocation is by less able or productive people whose exit from the regular wage sector increases wages on average. Conversely, a significant and negative coefficient on ρ_S signals lower wages in the regular wage sector than expected because of the selection of more able people out of regular wage and into self-employment.

Given the above selection model, our approach consists of two steps. In step 1, we estimate returns to human capital controlling for selection using equation (7). In step 2, we re-examine the decomposition exercises of the previous section treating the selection variables for men and women as attributes in UQRs decompositions based on RIF regressions.

Table 3 presents the selection coefficients obtained in step 1. The results show strong selection effects in both male and female labor markets. Both females and males face negative selection – individuals have selected out of wage earning based on unobservables that are associated with lower average wages (Table 3, row 1). The positive selection coefficient ρ_S for women suggests that the reallocation of some women out of regular wage into self-employment/casual wage has resulted in higher than average wages in the women's regular wage sector. This suggests that some low ability women move into self-employment/casual wage rather than regular wage employment. However, the coefficient ρ_S is imprecisely estimated and not statistically significant in the female regression. The opposite pattern holds true for men (Table 3, row 2). The negative selection coefficient ρ_O for women suggests that the reallocation of some women out of regular wage into unemployment and out of the labor force has resulted in lower wages for women in the regular wage earning sector (Table 3, row 3). This is consistent with some high ability women being shut out of the wage sector into unemployment or out of the labor force. There is no evidence of such effects for men.

Table 4 presents the results of step 2; unconditional wage gap decompositions with selection controls. We keep the scope of our analysis and focus on comparing the human capital coefficients with (Table 4) and without (Table 2) selectivity. Clearly, the coefficient

Table 3: Selection Coefficients – Equation 7

Selection coefficients	Female	Male
Regular wage ($\sigma\rho_R$)	-0.127**	-0.274***
	(0.050)	(0.025)
Self employed/casual wage ($\sigma\rho_S$)	0.195	-0.289***
	(0.137)	(0.067)
Out of labor market/unemployed ($\sigma\rho_O$)	-0.447**	0.198***
	(0.214)	(0.054)

Bootstrapped standard errors are reported in parenthesis.
Full results available upon request.
$*p < 0.10, **p < 0.05, ***p < 0.01$.

estimates are qualitatively identical throughout with a few exceptions. When controlling for selection, the return effects of education variables at the 50th and 90th quantiles significantly contribute to increasing the gender wage gap. These coefficients, however, are mainly statistically insignificant in the model without selection effects (Table 2). Another (subtle) difference is that primary education does not significantly affect the gap in the selection model (Table 4), while small effects are identified without selection controls (Table 2). Finally, for most education variables, endowment effects for the 10th quantile in the selection model tend to be slightly lower than our initial estimates. These findings allow us to conclude that while selection effects are a feature of Indian labor markets, there is much to be learned from a deeper analysis focused on the nature of selection effects among various subsectors. However, in the context of our study we conclude that the qualitative results of our decomposition analysis are similar with and without selection controls with the exceptions pointed out above.

7 Discussion

The paper contributes to the literature on gender inequalities on labor markets by exploring two types of heterogeneities in wage earning regressions: i) heterogeneity on the effects of education and experience on expected earnings along the distribution of these endowments (RHS heterogeneity); and ii) heterogeneity on the contributions of education and experience to the gender wage gap at different quantiles of the wage distribution (LHS heterogeneity). Our approach to tackle RHS heterogeneity is to use cubic splines to estimate non-linear returns to human capital on expected earnings of men and women. Next, we change the perspective from an approach with a non-linear expected earnings function to a linear function

Table 4: Wage decomposition between women and men at selected quantiles with selection controls

