ABSTRACT

Title of Dissertation: CHANGES IN THE WAGE GAP OF GENDER AND CASTE GROUPS IN INDIA

Marilyn Jacob, Doctor of Philosophy, 2006

Directed By: Professor Seth Sanders, Department of Economics

We explore the changes in the wage gap of caste and gender groups in India. Traditional Hindu society divided people into social classes based on the caste system. The lowest of the castes have traditionally been economically disadvantaged. Women in India have typically been restricted to the household and their participation in the formal labor market has begun expanding only recently. We explore the changes that these two groups have experienced over the years using a nationally representative dataset.

In the second chapter we decompose the wage gaps of these groups into explained and unexplained components based on the Blinder-Oaxaca (1973) decomposition technique. Our contribution to the literature here is the extension of

the analysis of discrimination to a society with a clearly established social hierarchy. We find that the gross wage gap has reduced over this period, and the extent of the gap attributable to discrimination has decreased over time. We further decompose the wage gap into components attributable to wage differences and occupational differences based on Brown et al. (1980). We find that the wage discrimination component has decreased over time and the job discrimination component is statistically insignificant.

In the third chapter we investigate whether there have been beneficial wage gains for women and lower castes because of increased competition following liberalization of trade in India. Based on Becker's model of taste-based employer discrimination, it is expected that as an economy becomes more competitive, employer discrimination should decline. The trade liberalization reforms that began in 1991 in India increased competition by lowering protection in certain manufacturing industries. Firms who could indulge a taste for discrimination when trade protection allowed supernormal profits may not have been able to continue to do so as competition eliminated such profits. Using individual-level data and tariff data from pre- and post-reform periods, we find that wage differences reduced for female workers relative to male workers in the more open manufacturing sector industries. However, there is no significant effect on the wage differential between low and high caste workers.

CHANGES IN THE WAGE GAP OF GENDER AND CASTE GROUPS IN INDIA

By

Marilyn Jacob

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2006

Advisory Committee:
Professor Seth Sanders, Chair
Professor Roger Betancourt
Assistant Professor Rodrigo Soares
Assistant Professor Chris McKelvey
Associate Professor Shawn Bushway

Dedication

I dedicate this thesis to my parents.

Acknowledgements

I am thankful to the guidance and support provided by Professors Seth Sanders, Roger Betancourt and Rodrigo Soares. I am grateful to May Boggess of Stata Corporation for help with programming, and to Utsav Kumar for providing concordance tables for manufacturing industries. I would also like to thank Judy Hellerstein, Jeffery Smith, Mark Duggan, Utsav Kumar, Beomsoo Kim and Julian Cristia for their comments and suggestions. I am especially grateful to Ashok Jacob, Jose John, John Joseph, KPR Nair and Renuka Ravindran.

Table of Contents

List of Tables

List of Figures

Chapter 1 Introduction

We examine the changes in the wage gaps of two groups in India, namely caste and gender groups over a period of roughly two decades. This time period has been marked by substantial changes in the policy outlook of successive governments. Traditional Hindu society divided people into social classes based on the caste system. The lowest of the castes have traditionally been economically disadvantaged. Women in India have typically been restricted to the household and their participation in the formal labor market has begun expanding only recently. We explore changes that these two groups have experienced over the years using a nationally representative dataset spanning two decades.

The second chapter contains a descriptive analysis of the actual wage gaps among these groups during this period. Using individual level data from the *National Sample Survey Organization* (NSSO) for the years 1983, 1987-88, 1993-94 and 1999-2000, we apply the Blinder-Oaxaca (1973) decomposition technique to decompose the wage gap into explained and unexplained components. We further use the Brown et al. (1980) decomposition technique to examine if there is discrimination in the occupational attainment process itself which would explain the wage differences between male and female workers and between low and high caste workers. We expect that both affirmative action programs, to benefit the lower castes, and urbanization, which provides greater anonymity to workers, would result in lower wage gaps over time. We find that the wage gaps decrease over time and that occupational discrimination is smaller in magnitude (and statistically insignificant) compared to wage discrimination within occupations. We find that most of the wage

gap between low and high caste workers is explained by differences in productive characteristics. On the other hand, almost half the wage gap between male and female workers is unexplained by differences in productive characteristics. For this reason, Chapter 3 focuses on examining how increased competition affected relative wages of female workers.

The third chapter examines the impact of the trade reforms of 1991 by directly examining the industries where tariffs were reduced as a part of the reforms. Based on Becker's model of taste-based employer discrimination, it is expected that as an economy becomes more competitive, employer discrimination should decline. We find, as expected, that there is a greater reduction in the wage gap in industries that experienced greater reductions in tariffs. We find that competition in fact did raise women's wages relative to men's wages. For example, the reduction in protection of manufacturing industries reduced the wage gap between men and women and accounts for roughly 20% of the narrowing of the wage gap over the time period in which the trade reforms were implemented. The effect for lower castes was not found to be important which is consistent with other programs already guaranteeing a certain level of protection from discrimination and consistent with the results in Chapter 2 that suggest that it was mostly productive differences between these workers and others that caused their lower earnings. Chapter 4 concludes.

In the remainder of this chapter, we present an overview of the Indian economy, the industrial and trade policies and the aforementioned reforms, and the economic and social position of castes and women in India.

1.1 Overview of the Indian Economy

Following independence in 1947, Indian leaders adopted a mixed economy framework in order to achieve an industrialized economy while embracing goals of achieving equality. Since private entrepreneurs couldn't be entrusted with the responsibility of caring for the needs of the multitudes, the government devised a means to control the private sector. The government set up a licensing system which regulated private production and reserved production in certain strategic areas for the government. The government therefore allowed for private production while adopting the central planning method of socialist regimes. The other concern that faced the government was developing the means to achieve greater equality between different groups of individuals.

1.1.1 Industrial Policy: Five Year Plans and Licensing Regime

India embarked on the path to industrialization and growth by adopting five year plans in the model of the socialist economies. The Planning Commission was set up in March 1950 to assess the material, human and capital resources and develop plans to use them in a balanced and efficient manner. The second five year plan is largely considered the most significant plan which directed greater focus on developing heavy industry and infrastructural support. The primary goals were to achieve higher national income and growth in employment opportunities to cater to the larger labor force generated by the increasing population, in addition to developing the industrial base. Subsequent plans focused on agricultural growth and poverty reduction among other goals.

The political leaders of the time under Jawaharlal Nehru desired to make India an industrial success and aimed to achieve this through an active role for the government. Nehru desired that the *commanding heights* of the economy should be in the public sector. In an attempt to develop a large machine-building industry, the government adopted a policy of import substitution and regulated private sector. The Industries Development and Regulation Act of 1951 installed the licensing system which required an entrepreneur to obtain a license to set up a new unit, expand an existing one, or to change the product mix. Licensing was to allow for better planned investment, prevent concentration of industrial power in the hands of a few, maintain regional balance of industries, protect small scale producers, encourage entry of new entrepreneurs, etc. However, bureaucratic procedures to get a license involved long delays while awaiting permission from several sources, which in turn opened up the opportunity to corrupt officials to speed-up the process as well as to obtain additional licenses. The system therefore ended up suppressing competition and entrepreneurship while promoting monopolies, which was exactly what it intended to avoid (Das, 2002). Thus the entire manufacturing sector was effectively controlled by the government which had discretionary powers regarding granting licenses to producers. There was however a gradual shift in industrial policy in the 1980s with an emphasis on cost-efficiency in Indian industry through domestic competition.

1.1.2 Trade Policy

On the trade front, export pessimism developed from the belief that exports of primary products (which were India's main exports) would face adverse terms of trade in the world market; and a fledgling domestic industry meant that exports of

manufactured goods (which stood to gain in the world market) would take time to develop. The belief was that a viable balance of payments account would require minimizing imports, which led to the adoption of import substitution policies during the late 1950s. Imports were limited through quantitative restrictions and high tariffs.

But by the early 1960s, planners recognized the importance of exports and adopted a number of export promotion measures. In addition there were attempts towards liberalization in the 1980s. This involved reductions in quantitative restrictions on imports, but were accompanied by increases in tariffs which effectively afforded greater protection to the domestic industry.

1.1.3 Crisis and Economic Reforms of 1991

A series of unfavorable domestic and international developments created the threat of an economic crisis in June 1991. The government (central, state and union territories) fiscal deficit as a percentage of Gross Domestic Product (GDP) rose from 9% in 1980-81 to 12% in 1990-91. The current account deficit as a ratio of GDP rose from 1.3% during 1985-90 to 3.3% in 1990-91. In addition the current account deficit was increasingly being funded by more expensive external commercial borrowings. The debt-service ratio (ratio of repaid interest on external debt to current receipts on the Balance of Payments) had increased from 10% in 1980-81 to 30% in 1990-91. The Gulf crisis and resulting uncertainties in oil prices edged India closer to a crisis combined with the political uncertainties at the time with frequent changes in the Central government. Foreign exchange reserves plummeted to about one billion dollars and India was on the verge of defaulting on its external debt. International commercial banks refused to extend new credit and international credit ratings were

downgraded. Inflation was at a high of 17% per annum, industrial production was falling, and overall economic growth had fallen to 1.1% (*New Economic Policies*, Lok Sabha Secretariat, 1998).