	10th Quantile		50th Quantile		90th Quantile	
	Endowment Effect	Return Effect	Endowment Effect	Return Effect	Endowment Effect	Return Effect
	(1)	(2)	(3)	(4)	(5)	(6)
Human Capital:						
Primary	0.0009	-0.0029	0.0021	0.0034	0.0033	0.0097
	(0.002)	(0.013)	(0.002)	(0.012)	(0.002)	(0.012)
Secondary	0.0142**	0.0446**	0.0244***	0.0578***	0.0179**	0.0495**
	(0.006)	(0.021)	(0.005)	(0.021)	(0.007)	(0.019)
Middle	0.0126***	0.0392**	0.0228***	0.0607***	0.0116***	0.0371**
	(0.004)	(0.018)	(0.006)	(0.020)	(0.004)	(0.018)
Higher Secondary	0.0075**	0.0484***	0.0157***	0.0756***	0.0063**	0.0445***
	(0.003)	(0.015)	(0.005)	(0.017)	(0.003)	(0.014)
Graduate/University	-0.0259***	0.2104***	-0.0562***	0.3749***	-0.0492***	0.3369***
	(0.008)	(0.073)	(0.015)	(0.083)	(0.014)	(0.066)
Technical Degree	-0.0032**	0.0017	-0.0021	-0.0006	-0.0122**	0.0204**
	(0.001)	(0.005)	(0.002)	(0.006)	(0.005)	(0.008)
Experience	0.0076	0.3268**	-0.0001	0.5529***	0.0042	0.5745***
	(0.006)	(0.136)	(0.011)	(0.123)	(0.012)	(0.127)
Selectivity:						
Regular wage	0.3117***	-0.2847	0.1180**	0.1787	-0.1553***	0.8326***
	(0.098)	(0.348)	(0.053)	(0.291)	(0.059)	(0.289)
Self employed/casual wage	0.2368**	0.3332	0.0109	0.1509	-0.4924***	-0.2554
	(0.112)	(0.225)	(0.072)	(0.198)	(0.105)	(0.211)
Out of labor market/unemp.	0.2519	-0.8075*	0.1712**	-0.7072*	-0.1270	-0.3364
	(0.154)	(0.423)	(0.082)	(0.377)	(0.107)	(0.311)
Occupation:						
High Skilled	0.0164*	-0.0660	-0.0099	-0.0108	-0.0959***	0.1699***
	(0.009)	(0.048)	(0.008)	(0.046)	(0.017)	(0.049)
Medium Skilled	0.0015	-0.0319*	0.0023	-0.0303	0.0034	-0.0279
	(0.005)	(0.019)	(0.003)	(0.020)	(0.003)	(0.019)
Low Skilled	0.0261**	0.0724	-0.0179*	0.0432	-0.0010	0.0545
	(0.012)	(0.060)	(0.010)	(0.054)	(0.008)	(0.058)
Industry:						
Construction	0.0162***	-0.0011	0.0203***	-0.0001	0.0127***	-0.0019
	(0.004)	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)
Low Tech Mfg.	0.0121**	-0.0636	0.0109**	-0.0683	0.0030	-0.0987*
	(0.005)	(0.056)	(0.005)	(0.050)	(0.002)	(0.053)
High Tech Mfg.	0.0251***	-0.0037	0.0253***	-0.0036	0.0183***	-0.0045
	(0.005)	(0.003)	(0.004)	(0.003)	(0.005)	(0.003)
Modern Services	-0.0352***	0.0707	-0.0526***	0.1790**	-0.0119***	-0.0740
	(0.008)	(0.079)	(0.012)	(0.077)	(0.004)	(0.076)
Computer & IT	-0.0047**	0.0046*	-0.0060*	0.0076**	-0.0025	-0.0007
	(0.002)	(0.003)	(0.003)	(0.003)	(0.002)	(0.004)
Social/Religion:						
Social/Religion	0.0089**	0.0014	0.0054*	-0.0281	0.0001	-0.0284
	(0.004)	(0.049)	(0.003)	(0.044)	(0.003)	(0.049)
Region/Constant:						
Region	0.0041	0.0965	-0.0008	0.1765	0.0121**	0.0201
	(0.006)	(0.167)	(0.006)	(0.149)	(0.006)	(0.164)
Constant		-0.3070		0.1588		2.2094***
		(0.744)		(0.653)		(0.688)
Total Effect	0.8845***	-0.3186*	0.2838***	1.1708***	-0.8545***	3.5311***
	(0.149)	(0.178)	(0.101)	(0.107)	(0.120)	(0.136)