To pull the Indian economy out of this crisis, a Stand-By Arrangement was worked out by the new government with the IMF, subject to India undertaking wide ranging reforms. The first part of the reforms required stabilizing the economy by reducing the fiscal and current account deficit and raising GDP growth. The second part of the reforms was structural adjustments that included industrial de-licensing, liberalizing trade policy, and de-regulation of the financial sector. A New Industrial Policy was announced in July 1991, which among other things abolished licensing for all but 18 industries; industries restricted for public sector investments were cut from 17 to 8; and allowed small scale enterprises to offer up to 24% of shareholding to large enterprises.

In the external sector, there was an immediate devaluation of the rupee by 22% and the introduction of a dual exchange rate system in July 1991. Import liberalization was undertaken such that except for consumer goods, almost all items of capital goods, raw materials, intermediates, etc. became freely importable. Different import lists were consolidated into a single negative list of imports requiring a license. Beginning in 1991-92, the government also began the phase of tariff reductions, bringing down the maximum from 300% in 1990 to 50% by March 1995. The dispersion of tariffs was also significantly reduced and quantitative restrictions

were eased. Export restrictions were eased and additional export promotion plans, such as the Duty Drawback Scheme and Advance Licenses Scheme, were initiated. [1],[2]

Reforms were also undertaken with respect to the financial sector and monetary policy. Financial sector reforms included decontrolling interest rates on loans (above a certain amount), freedom to banks to open new branches, capital market reforms, etc. aiming towards eliminating administered interest rates and removing government interference in financial institutions.

Almost 70% of the population is in rural areas, most of who work in the agricultural sector. The growth of this sector is therefore important not only for food-security reasons in a country of one billion people, but also for the overall growth prospects of the country. The agricultural sector reforms removed most restrictions on movement of agricultural goods both domestically and for exports. Restrictions on agricultural exports were also substantially reduced. The lowered protection to the domestic industrial sector should reduce the anti-agriculture bias of the earlier protectionist policies. The competitiveness of agricultural exports should improve at the new exchange rate, and stimulate the growth of agro-processing industries. The development strategy during the ninth five year plan (1997-2002) for agriculture targeted sustainability of employment generation, food security and poverty alleviation. Regional plans were targeted to reach the full potential in each area and

[1] Drawback means the rebate of duty chargeable on any imported materials or excisable materials used in manufacture or processing of goods that are manufactured in India and exported. Duty Drawback is equal to (a) customs duty paid on imported inputs including SAD (Special Additional Duty) plus (b) excise duty paid on indigenous inputs.

[2] Inputs required to produce exported products can be imported without payment of customs duty under Advance License. Advance License can be granted to merchant exporter or manufacturer exporter to import raw materials. Since the raw materials can be imported before exports of final products, the licenses issued for this purpose are called ‘advance licenses’.

infrastructure development was considered highly important. While the agricultural sector is a larger player than the manufacturing sector in their contributions to GDP, insufficient data on the magnitude of agricultural reforms precludes an analysis of the direct impact of the reforms in agriculture. We will therefore focus on the manufacturing sector and the impact of the trade reforms on the wage gaps among different groups.

1.2 The Social Dynamics

1.2.1 The Caste System

The word *caste* is derived from the Portuguese word *casta* meaning lineage, breed or race.[3] In the Indian context, traditional Hindu society divided people into social classes based on the caste system. It divides people into hierarchical and hereditary groups which determined the social and economic status of individuals. There are four main caste (*varnas)* groups: *Brahmins* (priests, teachers)*, Kshatriyas* (warriors, rulers)*, Vaishya or Bania* (businessman) and *Shudras* (laborer, artisan). Individuals further outside this order were referred to as the *untouchables.* Traditionally the three upper castes have dominated over the other two groups. The upper castes have traditionally been richer and more influential, while the lowest caste (*untouchables*) was relegated to menial jobs. Higher castes practiced the system of *untouchability* wherein they believed that contact with *untouchables* would defile them. The higher castes didn't allow the lower castes to share the same resources (e.g., school, religious functions, etc.). Lower castes were therefore both economically and socially disadvantaged.

[3] The caste system exists in India, Nepal, Sri Lanka, Indonesia, and Japan among other countries. Anthropologists use the term *caste* to refer to a social group that is endogamous and occupationally specialized. (Wikipedia.org)

Table 1.1 presents work force participation rates for social groups in urban India. Between 1983 and 1999, comparing columns (1) and (2), work force participation rates have gone up slightly for male scheduled castes (SCs) but have decreased slightly for female SCs. The work force participation rates are higher among lower caste women compared to higher caste women, as seen by comparing columns (2) and (4).[4] This can be partly explained by the fact that higher caste women may face greater social stigma in working outside the household, unlike lower caste females. Work force participation rates of male scheduled castes are slightly lower than of other male workers.

1.2.2 Women in India

Women in India have traditionally been restricted to the household, but formal labor force participation rates of women have improved (since independence in 1947) just as observed in other countries across the world. Trade theory predicts that relative wages of unskilled labor will rise in the long run in an unskilled-labor abundant country that opens up to trade. It was expected that labor-intensive industries in labor abundant economies would become export competing and women who tended to dominate labor-intensive industries would stand to gain. It is expected that opening up to trade benefit women in a low-skill abundant country like India.[5] Joekes (1999) in her review of the literature on gender effects of trade reforms found a positive

[4] The term *lower caste* is used interchangeably here to refer to *Scheduled Castes*, explained further in section 1.2.3. It must however be noted that the scheduled castes are not the only 'lower' castes in India.

[5] Nordås (2003) found evidence of increased likelihood of trade liberalization raising employment and relative wages of women based on a case study of Mauritius, Peru, Mexico, Philippines and Sri Lanka.

relation between trade growth and women's employment and relative wages, with highest impact in the lowest income countries.

The work force participation rates for women are lower than for men throughout the periods being studied, but the work force participation rates are higher among lower caste women compared to higher caste women. The numbers are presented in Table 1.2. While it's expected that opening up to trade benefits women and lower castes in a low-skill abundant country like India, the work force participation numbers don't reflect it. This may be partly due to the fact that a large proportion of women work in the unorganized sector, which means that they may not be getting counted into the official statistics on employment. Thus while women's employment may actually be increasing during this time, official numbers may not reflect it.

1.2.3 Anti-discrimination Policies

The Constitution was written by members of the Constituent Assembly which was convened in December 1946. The Constitution was fully adopted on January 26, 1950. The preamble to the Constitution declared the country a sovereign democratic republic which promised justice, liberty and equality for all.

As part of ensuring equal opportunity for all groups, affirmative action (article 335 of the Constitution) for *untouchables* was introduced in the form of reservation of a certain percentage of jobs in the public (government) sector jobs for these groups. The lowest castes which were thought to be candidates for such reservation were listed in a schedule of the Constitution and hence they are referred to as the *Scheduled Castes*. They constituted about 12-15% of the population. Hence 12.5% of all

government jobs were reserved for scheduled castes. This was modified to 15% in 1970. Free schooling, subsidized higher education, as well as free training to prepare for the entrance examinations for the government services were made available to scheduled castes. Job reservations which were originally intended for a decade continue till date.

Affirmative action was also extended to tribal communities that had largely functioned outside the purview of modern communities. Tribes that were considered economically and socially disadvantaged and in need of programs to help their assimilation into society were identified and listed in another schedule of the Constitution. 7.5% of all government jobs were reserved for members of these *Scheduled Tribes*. In addition, article 46 of the Constitution noted that the State shall promote the educational and economic interests of the weaker sections of the people and in particular that of the Scheduled Castes and Scheduled Tribes. Articles 338 and 338A specified the creation of National Commissions for Scheduled Castes and Scheduled Tribes to evaluate and monitor the working of safeguards for these groups.

Equality of sexes is guaranteed by Article 15 which declares that the government shall not discriminate against any citizen on the grounds of sex. Article 15(3) allows the government to undertake affirmative action programs in favor of women and article 42 directs the State to ensure humane conditions of work. Article 39 directs the State to ensure equal pay for equal work for both men and women.

While there is legislation in place to ensure equal treatment of all citizens, there is a substantial difference in the outcomes of different groups. While article 45 of the Constitution directs the State to provide free and compulsory education to all

children till they are 14 years of age, there are substantial differences in the educational achievements of women compared to men and of lower castes compared to higher castes. The overall literacy rate for women is 39% compared to 64% for men. While differences in such productive characteristics are one potential explanation of wage gaps between men and women workers, other factors, such as discrimination in access to education, jobs, etc. are also plausible explanations. In the next chapter, we examine the extent of the wage gap between these groups and determine the part of the gap that is unexplained by these productive characteristics.

Chapter 2 Discrimination in Urban Indian Labor Market: Decomposing the Wage Gap for Gender and Caste Groups

2.1 Introduction

Wage inequalities among race, gender and social groups have been extensively studied in both developed and developing countries. The explanations provided for such wage differences are several, including the human capital theory, compensating differentials, search models, and discrimination. Human capital is the embodiment of productivity in people. The human capital theory predicts that earnings are higher for those with higher education and experience. Even after controlling for all observable factors such as education, age, experience, marital status, occupation and industry in earnings regressions developed by the human capital theory, wage differentials between workers may not be fully explained. The segment of the wage gap that is not explained by observable differences in worker characteristics is typically attributed to discrimination in the Blinder-Oaxaca (1973) decomposition technique.

Discrimination in the labor market has been studied extensively, particularly in developed countries. The main economic theory to study discrimination was developed by Gary Becker, in the 1950s. Becker developed a neoclassical model, with typical neoclassical assumptions of perfect competition in labor markets and utility maximization, using the concept of a *taste for discrimination* on the part of employers, employees or customers to examine the consequences of discrimination, where it exists. Becker proposes that this *taste for discrimination* creates a wage differential in the short run, as a result of the willingness on the part of discriminators

to bear an additional cost in order to avoid any association with certain other agents. But employer discrimination is predicted to disappear in the long run, subject to constant returns to scale production technology and the distribution of tastes. Becker highlights that in an open market economy there will be competitive forces working against discrimination. He views economic discrimination as a *reflection of market imperfections*. This theory was mainly applied to study racial discrimination but was extended later to analyze other forms of discrimination including gender discrimination.

Oaxaca (1973) and Blinder (1973) both independently developed a statistical model to estimate discrimination. They decomposed the earnings differential into a component explained by differences in personal characteristics of workers that affected their productivity and a component unexplained by observable productive differences and therefore attributable to discrimination. This has been a widely used technique to assess the extent of discrimination and has been applied to study both racial and gender discrimination.[6]

Studies on developing country discrimination have mostly focused on a select country. We follow along this line and focus on studying discrimination in the Indian labor market on the basis of gender and caste. We focus on these groups because they have traditionally been economically disadvantaged. The caste system is a hierarchical and hereditary system and the lowest castes have typically been economically and socially disadvantaged. Women have traditionally been restricted to working within the household and have much lower formal labor force participation

[6] We discuss the limitations of this strategy in Section 2.4.

rates than men. Females and lower castes in India have lower levels of education and experience than males and higher castes. This difference in human capital can be one explanation of wage gaps between gender and caste groups. However, studies (Sambamoorthi, 1984; Banerjee and Knight, 1985) have found the existence of wage and job discrimination against female workers and lower caste workers in India using decomposition analyses developed by Blinder (1973) and Oaxaca (1973). Banerjee and Knight (1985) found that there was wage discrimination against lower caste workers in a study of migrant workers in Delhi. This is particularly relevant since affirmative action programs which were begun in 1950 have reserved a certain percentage of jobs in the public sector to lower caste workers.

We contribute to the literature by conducting a systematic analysis of the wage gaps between male and female workers and between low and high caste workers. While some analysis of wage gaps has been conducted for specific regions or firms in India, there has been no country-wide analysis. We use a nationally representative dataset to analyze the changes, if any, in the extent of the wage gap for caste groups and gender groups over a period of almost two decades. We find that the wage gap between male and female workers has decreased and between low and high caste workers has increased slightly over the period under study. We also find that the explained component of the wage gap has increased over time. However based on the occupational decomposition of the wage gap, we find that discrimination in wages within occupations is more important than discrimination in access to occupations.

The chapter is organized as follows: Section 2 briefly describes the literature on labor market discrimination; Section 3 presents the methodology; Section 4

describes the data; Section 5 presents and analyzes the results on the existence of discrimination, and Section 6 concludes.

2.2 Literature on Labor Market Discrimination

Labor market discrimination consists of treating equally productive people differently in the labor market based on characteristics unrelated to an individual's productivity. Most studies on discrimination in developed countries have focused on explaining wage gaps between men and women and between races. The primary approaches in the literature include the neoclassical approach [primary among them Becker (1957), Arrow (1972), Blinder (1973) and Oaxaca (1973)], statistical discrimination approach [Aigner and Cain (1977)] and segmented labor market approach or the dual labor market hypothesis [(Doeringer (1986); Reich et al. (1973)].

2.2.1 Theoretical Literature

Becker models discrimination as arising out of the utility maximizing behavior of employers who suffer a disutility from employing certain workers and are therefore willing to give up some profits in the bargain. In an otherwise perfectly competitive world, the *distaste* of some discriminating firms from employing certain workers generates the *discrimination coefficient*, a monetary premium that such firms are willing to pay to not associate with those workers. Let us suppose firms can choose between hiring white or black workers, both groups being equally productive. However some employers dislike hiring black workers such that if the market wage rate for white workers is w, then the effective wage rate for black workers becomes w

$+ d$, where d is the *discrimination coefficient*. This results in the discriminating firms hiring fewer black workers than a non-discriminating firm.

We illustrate this outcome using the analogous case of male and female workers in Figure 1 which illustrates the discrimination of employers against female workers. The horizontal axis represents the number of female workers employed and the vertical axis represents wages and productivity. Both male and female workers are assumed to be equally productive and therefore have the same marginal revenue product of labor represented by the *MRP* curve. The equilibrium is determined by the profit maximizing condition in the labor market which requires that the wages paid to workers should equal their marginal productivity. Assuming a perfectly competitive labor market, the firm cannot alter the market wage w_F. The non-discriminating firm applies the profit maximizing condition and chooses to operate at point A where the market wage equals *MRP*. Discriminating firms dislike hiring female workers, and this makes them act *as if* women are less productive than men. The more biased (larger d) an employer, the more the actual productivity of female workers is undervalued by discriminating firms. This results in discriminating firms equalizing the market wage to a modified (subjective) estimate of female worker productivity which is less than actual female productivity. Therefore discriminating firms maximize profits at B where the market wage equals $MRP - d$, and therefore employ fewer workers than a non-discriminating firm. Since male and female workers are equally productive this equilibrium implies that the *discrimination coefficient*, d, is the extent by which male wages exceed female wages. Notice that the profit of firms equals the difference between total revenue and total cost. Figure 1 shows that in

equilibrium discriminating firms make lower profits than non-discriminating firms by the amount of the triangular area ABC.

Since discriminating firms are not utilizing factors of production efficiently, they suffer losses in the short run which cannot be sustained in the long run. Therefore discriminating firms, in the Becker model, will have to either shut down or sell off to non-discriminators in the long run, subject to certain conditions pertaining to the production function and distribution of tastes.

Oaxaca (1973) and Blinder (1973) have contributed to this literature by devising the technique to decompose earnings differences between groups. If we suspect wage differences to be due to discrimination, this technique can identify the possible extent of such discrimination. The technique depends on identifying characteristics of individuals that affect their productivity and therefore their wages, such that the share of the wage gap not accounted for by such productive factors is the unexplained wage gap, all or some of which may be due to discrimination. Though the technique suffers from the index number problem, it is definitely a valuable first step towards analyzing discrimination which was further developed by others (e.g. Brown et al. 1980).[7]

Several economists including Arrow (1972) and Phelps (1972) argue that traditional neoclassical theory fails to explain the existence of discrimination. They propose the statistical model of discrimination based on employer uncertainty about a worker's productivity and therefore reliance on group characteristics to judge an individual. This can be explained by the following example.

[7] The index number problem is discussed in the methodology section 2.4.1.

Assume there are two races in the economy: W and B, and that the true productivity of each individual cannot be observed, but we know the average productivity of the races. Thus in the event that two people from these two groups seek employment at the same firm, the firm lacking any other information on the individual, will fall back on the average performance of individuals from that group to conclude what the individual's productivity is likely to be, and choose the one who belongs to the higher average productivity group. In doing so the firm may be hiring someone who is below average from the high productivity group and meanwhile reject someone who is above average from the low productivity group. They are thus wrongly attributing *low* group characteristics to an individual who might be above average in the lower average productivity group, and therefore there is an additional problem of a negative feedback effect wherein such discriminated-against individuals will not have enough incentive to invest in human capital. This form of discrimination is however not based on any distaste on the part of the employer and as such is expected to be corrected by *learning* over time as employers observe the actual productivity of workers.

Another view is that internal labor markets can explain the existence of labor market discrimination. Doeringer (1986) analyzed the segmentation of the labor market by firms and trade unions via the operation of an internal labor market. The uncertainties of the economy may make it efficient to offer workers an implicit wage contract which ensures stability for risk-averse workers. Efficiency wages (offering workers a wage above the prevailing market wage in order to induce them to not shirk and face the prospect of losing a 'valuable' job) might explain why wages may not

fall within internal labor markets. The bargaining nature and importance of social work groups within internal markets allows the existence of non-competitive forces. Higher productivity under the non-competitive conditions can generate higher rents that can be shared with *insiders*.

2.2.2 Empirical Literature

Several authors have analyzed the existence of discrimination in developing countries in recent years.[8] The primary technique employed in these studies is the Blinder-Oaxaca decomposition technique which classifies the part of the wage gap unexplained by productive characteristics as the discrimination component. Several studies have further used the Brown et al. (1980) method to distinguish between the components of discrimination (whether it is due to unequal access to occupations or due to unequal pay within occupations).