Standard errors are reported in parenthesis.
$*p < 0.10, **p < 0.05, ***p < 0.01$.

approach that examines earnings differentials along the wage distribution. To explore LHS heterogeneity, we use unconditional quantile regressions to decompose the gender wage gap into endowment and return effects at various quantiles of the wage distribution.

Our analysis of RHS heterogeneity reveals that returns to education vary significantly both along their distributions but also between men and women. We find that, at the bottom of the education distribution, returns to education on women's expected wages are decreasing. This result indicates that Indian labor markets create gender distortions when it comes to rewarding early education efforts, and education policy should focus on understanding and addressing this issue. On the other hand, women's marginal return of education on expected earnings at the top of the education distribution is greater than that of men. This result suggests that, holding other factors constant, policies that facilitate access of women to higher education may be proven fruitful in reducing the gender wage gap among highly educated workers. We also find that the interaction of education and experience has a constant effect on men's wages but an increasing effect on women's. This result suggests that experience and education are substitute inputs for men and complement inputs for women.

Our UQR-based decompositions uncover significant LHS heterogeneities on the relationship between education and gender earnings differentials. In general, for pre-university educational levels, endowment differences increase the gap along the entire distribution, but with a larger effect just above the median. However, women's larger university education endowment decreases the gender wage gap along the entire distribution. In terms of return effects to post-primary education, we generally find a narrowing effect for women with earnings above the median. However, for primary education, the return effect is generally negative below the 80th quantile, but positive above.

A central finding of the quantile decomposition analysis is the heterogeneity in the educational endowment and returns effects across the wage distribution. We show, for example, that the return effects of both education and experience fall sharply after the median but then almost recover. This reveals that, even though education and experience contribute to a decrease of the gender wage gap as we move up from the bottom of the wage distribution, at the very top, return effects increase indicating that women labor markets in India have important glass-ceilings. This asymmetry is relevant for policy makers in terms of devising effective means for narrowing the wage gap. A conventional decomposition at the mean would miss this with its 'one-size-fits-all' policy implications.

The wage gap decompositions show varying contributions of education, and returns thereof, to the wage gap thus again suggesting an important role for education policy in reducing the gender wage gap. But in considering impacts of investments in education, the LHS heterogeneity evident in our results suggests that there are specific cohorts who are likely to benefit more than others. For example, at the primary level, increasing education among women will likely contribute significantly to reducing the wage gap through changing educational endowments, which in turn, results in higher returns. But such results are not likely to be repeated for post-primary education, except for those who are among the highest wage earners. Moreover, though investing more in the university/college education of women would likely decrease the overall mean wage gap, maximum effects occur at the highest wage levels, among women who are already doing quite well relative to men.

Note that our policy conclusions hold general equilibrium effects constant. That is, the human capital earnings functions used in our analyses represent a partial equilibrium representation of education choices in the labor market. Specifically, if government policy were to encourage greater post-secondary education then the increased supply of women with that level of education is likely to lower the education returns at that level possibly offsetting part of the intended effect of the policy. The education returns are informative and can be thought of as exogenous, however, for a cohort of young adults who are about to make their education decisions, conditional on their family backgrounds, institutional characteristics and economic conditions (Card, 1999).

As our paper sets forth to examine heterogeneous effects of education on wages along the distribution of both variables, it faces an important limitation that both LHS and RHS heterogeneity are not addressed in the same empirical framework. Our first analysis uncovers heterogeneous effects along the distribution of education, but these estimates relate to expected earnings. Our second analysis uncovers heterogeneous effects of education on different quantiles of wages, but these estimates rely on a linear wage earning function.