Not much empirical work has been done on caste-based and/or gender discrimination in India. One of the earliest empirical analyses of caste discrimination in India was carried out by Banerjee and Knight (1985) who decomposed the caste wage-gap using the occupational attainment framework by Brown et al. (1980). They restricted their analysis to male workers in the capital city of Delhi, and designed their own survey. Using 1974-75 data they found evidence of the existence of discrimination against lower caste workers in Delhi, and using the occupational attainment model they found that discrimination accounted for about two-thirds of the

[8] See Scoville (1991) for a collection of studies on status influences in third world countries. See Birdsall and Sabot (1991) for an additional collection of analysis of discrimination in urban labor markets in developing countries.

gross earnings differences, and that wage discrimination within an occupation was much more important than discrimination in access to jobs.

Bhattacherjee (1985) estimated the extent of wage discrimination against low caste workers in a Bombay automobile firm. Based on the Oaxaca decomposition of the wage gap, he found evidence indicating that wage discrimination was reflected in the under-valuation of lower caste human capital and other characteristics, and this discrimination mostly stemmed from differential access to resources important to developing skills for the labor market.

More recently, Deshpande (2000) examined the role of caste affiliation as an indicator of inter-group disparity. She formulated the *Caste Development Index* (*CDI*), based on the *Theil* index, to measure caste inequality and defined caste disparity as the distance between the *CDI* for low caste groups and for others (non-low castes). She analyzed between-group and within-group disparity focusing on the southern Indian State of Kerala which has experienced greater social and labor reforms thanks to a communist regime and where caste is expected to be less important in the society. She used the *Consumption and Expenditure Data* from the *National Sample Survey Organization* for 1993-94, and considered three groups - the scheduled castes (SC), scheduled tribes (ST) and others.[9] She found that overall inequality was not on the higher side, as was expected given the State's social and political history. Contrary to expectations, she found evidence of inter-caste disparity in both rural and urban areas. But she found that the problem of within-group

[9] Scheduled castes and tribes are castes and tribes listed in the Constitution which are deemed to be economically disadvantaged because of past exploitation and lack of opportunities. There are specific policies formulated for their development which include free schooling, reservations in government jobs, etc.

inequality was more a concern for 'others' than for the scheduled castes and scheduled tribes.

We contribute to the empirical literature on labor market discrimination by extending the analysis to a country with a well known social hierarchy. The caste system has been the foundation of traditional Hindu society and while urbanization might provide greater anonymity to individuals, caste affiliations still affect their interactions. We also examine the changes with respect to women who are increasing their labor force participation compared to earlier years. We provide a thorough investigation of the changes in the wage gap over almost two decades for these different groups using the most comprehensive dataset available for the country. While past studies on India have focused on particular cities, we are able to conduct the analysis for almost the entire country.

2.3 Data

The data used in this study comes from the *Employment and Unemployment Survey* (*Schedule 10*) of the *NSSO* (National Sample Survey Organization). This quinquennial survey is divided into four sub-rounds and covers both urban and rural areas. It is a time-series of cross-sections with different households being interviewed each year. It contains detailed information on demographic and household characteristics of individuals (like age, education, household size, industry, occupation, etc.).[10] The survey adopts a stratified two-stage design.[11] This survey is available for four years, namely 1983, 1987-88, 1993-94 and 1999-2000.

[10] The sub-rounds are from July-September, October to December, January to March, and April to June. Equal number of sample villages and blocks are allotted for survey in each of these four sub-rounds.

We focus on urban areas where there is more reliable wage reporting. India is a union of 29 States and 6 Union Territories. The analysis is restricted to 16 States which are typically the focus of studies on India. The omitted States are primarily those in north-east India, where frequent insurgency problems have affected data collection. The list of States is available in the Data Appendix.

The education variables are defined based on completed years of schooling. We define six categories of education: not literate; literate but below primary; completed primary school; completed middle school; completed secondary school and graduated from college. The survey reports total weekly earnings of each individual. These are converted to real weekly earnings which are the primary outcome variable of interest. We focus on those aged 15-65 years and include marital status and household size information in the regressions to control for selection bias. We use a fourth order polynomial in age to account for gains in wages due to experience.

Occupation of workers is listed based on the *National Classification of Occupations (NCO)*, which is classified into ten one-digit level occupations; and on the basis of skill into six occupational groups: professional, production, clerical, service, skilled and unskilled workers (based on Banerjee and Knight, 1985). Professional workers include all professional workers and managerial, executive and administrative workers. Unskilled workers are from all other occupation groups,

---

[11]The first-stage units are census villages in the rural sector and the NSSO urban frame survey (UFS) blocks in the urban sector. In 1993-94, the survey covered more than 69000 rural and 46000 urban households. The total sample size is determined by the relative population sizes of rural and urban areas with double weight to the urban sector. More than 97000 rural and 67000 urban households were surveyed in 1999-2000.

including the 'laborers not classified elsewhere'. Government workers get included here in the professional categories. We use ten one-digit industry codes, based on the *National Industrial Classification (NIC)* to identify the industry affiliation of each worker.

2.4 Methodology

We apply the Blinder-Oaxaca (1973) and Brown et al. (1980) methodology to estimate the extent of discrimination and the changes in the components of the wage gap over time.

2.4.1 Blinder-Oaxaca Decomposition Technique

The Oaxaca (1973) and Blinder (1973) decomposition technique involves identifying all the productive factors that affect an individual's wages and examining whether different individuals with the same characteristics are rewarded differently. We first run earnings regressions for each group of workers, namely non-scheduled and scheduled castes, and male and female workers. For sake of brevity, we present the equations for caste workers. The equations for male and female workers are similar. The equations for each individual are:

$$\begin{aligned} w_i^n &= X_i^{n'}\beta^n + u_i^n \\ w_i^s &= X_i^{s'}\beta^s + u_i^s \end{aligned} \quad . \qquad (1.1)$$

The superscript *n* denotes non-scheduled castes and *s* denotes scheduled castes. w_i^n is the natural log of earnings for individual i, $X^{n'} = (1, X_{1i}^n, ..., X_{Ki}^n)'$ is the vector of a constant and *K* observable characteristics for individual *i*, $\beta^n = (\beta_0^n, \beta_1^n, ...\beta_K^n)$ is the corresponding vector of coefficients, and u_i^n is the random

error term, and similarly for scheduled castes. The gross difference in average earnings can then be written as $G = \overline{X}^n b^n - \overline{X}^s b^s$, where *b* is the vector of estimated coefficients. We can add and subtract the term $\overline{X}^s b^n$ and rearrange G as follows:

$$G = \overline{w}^n - \overline{w}^s = (\overline{X}^n - \overline{X}^s)b^n + \overline{X}^s(b^n - b^s) + (\overline{u}^n - \overline{u}^s), \qquad (1.2)$$

where the first term on the right hand side represents the part of the earnings differential due to difference in characteristics (and thus the explained part), and the second term represents differences due to differential returns to the same characteristics (captured by different coefficients for non-scheduled and scheduled castes), and the third term is the difference in the averages of the error terms. The second and third term are normally attributed to discrimination. Alternatively we could subtract the term $\overline{X}^n b^s$ and rearrange G to obtain:

$$G = \overline{w}^n - \overline{w}^s = (\overline{X}^n - \overline{X}^s)b^s + \overline{X}^n(b^n - b^s) + (\overline{u}^n - \overline{u}^s). \qquad (1.3)$$

In equation (1.2) it is assumed that everyone is getting treated as if they were non-scheduled castes (by getting the same rate of return b^n), and therefore the discrimination was the difference in the *b* coefficients. In equation (1.3) it is assumed that everyone is treated as if they were scheduled castes. Either method can be used to decompose G, though clearly the explained and unexplained components will vary numerically across the two methods. This creates an index number problem since the components of the decomposition can vary based on which group is used as the reference group.

Researchers (Neumark 1988; Oaxaca and Ransom 1988; Oaxaca and Ransom 1994) have attempted to create other reference wage structures for the decomposition to solve this index number problem. Neumark (1988) suggests that the correct

reference group to be used for the decomposition technique can be determined if enough is known about the market which would enable one to decipher if, say females, are being underpaid (discrimination) or if males are being overpaid (nepotism). If females were underpaid, then the coefficients from the male earnings regressions should serve as the reference. However if employers are nepotistic, then they overpay men and therefore coefficients from the female earnings regressions should serve as the no-discrimination reference for the decomposition analysis. Neumark (1988) suggests using a weighted average of the male and female wage structures to estimate the non-discriminatory reference wage structure if enough is not known about the labor market and employer preferences. This involves assuming that employers only care about the proportion of each type of labor employed (that is employer preferences are homogeneous of degree zero). The Neumark decomposition of the wage gap between males (m) and females (f) can be represented by the following equation:

$$\overline{w}_m - \overline{w}_f = \beta(\overline{x}'_m - \overline{x}'_f) + \overline{x}'_m(\beta_m - \beta) + \overline{x}'_f(\beta - \beta_f). \quad (1.4)$$

The first term represents differences in characteristics between males and females and is therefore the explained part of the wage gap. The second and third terms represent the share of the wage gap due to differences between the actual wage structure and the pooled wage structure (β) for male and female workers respectively. While this technique can avoid the index number problem, it might still be subject to problems, including omitted variable bias.[12]

[12] The pooled coefficient may not be a good estimator of the non-discriminatory wage structure if there isn't zero-homogeneity restriction on employer preferences (Appleton et al. 1999).

2.4.2 Limitations

The primary drawback of any decomposition technique, including the one described above is that it attributes everything that is not explained by observable characteristics, such as education, experience, occupation, etc., to discrimination. As pointed out in the literature, the problem with this technique arises if there are factors, such as an individual's ability, which are not included in the list of productive characteristics (*X)* but affect wages earned. One solution would be to use a proxy to measure the relevant factor, for example using a test score on an IQ test as a measure of ability. Another option would be to use an instrument variable for the missing factor. By definition, the instrument should be highly correlated with the factor it is representing but should be unrelated to the outcome variable. Examples of instruments for ability would include parents' education or family background that should explain each individual's ability but at the same time doesn't dictate the wages that they would actually earn in the labor market.

The data being used in this study does not contain enough information on any proxy or instrument variable measure of such unobservable factors which could affect wages. Thus the results have to be interpreted keeping in mind that the entire amount of the wage gap unexplained by the productive factors is not necessarily discrimination. It is important to note that as long as we don't expect any systematic changes in ability over time, the analysis of changes in the wage gap over time should not be systematically affected by this drawback.

Another problem arises if people select in and out of the labor force based on some characteristic that is related to some observable productive characteristic, such

as education. This correlation between education and regression residuals would result in biased coefficients if we use ordinary least squares. We include household size and marital status to capture the effects of such selection bias.[13]

2.4.3 Occupational Attainment Model

The usual characteristics that get included in the list of explanatory variables are age, education, experience, etc. Occupational choices also explain wage differences, and are usually incorporated in this list. However if there was pre-labor market discrimination which in turn affected occupational choices, the traditional strategy would not give correct estimates of discrimination. It would generate lower estimates of discrimination since it would not have accounted for differences in occupational attainment. Brown et al. (1980) develop the strategy to further decompose the wage differential using a weighted average of proportions of workers of each group in every occupation to account for discrimination in the occupational choice itself.

The sample proportions of each group are included to generate the wage differential $\overline{w}^n - \overline{w}^s = p^n \overline{w}_j^n - p^s \overline{w}_j^s$, where $p^n = (p_1^n, p_2^n, ..., p_J^n)$ are the proportions of non-scheduled caste workers in each occupational group j=1,…,J, and similarly p^s is the proportion for scheduled castes, and $\overline{w}_j^n$ are the average wages of non-sc workers in each occupation, and similarly for SC workers. Add and subtract $p^s \overline{w}_j^n$ and re-arrange:

[13] In a future extension of this study we propose to incorporate the Heckman (1979) correction for self-selection bias.

$$\overline{w}^n - \overline{w}^s = p^s(\overline{w}_j^n - \overline{w}_j^s) + \overline{w}_j^n(p^n - p^s) = W+J. \qquad (1.5)$$

The first term on the right hand side is the gross difference in wages due to difference in wages across groups within occupations, while the second term is the part due to differences in occupational composition. Each of these terms can be further decomposed into explained and unexplained parts. Add and subtract $\overline{X}_j^s b_j^n$ to the first term (W) and re-arrange:

$$\begin{aligned} p^s(\overline{w}_j^n - \overline{w}_j^s) &= p^s(\overline{X}_j^n b_j^n - \overline{X}_j^s b_j^s) \\ &= p^s(\overline{X}_j^n - \overline{X}_j^s)b_j^n + p^s\overline{X}_j^s(b_j^n - b_j^s). \end{aligned} \qquad (1.6)$$

The first term in equation (1.6) is the explained part of the difference in wages, while the second term is the part of the wage difference due to differential returns between castes and thus is attributed to discrimination.

Next, let $\hat{p}_j^n$ be the proportion of non-scheduled castes that would be in occupation j if non-scheduled castes faced the same occupational choices as scheduled castes, and similarly $\hat{p}_j^s$ be the proportion of scheduled caste workers who would be in occupation j if scheduled castes face the same occupational structure as the non-scheduled castes. The term J can be re-arranged by adding and subtracting $\hat{p}_j^s$:

$$(p^n - p^s)\overline{w}_j^n = (p^n - \hat{p}^s)\overline{w}_j^n + (\hat{p}^s - p^s)\overline{w}_j^n. \qquad (1.7)$$

The first term in the above equation reflects differences in qualifications for the occupations, and the second term is due to differences in the structure of occupational attainment between non-scheduled castes and scheduled castes. Thus the full model becomes:

$$\overline{w}^n - \overline{w}^s = p^s(\overline{X}_j^n - \overline{X}_j^s)b_j^n + p^s\overline{X}_j^s(b_j^n - b_j^s) + (p^n - \hat{p}^s)\overline{w}_j^n + (\hat{p}^s - p^s)\overline{w}_j^n \ . \quad (1.8)$$

In the above equation, the second and fourth terms represent the parts due to discrimination in wages and occupations respectively. This method therefore gives a better picture of the true extent of discrimination. This exercise could also be done by adding and subtracting $p^n\overline{w}_j^s$ in the gross wage differential equation instead to get equation (1.5) and adding and subtracting $\hat{p}_j^n$ in the *J* term to get equation (1.8). The probability that individual *i* is in occupation *j* as a function of exogenous variables is calculated using a multinomial logistic model:

$$p_{ij} = prob(y_i = oc_j) = e^{X_i'\gamma_j} / \sum_{k=1}^{j} e^{X_i'\gamma_k} \ . \quad (1.9)$$

Here *i= 1, ..., n*; and *j= 1, ..., J*, oc_j is occupation *j*, and *X* is the vector of independent variables. Similar equations are applied for male and female workers.

2.5 Summary Statistics and Earnings Functions Analysis

We begin by looking at the summary statistics (Table 2.1 and Table 2.2) for scheduled caste and non-scheduled caste workers, and male and female workers. The sample is restricted to those aged between 15 and 65 in urban areas in 16 States. The average real weekly earnings are much higher for higher caste workers than for scheduled caste workers, and for males compared to female workers. The ratio of average wages for lower castes relative to higher caste workers improves from about 0.72 in 1983 to 0.77 in 1999-2000. The relative average wage ratio for female to male workers improved from 0.56 in 1983 to 0.74 in 1999-2000. Higher castes and male workers have smaller proportion of illiterates than lower castes and female workers, but there is an improvement in educational achievements for each group over time.

The highest proportion of workers from each group is those who have completed secondary education. The proportion of lower caste and female workers completing college doubles between 1983 and 1999-00.

The summary statistics also list the proportion of workers in some of the industries and occupations. There are ten one-digit industries and we report the proportions employed in three sectors, namely agriculture; manufacturing; and community, social and personal services. The proportion of low and high caste workers in manufacturing is roughly similar and decreases between 1983 and 1999. A larger proportion of lower caste workers are engaged in agriculture (14% in 1983) compared to higher caste (11.6%) workers. The services sector is also a large employer of workers in all years except 1999-00. The proportion of men in agriculture is much lower than that of women, while both show a decline over time. Trends in manufacturing and services sectors are similar to that of caste workers.

Of the ten one-digit occupations, we report the proportions employed in professional, services and production jobs. There is a larger proportion of non-scheduled caste (6.4% in 1983) than scheduled caste (1.5%) workers in professional jobs. On the other hand there is a larger proportion of female (10.5% in1983) than male (4.7%) workers in professional occupations. The proportion employed in professional occupations has increased slightly over time for all groups of workers. There is a larger proportion of scheduled caste (25% in 1983) and male (16%)

workers in production jobs than higher caste (14%) and female (11%) workers respectively, and these proportions have been declining over time.[14]

2.5.1 Earnings Functions Analysis of the Entire Sample

We begin the analyses by examining the wage regressions for each year for the entire sample of workers and including dummy variables to identify wage differentials by caste and gender. The equation being estimated looks as follows:

$$\ln w_{ikhj} = \alpha + \beta X_{ikhj} + \theta_1(male_i) + \theta_2(caste_i) + \sigma(X_{ikhj} * male_i) + \omega(X_{ikhj} * caste_i) + \lambda_k + \gamma_h + \pi_j + \varepsilon_{ikhj} \quad (1.10)$$

The natural log of real weekly earnings (*lnw*) of each individual *i* is regressed on a set of demographic (X) and institutional (state λ, industry γ, occupation π) factors. The *caste* variable equals one if individual belongs to the scheduled castes, and zero otherwise. The interaction terms of X with the *male* and *caste* dummy variables is to allow for the educational and age characteristics to vary by gender and caste.

State fixed effects are included to account for differences across states in prices; any state-specific differences in the labor market; or any other relevant factors that might affect wages earned. As mentioned earlier, the sample is restricted to 16 States and the southern Indian State of Andhra Pradesh is the omitted category. The individuals are surveyed in one of four sub-rounds. Therefore a *sub-round* dummy variable is included to control for the season when the individual was surveyed. The omitted season is July to September.

[14] The numbers in parentheses are for the same year as the first reference in each case, and there are similar trends over time.