Nevertheless, we are able to detect common trends when exploring these heterogeneities in their own models. For instance, the university educational sector seems to include cohorts of women that are performing well in labor markets. At this level, returns to education are relatively high, increasing, and convex. Also, university education endowments favor women in all wage quantiles, while return effects help to reduce gaps for women on the higher wage quantiles. For women with small education endowments, our non-linear regression shows decreasing returns to education while, for the same range, men's increasingly benefit from education. The UQR estimates indicate that the contribution of primary education to wage gaps is due to endowment gaps as opposed to "structural" gaps.

References

Ahmed, S. and P. Maitra (2015). A distributional analysis of the gender wage gap in Bangladesh. *Journal of Development Studies 51*(11), 1444–1458.

Aslam, M., F. Bari, and G. Kingdon (2012). Returns to schooling, ability and cognitive skills in Pakistan. *Education Economics 20*(2), 139–173.

Baker, M. and M. Drolet (2010). A new view of the male/female pay gap. *Canadian Public Policy 36*(4), 429–464.

Bayer, A. and C. E. Rouse (2016). Diversity in the economics profession: A new attack on an old problem. *Journal of Economic Perspectives 30*(4), 221–242.

Bhaumik, S. K. and M. Chakrabarty (2008). Does move to market have an impact on earnings gap across gender? some evidence from india. *Applied Economics Letters 15*(8), 601–605.

Blau, F. D. and L. M. Kahn (2000). Gender differences in pay. *Journal of Economic Perspectives 14*(4), 75–99.

Blau, F. D. and L. M. Kahn (2016). The gender wage gap: Extent, trends, and explanations. Technical report, National Bureau of Economic Research.

Bourguignon, F., M. Fournier, and M. Gurgand (2007). Selection bias corrections based on the multinomial logit model: Monte carlo comparisons. *Journal of Economic Surveys 21*(1), 174–205.

Card, D. (1999). The causal effect of education on earnings. *Handbook of Labor Economics 3*, 1801–1863.

Chi, W. and B. Li (2008). Glass ceiling or sticky floor? examining the gender earnings differential across the earnings distribution in urban china, 1987–2004. *Journal of Comparative Economics 36*(2), 243–263.

DiNardo, J., N. M. Fortin, and T. Lemieux (1996). Labor market institutions and the distribution of wages, 1973-1992: A semiparametric approach. *Econometrica 64*(5), 1001–1044.

Dubin, J. A. and D. L. McFadden (1984). An econometric analysis of residential electric appliance holdings and consumption. *Econometrica*, 345–362.

Duraisamy, M. and P. Duraisamy (1996). Sex discrimination in Indian labor markets. *Feminist Economics 2*(2), 41–61.

Duraisamy, P. (2002). Changes in returns to education in India, 1983–94: by gender, age-cohort and location. *Economics of Education Review 21*(6), 609–622.

Firpo, S., N. Fortin, and T. Lemieux (2007). Decomposing wage distributions using recentered influence function regressions. *Mimeo. University of British Columbia (June).*

Firpo, S., N. M. Fortin, and T. Lemieux (2009). Unconditional quantile regressions. *Econometrica 77*(3), 953–973.

Gates, M. F. (2014). Putting women and girls at the center of development. *Science 345*(6202), 1273–1275.

Hanushek, E. A. and D. D. Kimko (2000). Schooling, labor-force quality, and the growth of nations. *American Economic Review*, 1184–1208.

Harmon, C., V. Hogan, and I. Walker (2003). Dispersion in the economic return to schooling. *Labour Economics 10*(2), 205–214.

Heckman, J. J. and E. Vytlacil (2005). Structural equations, treatment effects, and econometric policy evaluation. *Econometrica 73*(3), 669–738.