Industry fixed effects are included to account for differences in returns to individuals across industries. There are 10 one-digit industries and agriculture is the omitted category. Occupation fixed effects are included to account for wage differences between occupations. There are 10 one-digit occupations and the omitted group consists of scientists, engineers, etc. The last term is the random error term. As mentioned earlier omitted variable bias is a common constraint of decomposition analyses. While there are remedies in the form of a proxy or instrument variable being used for the omitted variable, the dataset used in this study doesn't have sufficient information to allow either of these solutions. We however contend that the analysis here is not necessarily hampered since we are interested in examining the changes over time. As long as the extent of the biases due to any omitted variables is not changing over time, we should still be able to derive meaningful implications from these decomposition results.[15]

We begin by analyzing the earnings regressions assuming that everyone has the same returns to productive characteristics and examine the magnitude of the caste and gender variables. With reference to equation (1.10), we begin with the case where σ=0 and ω=0. The results in Table 2.3, Panel A show that the coefficient of *male* is positive and statistically significant at the 1% level in all four years. The magnitude of the *male* coefficient in 1983 is 0.529 implying that on average male workers earn almost 53% higher wages than female workers, having controlled for demographic and institutional factors. The magnitude of the *male* coefficient reduces to 0.358 in 1999-00, but remains statistically significant. We therefore analyze this wage gap

[15] Jolliffe and Campos (2003) conduct the Blinder decomposition with a similar drawback of limited information in Hungarian data.

between male and female workers and determine what extent of this gap is due to observable characteristics. The coefficient of the *caste* (*caste* equals one if individual belongs to lower caste, zero otherwise) variable is negative and statistically significant in all years except in 1993-94. The magnitude of the *caste* coefficient is slightly lower in 1999-00 compared to 1983, at 0.031 implying that lower caste workers earn about 3% lower wages than higher caste workers.

We also present regression estimates with and without the inclusion of the occupation fixed effects in Table 2.4. When occupation fixed effects are included (Panel B), we observe that the magnitude of the caste coefficient reduces while the magnitude of the male coefficient increases. The explanatory power (in terms of the adjusted R squared values) of the regressions increases slightly with the inclusion of occupation fixed effects. The average returns to higher education are slightly lower after controlling for occupations.

Next we allow for different rates of return to education for different groups of workers. Thus we estimate the full model specified in equation (1.10). Table 2.3, Panel B, shows that the *caste* variable in 1983 becomes statistically insignificant once we allow for differential educational returns across groups. However for all the other years, the coefficient of the *caste* variable is larger (than in the previous specification with σ=0 and ω=0) in magnitude and statistically significant. This implies that even after accounting for lower caste and higher caste workers having different returns to education, lower caste workers earn 10% lower wages in 1993-94 which reduces to 3.6% lesser wages than higher caste workers in 1999-00. All the coefficients of *education x caste* variables were negative and statistically significant (not reported in

the table) implying that lower castes had lower returns to each level of education than higher caste workers.

Similar results hold for male and female workers as well. The magnitude of the *male* coefficient increases once we allow for differential returns to education across genders. The estimate of the wage gap due to gender alone rises from 35.8% to 49.8% in 1999-00 once we allow for different educational returns. Across the years, the coefficients of *education x male* are statistically significant (not reported in the table) for those with secondary and graduate education, but they are negative in magnitude implying that female workers get higher returns to these levels of education than male workers. However the returns to middle school and higher levels of education are falling over the seventeen year period.[16] Therefore after controlling for all the productive and institutional factors, there is a much lower wage gap between low and high caste workers compared to male and female workers.

2.5.2 Earnings Functions for Each Group of Workers

Next we analyze the earnings functions for each of the four groups of workers for each year. We estimate the equation separately for each group of workers and include a *caste (male)* dummy variable in the gender (caste) equations. The earnings function for each group (male, female, low and high castes) is estimated using the following regression:

$$\ln w_{ikhj} = \alpha + \beta X_{ikhj} + \lambda_k + \gamma_h + \pi_j + \varepsilon_{ikhj}. \qquad (1.11)$$

[16] This can be seen by reading across the row for middle, secondary and graduate school coefficients. Average returns to middle school decrease from 0.499 to 0.335, and for college graduates from 1.457 to 1.245 (Table 2.3, Panel B).

The dependent variable is the natural log of real weekly earnings for each individual *i* in state *k* in industry *h* and occupation *j*. *X* includes five education dummy variables for each level of education completed; interaction terms between education variables and age to account for changes in the schooling quality across cohorts; a fourth order polynomial in age; marital status and household size. The same categories are excluded as detailed above in section 2.5.1.

Table 2.5 and Table 2.6 present the results for gender and caste groups respectively. The earnings regressions in these tables include state, industry, occupation and sub-round fixed effects. They also allow for interaction terms between *education* and *age*. The regressions for male and female workers include a *caste* dummy variable. Notice in Table 2.5, the coefficient of *caste* is negative and statistically significant for male (Panel A) workers while it's positive and statistically significant for female (Panel B) workers. This implies that after controlling for productive and institutional factors, lower caste male workers earn lesser than higher caste male workers, while lower caste female workers earn more than higher caste female workers. The magnitude of the *caste* coefficient changes substantially for female workers, indicating a 6.2% wage differential between caste groups in 1983 which rises to 14.6% in 1993 and falls back to 5.5% in 1999-00. Notice that household size coefficients are negative and statistically significant for male workers in all years, but it's statistically significant for female workers only in the second and last years.

The regressions for low and high caste workers include a *male* dummy variable. The *male* coefficient, in Table 2.6, is positive and significant for both low

and higher caste workers. Household size variable is statistically insignificant for lower caste (Panel A) workers. The R-squared values for all (except for lower castes) the regressions indicate that the demographic and institutional factors account for more than 50% of the variation in earnings.

Table 2.7 and Table 2.8 present the estimates for each group separately when we also include interaction terms between *education* and *caste* for gender group regressions, and interaction terms between *education* and *male* for caste groups. In Table 2.7, Panel A, we observe that controlling for differences in education between castes in examining wages of male workers renders the *caste* coefficient smaller in magnitude and statistically insignificant in 1999. We also observe an increase in the returns to higher levels of schooling. In Table 2.7, Panel B, we observe that the *caste* coefficient is still positive and statistically significant for female wage earners. We also observe that returns to higher levels of education are falling over time. In Table 2.8, the *male* coefficient is larger in magnitude and statistically significant in all the years for both low and high caste workers.

2.6 Blinder-Oaxaca Decomposition Results

Having found that the *male* and *caste* coefficients are statistically significant, we present the results for the Blinder-Oaxaca decomposition of the wage gap between male and female workers and between low and high caste workers in Table 2.9 and Table 2.10 respectively. As mentioned earlier, several reference groups can serve as the case where there is no discrimination. We present results based on the higher-income groups (males and higher castes) being considered the no-discrimination case (thus assuming that females and lower castes get paid less, rather than assuming

nepotism favoring men and higher castes) and the pooled wage structure suggested by Neumark (1988).

2.6.1 Decomposition Results for Gender Groups

Table 2.9, Panel A, presents the wage gap between male and female workers and the explained and unexplained components of this gap, assuming males to be the reference no-discrimination group. We also present the standard errors of these components. The regressions used to determine the wage gap controls for state, industry and occupation fixed effects. We use the entire sample of Hindu workers and include a *caste* dummy variable. We also include household size and a dummy variable for being married to control for the selection bias. Only two components (education and occupation) of the explained part of the wage gap are presented for the sake of brevity. The log wage gap in the second column equals the sum of the total explained and unexplained parts in the fourth and seventh columns. The second column shows that the log wage gap has fallen substantially over time from 0.825 to 0.55 between 1983 and 1999-00. We also observe that the extent of the wage gap that was explained (unexplained) declined (increased) in 1987-88 and 1993-94, but increased by 1999-00.[17] Based on the male reference group estimates, about 55% of the wage gap in 1999 is unexplained by differences in productive characteristics and endowments between men and women. We can also observe that education is a significant part of the explained component. Note also that all the estimates are statistically significant at the 1% level.

[17] The explained component was 37%, 35%, 26.6% and 45.5% of the total wage gap in 1983, 1987-88, 1993-84 and 1999-00 respectively.

Table 2.9, Panel B, presents the wage gap using a weighted average of the male and female wage structures. The trends remain the same with an increased share of the wage gap explained in 1999 relative to 1983. The explained component is slightly larger and the contribution of the education component in the explained part increases slightly in all the years. Based on the pooled-weighted estimates, 50% of the wage gap is unexplained by differences in productive characteristics between men and women. All the estimates are statistically significant at the 1% level. We also decomposed the wage gap based on the female wage structure (not reported here) as the reference no-discrimination group (it assumes nepotism favoring males). It resulted in similar but slightly lower estimates of the explained components than the case assuming that the male wage structure was at the competitive levels.

2.6.2 Decomposition Results for Caste Groups

Table 2.10, Panel A, presents the wage gap between low and high caste workers and the explained and unexplained components of this gap, assuming higher caste workers to be the reference no-discrimination group. We use the entire sample of Hindu workers and include a *gender* dummy variable. We also include household size and a dummy variable for being married to control for the selection bias. We observe that the wage gap between castes is smaller than between male and female workers (as was also indicated by the smaller coefficients on the *caste* dummy compared to the *male* dummy in the earnings regressions). The wage gap increases in 1987-88 and decreases slowly by 1999-00, but ends up slightly higher than the wage gap in 1983. However, most of the wage gap (almost 95%) is explained by the

differences in productive characteristics between workers. The unexplained component is not statistically significant in 1987 and 1993.