Ito, T. (2009). Caste discrimination and transaction costs in the labor market: Evidence from rural north India. *Journal of Development Economics 88*(2), 292–300.

Jann, B. (2008). The blinderoaxaca decomposition for linear regression models. *Stata Journal 8*(4), 453–479.

Juhn, C., K. M. Murphy, and B. Pierce (1993). Wage inequality and the rise in returns to skill. *Journal of Political Economy*, 410–442.

Kingdon, G. G. (1998). Does the labour market explain lower female schooling in India? *Journal of Development Studies 35*(1), 39–65.

Kingdon, G. G. and J. Unni (2001). Education and women's labour market outcomes in India. *Education Economics 9*(2), 173–195.

Kishor, S. (1993). "May God Give Sons to All": Gender and child mortality in India. *American Sociological Review*, 247–265.

Koenker, R. and G. Bassett (1978). Regression quantiles. *Econometrica 46*(1), 33–50.

Koop, G. and J. L. Tobias (2004). Learning about heterogeneity in returns to schooling. *Journal of Applied Econometrics 19*(7), 827–849.

Lee, L.-F. (1983). Generalized econometric models with selectivity. *Econometrica*, 507–512.

Lemieux, T. (2006). Postsecondary education and increasing wage inequality. *American Economic Review 96*(2), 195–199.

Levin, H. M. (2009). The economic payoff to investing in educational justice. *Educational Researcher 38*(1), 5–20.

Machado, J. A. and J. Mata (2002). Earning functions in Portugal 1982–1994: Evidence from quantile regressions. In *Economic applications of quantile regression, Studies in Empirical Economics*, pp. 115–134. Springer.

Machado, J. A. and J. Mata (2005). Counterfactual decomposition of changes in wage distributions using quantile regression. *Journal of Applied Econometrics 20*(4), 445–465.

Manning, A. and H. Robinson (2004). Something in the way she moves: a fresh look at an old gap. *Oxford Economic Papers 56*(2), 169–188.

Melly, B. (2005). Decomposition of differences in distribution using quantile regression. *Labour Economics 12*(4), 577–590.

Menon, N. and Y. Van der Meulen Rodgers (2009). International trade and the gender wage gap: New evidence from India's manufacturing sector. *World Development 37*(5), 965–981.

Poirier, D. J. and S. G. Garber (1974). The determinants of aerospace profit rates 1951-1971. *Southern Economic Journal*, 228–238.

Rosenzweig, M. R. and T. P. Schultz (1982). Market opportunities, genetic endowments, and intrafamily resource distribution: Child survival in rural India. *American Economic Review 72*(4), 803–815.

Royston, P. and W. Sauerbrei (2004). A new approach to modelling interactions between treatment and continuous covariates in clinical trials by using fractional polynomials. *Statistics in Medicine 23*(16), 2509–2525.

Royston, P. and W. Sauerbrei (2008). *Multivariable model-building: a pragmatic approach to regression anaylsis based on fractional polynomials for modelling continuous variables*, Volume 777. John Wiley & Sons.

Söderbom, M., F. Teal, A. Wambugu, and G. Kahyarara (2006). The dynamics of returns to education in Kenyan and Tanzanian manufacturing. *Oxford Bulletin of Economics and Statistics 68*(3), 261–288.

Sperling, G. B. and B. Herz (2004). What works in girls education: Evidence and policies from the developing world. *Washington, DC: Council on Foreign Relations*.

United States Agency for International Development (2012). Food security and gender. Technical report.

Weichselbaumer, D. and R. Winter-Ebmer (2005). A meta-analysis of the international gender wage gap. *Journal of Economic Surveys 19*(3), 479–511.

World Bank (2011). World Development Report 2012: Gender equality and development. Washington D.C. Technical report.

Zeng, J., X. Pang, L. Zhang, A. Medina, and S. Rozelle (2014). Gender inequality in education in China: a meta-regression analysis. *Contemporary Economic Policy 32*(2), 474–491.