Table 2.10, Panel B, presents the wage gap using weights from the pooled model. The trends and magnitudes of the estimates are similar to those based on the higher-caste reference group. The magnitude of the wage gap between low and high caste workers is larger than previous estimates for the city of Delhi by Banerjee and Knight (1985). The main difference however is that most of the wage gap, in these estimates, is explained by differences in characteristics between low and high caste workers, while Banerjee and Knight had found larger unexplained components. We also decomposed the wage gap based on the lower caste wage structure (not reported here) as the reference no-discrimination group (assumes nepotism favoring higher castes). We get similar but slightly lower estimates of the explained components than in the case assuming that the high-caste wage structure was at the competitive levels.

2.7 Occupational Attainment Decomposition

Having obtained results that show the existence of 'unexplained' differences in wages, we go on to a model of occupational attainment which tries to examine if there is any discrimination in terms of which occupation one might be forced to take up and hence have 'crowding into certain occupations' (Bergmann, 1971) for lower castes based on pre-labor market discrimination. We examine the occupational distribution of workers and whether there is a difference in the occupational attainment between male and female workers which can explain the continuing wage gap based on the technique developed by Brown et al. (1980).

There might be differences in characteristics that stem from differential access to resources *before* entering the labor market. Then the differences in characteristics would themselves be influenced by discrimination, in which case one would have understated the true extent of discrimination. The occupational attainment model is based on predicting the occupational distribution of one group assuming that it is determined in the same way as that of members of the other group (that is, lower caste workers choose occupations just like higher caste workers, and female workers choose occupations just like male workers), and involves estimation of within-occupation earnings functions.

Multinomial logistic regressions are used to predict the occupational distribution of one group assuming they chose occupations as if they were members of the other group. The multinomial logistic regression is first run for one group. The dependent variable is occupation which is re-defined into six categories based on Banerjee and Knight (1985).[18] The independent variables are age, education, marital status, household size; state, industry and sub-round dummy variables. Employing these multinomial logistic estimates, predicted distributions for the other group of workers of belonging to these six occupations are obtained**.** The predicted probabilities are used as the $\hat{p}_j^s$ values, which show the proportion of scheduled caste workers who would be in occupation j if scheduled castes (SCs) faced the same occupational structure as the non-scheduled castes (non-SCs). Predicted probabilities are similarly calculated for female workers relative to male workers. This is used to predict what the occupational pattern of lower caste (female) workers would have

[18] The six occupations are: skilled, professional, clerical, production, unskilled and service.

been in the absence of any pre-labor market discrimination against them. This exercise could also be done assuming that everyone is making occupational choices as if they were lower caste (female) workers.

It is assumed that $P_n = \hat{P}_n$, since this means that occupational choices of higher caste (male) workers are not subject to discrimination, and their probability of choosing an occupation is the same as is observed in reality and therefore in the sample; while the lower caste (female) workers might exhibit occupational choices different in reality (P_{sc}) than what they might have ($\hat{P}_{sc}$) ended up with if they were allowed to choose occupations freely just like the higher caste (male) workers. The *explained* part of occupational distributions is therefore $P_n - \hat{P}_{sc}$ since it incorporates actual occupational distributions of non-SCs and of SCs if they chose occupations as if they faced the same choices as the non-SCs. Thus $\hat{P}_{sc} - P_{sc}$ is the *unexplained* component since it reflects the difference in occupational distribution arising from actual choices (P_{sc}) of lower caste workers being different from what they should have been ($\hat{P}_{sc}$) in the absence of discrimination. Similar results are applicable for male and female workers.

Within-occupation earnings regressions are run for each of the six occupations with the same independent variables. For each occupational group, the actual earnings difference between lower caste (female) and higher caste (male) workers is broken down into explained and unexplained parts assuming that everyone is paid like higher caste (male) workers in the labor market. Table 2.11 contains notes to interpreting the occupation-based decomposition results presented in Table 2.12 to Table 2.19. The

decomposition includes explained differences in wages and jobs and the residual components which are attributed to discrimination. Standard errors are also reported. The format of these tables is based on Banerjee and Knight (1985).

2.7.1 Occupation-based Decomposition for Gender Groups

The decompositions are presented for each year in Table 2.12 to Table 2.15. We observe a larger proportion of male workers engaged in skilled and clerical jobs compared to females. While the proportion of male workers in skilled occupations increases over the seventeen year period, the proportion in clerical occupations decreases over the same period. In each of these tables, Column (3) lists the predicted occupational distribution of female workers assuming that they choose occupations in a manner similar to male workers. Column (5) shows the observed difference in the occupational attainments between male and female workers with higher proportions of females in professional, service and production jobs and a higher proportion of males in skilled, clerical and unskilled jobs. We can divide this observed difference in occupational distributions in column (5) into two parts as presented in the next two columns. Column (6) represents the part of this difference in occupational distribution that is explained, that is, the difference between the predicted occupational distributions of women assuming they choose occupations based on the same considerations as men, and the actual occupational distribution of men. Column (7) on the other hand represents the part of the difference in occupational distribution between the predicted and actual occupational distribution of females. A positive number in column (7) implies that the actual proportion of women in the corresponding occupation is greater than if they had chosen subject to a male

preference structure. A negative number in column (7) however indicates that the proportion of females is less in the corresponding occupation than what would have been in females could choose occupations like men. We observe negative numbers in column (7) for professional and service workers in all the four years. This implies that the proportion of females in professional and services would be higher if females could choose occupations based on the male occupational-choice structure. This may be due to differences in educational qualifications between male and female workers, or other barriers (say social norms) which preclude women from certain occupations. In some years, there are negative numbers in column (7) for production workers as well; implying that the actual proportion of females in production would have been higher if there were no difference between the occupational-choice structures of men and women.

Column (8) presents the actual average wage gap between male and female workers in each occupation. We observe that the wage gap is always positive, implying that females earn less than men, and the largest wage gaps are mostly in skilled, service and production jobs. This wage gap can be decomposed into explained and unexplained components as is done in columns (9) and (10) respectively. The occupational attainment model further decomposes the explained and unexplained components into wage and occupational components.

Column (11) presents the explained part of the difference in wages within occupations and the column total is denoted WE. Column (12) presents the unexplained part of the wage gap within occupations and the column total is denoted WD. Column (13) presents the difference in qualifications between males and

females for a job, and therefore is the explained part of job differences and the column total is denoted JE. Column (14) presents the part of the wage gap due to differences in occupational structure between the two groups of workers and the column total is denoted JD. Therefore WE and JE are the explained components and WD and JD are the unexplained components of the wage gap between male and female workers.

We find that the job discrimination component is negative and statistically insignificant, in all the years, implying that the difference in wages (across occupations) caused by differences in occupational choice structures between men and women is not significant. The explained part of job differences is also statistically insignificant in all the four years. However the within-occupation components are both statistically significant in all the years. We find that the unexplained wage differences (WD in column 12) are quite large and greater in magnitude than all the other components. All the within-occupation wage discrimination components are statistically significant. We also notice that the magnitude of this within-occupation discrimination component reduces over the seventeen year period under study. While wage discrimination within occupations remains the primary reason for wage gaps between male and female workers, the extent of wage discrimination and the wage gap itself is reducing over time.

2.7.2 Occupation-based Decomposition for Caste Groups

Results for caste groups are presented in Table 2.16 to Table 2.19. We observe in column (5) that a larger proportion of lower caste workers are in service and unskilled jobs. Column (7) shows the residual difference between actual and

predicted occupational distributions of lower caste workers. The negative numbers for service and unskilled occupations imply that the actual proportion of lower caste workers in these occupations is larger than the predicted proportions based on higher caste occupational choice structure. Columns (11) through (14) decompose the wage gap between low and high caste workers into explained and unexplained components.

We find that the explained job difference and explained wage difference components are the larger components in all the four years for caste workers. However explained and unexplained components of occupational differences are statistically insignificant. We also notice that the magnitude of wage discrimination is much smaller than for gender groups, and is declining over time. Wage discrimination reduces from about 22% in 1983 to 10% of the wage gap in 1999. Therefore most of the wage gap for caste workers is explained by explained wage differences, that is, differences in productive characteristics such as education and experience. Barriers to certain occupations (reflected in the job discrimination component) are found to be unimportant for both gender and caste groups, though the magnitude of this component is larger for caste groups.

2.8 Conclusion

The wage decomposition exercise we undertook in this analysis has shed new light on understanding the reasons for the existence of wage gaps. We can think of several reasons for wage gaps between groups, for example differences in their abilities and productivity; differences in occupations; geographic variations; efficiency wages; or even discrimination. While differences in ability are hard to capture, we control for differences in education and age to account for some of the

productive characteristics of individuals. We find that even after controlling for all observable individual characteristics, and state, industry and occupational differences, there still exists a wage gap. Even if the entire unexplained component cannot be attributed to discrimination, we do have evidence of unexplained wage gaps.[19]

The results from Blinder-Oaxaca (1973) decomposition analyses show that unexplained wages are much more important in explaining wage gaps between male and female workers than between low and high caste workers. Almost half the wage gap between male and female workers is unexplained even after controlling for occupations, industries, workers' and jobs' characteristics.

We re-examine the wage gap by considering the possibility that occupational differences may not be completely voluntary. It might be that females or lower caste workers have lesser access to education or training programs that would allow them to join certain occupations. In this scenario the Blinder technique underestimates the unexplained wage difference. Brown et al. (1980) have developed the strategy to examine wage differences particularly with reference to differences in occupational attainment. That is, their strategy assumes that certain groups (say, females) might have been subject to pre-labor market discrimination. To correct this, they estimate a predicted occupational distribution for the discriminated-against group. They also decompose the total wage gap between groups into components explained and unexplained by wage and occupational differences. The occupation-based decomposition analyses show that the differences in occupational attainment are not statistically significant, and wage discrimination is the primary cause of wage gaps

[19] This is especially true in light of our restricted list of individual characteristics, in particular the absence of any proxy or instrument variable of ability.

for female and male workers, while most of the wage gap is explained for low and high caste workers.

While one might have expected that wage discrimination in India is largely influenced by limited access of minorities to certain occupations, given that they have lower levels of education and face greater social barriers, it turns out that discrimination in access to occupations has a smaller final impact on wages than discrimination in wages within the same occupation. This aspect of discrimination in the Indian labor market is particularly important for female workers. We conclude that within-occupation discrimination in India is particularly important for women and not so important for lower castes, while discrimination in access to occupations is not statistically significant for any of the groups.

These results could explain the effect of the government policies and programs on the extent of discrimination since the previous study in this field by Banerjee and Knight (1985).[20] In particular it is some respite that access to occupations itself is no longer a significant factor. Part of this may be explained by the affirmative action programs, some of which include providing subsidized schooling and higher education, and coaching to prepare for entrance exams, besides reservations in government sector jobs for lower castes. These wage decomposition results indicate that for those who do manage to get the education, there are clearly fewer hurdles in accessing jobs. The Blinder decomposition results suggest that we need to re-examine policy with respect to improving productive characteristics since

[20] Banerjee and Knight had found that there was both wage and job discrimination and wage discrimination was larger in magnitude) in the Delhi labor market.

almost half the wage gap is due to explained differences in educational and other productive characteristics.

The results for caste groups are consistent with implications of the affirmative action model (Coate and Loury, 1993). They establish that in the scenario where affirmative action programs on the part of the government require employers to ensure equal rates of promotions for members of the disadvantaged group (B) and others, there might also be a case of employers patronizing members of the disadvantaged group. If employers feel that B workers are not likely to be qualified, they might lower the standards required to meet the promotion requirements in order to satisfy the affirmative action policy requirement. Then B workers would invest lesser in human capital since they can now achieve promotions with lower standards. This results in lower productive characteristics for B workers as a result of employer patronization and therefore results in a self-fulfilling prophecy of disadvantaged workers being less qualified. In the Indian scenario, there are provisions as part of the affirmative action policies in favor of scheduled castes that allow for employer discretion in lowering qualification requirements for jobs where they feel that enough members of the SC group don't qualify under the regular conditions. Thus our results that almost all the wage gap between low and high caste workers being explained by differences in productive characteristics is consistent with Coate and Loury's prediction that members of the disadvantaged group will end up choosing to invest lesser in human capital.

There were significant changes in policy over the period being studied. Most of the 1980s were a period of industrial controls and regulation in India with pro-big

business policies being the primary emphasis of the government. There was a change in power in 1989 to a more pro-labor government, and simultaneous change in reservation policies as well, with an extension of reservation to more groups of backward castes (and not just those which were originally listed in the Constitution under Scheduled Castes). Furthermore, in mid-1991, following a foreign exchange crisis, India adopted several wide-ranging reforms in keeping with a Stand-by agreement with the IMF. As part of these reforms there was significant liberalization on the industrial and trade front which increased competition. We examine the role of this potential factor in influencing the decrease in wage gaps of these groups in the next chapter. Since almost half the wage gap between male and female workers is unexplained by differences in productive characteristics, Chapter 3 focuses primarily on examining how increased competition affected relative wages of female workers.

Chapter 3 The Impact of Trade Liberalization on Gender and Caste Groups in India

3.1 Introduction

The analysis in this chapter uses changes in trade policies in India, beginning in 1991, to test the effect of the resulting increase in competition on wage gaps between different groups. While wage gaps can be generated by several factors, one of the factors examined in the literature in India is discrimination. Taste based theories of discrimination (Becker, 1957) predict that increased competition should reduce the scope for discrimination, since discriminating firms don't employ inputs efficiently. We use the trade liberalizing reforms implemented by means of lower tariff and non-tariff barriers as an exogenous increase in competition and evaluate its impact on relative wages of females and lower caste workers. We find that industries that experienced larger reductions in trade barriers experienced a proportionately higher reduction in the wage differential between men and women. The wage differential between low and high caste workers however wasn't significantly affected by trade liberalization.

Discrimination models include neoclassical models like Becker's (1957) model of taste based discrimination and statistical discrimination models proposed by Phelps (1972) and Arrow (1972). Taste-based discrimination models assume that some agents dislike working with certain others and therefore are willing to pay a premium to avoid contact with disliked factors. For example, employers who dislike certain workers are willing to lose profits and not employ such workers, or employ them only at a lower wage than other workers. The statistical discrimination model

assumes that information about individuals is inadequate and therefore commonly known characteristics of the group they belong to may be used to infer information about individuals.

Becker's model predicts that as the economy becomes more competitive, employer discrimination should decline since rents that are implicitly used to 'pay' for a firm's discriminatory behavior are reduced. In the long run, with free entry and zero profits, the existence of potential non-discriminating firms implies that discrimination should not be sustainable in equilibrium. If some of the labor market discrimination in India is taste-based employer discrimination, then we should see a decline in discrimination as competition increases. Higher caste employers may dislike hiring lower caste workers, and male employers may dislike hiring female employees to conform to social norms.[21] It is plausible that there exists taste-based employer discrimination in the Indian labor market. We examine if the long-run (increased competition) implications of Becker's theory of employer discrimination are validated in the Indian case. The key to testing Becker's theory is to find instances of increases in competition in the Indian economy. Trade reforms of 1991 provide such an instance. The underlying premise is that trade reforms increase domestic competition and therefore create pressures on all employers to set wages competitively and to employ factors of production efficiently.

[21] In India, traditional Hindu society divided people into social classes based on the caste system. A person's caste identifies his social and economic status. Traditionally lower castes were restricted to menial jobs and their presence was considered to defile a higher caste person. While the caste system was abolished after independence, it continues to play an important role in modern Indian society. Women on the other hand have traditionally been restricted to remaining outside the formal labor market. It is considered a taboo if women have to work to supplement household income. This perception has been changing and probably more so in urban areas.

Trade reforms were undertaken in India in the face of a foreign exchange crisis. But the reforms faced opposition from several fronts which feared that opening the economy would worsen inequality. Given the debate on advantages and drawbacks of the reforms, it is also an important policy exercise to examine the impact of the reforms on wages. Studies have examined the effect of trade reforms in India on wage inequality between skilled and unskilled workers (Banga, 2005); poverty (Topalova, 2005); and industry wage premiums (Mishra and Kumar, 2005). Existing studies provide mixed evidence of the effect of trade reforms. Assuming that one of the reasons for wage gaps between groups is the existence of taste-based discrimination by employers, the increased competition brought by the reforms should reduce the wage gap. Particularly, the extent of the reduction in the wage gap should vary according to the change in the degree of competition. Therefore, larger effects on the wage gap should be seen in sectors that experienced larger reductions in tariff protection and were relatively more concentrated before the reforms.

We contribute to the empirical literature on testing the neo-classical theory of discrimination. Black and Brainerd (2004) found statistically significant reductions in the residual gender wage gap as a result of increased competitiveness in previously more concentrated sectors in the United States. However, Berik et al. (2003) found increases in the gender wage gap in more concentrated sectors that were opened up to greater competition in Taiwan and South Korea. Artecona and Cunnigham (2002) did not find any statistically significant effects in Mexico. By extending the analysis to India, we add to the debate on the impact of increased competition on the gender and caste wage gaps.

We examine the impact of the reforms on wages of female relative to male workers, and lower caste relative to higher caste workers. While results from Chapter 2 indicate that wage gaps between low and high caste workers is mostly explained by differences in productive characteristics, we explore caste groups to confirm that there is no particular effect due to reforms. We use a nationally representative dataset, spanning almost two decades, with information on individuals prior to and after the reforms. We examine relative wage differentials over time, after accounting for an array of productive and demographic factors, as a measure of potential discrimination. We use data on tariff and non-tariff trade barriers to ascertain the change in the degree of protection offered to each industry. Both types of trade barriers were reduced as part of the reforms in 1991. Trade reforms affected both manufacturing and agricultural sectors. Information on tariffs is however unavailable for agriculture, and therefore we focus our attention on the manufacturing sector. Each industry experienced lower protection but the extent of the decrease varied based on initial levels of protection. While past studies on discrimination among gender and caste groups in India have focused only on specific states or even cities, we conduct the analysis for the entire country.

We find that wage differences between female and male workers were reduced more in sectors that experienced larger reductions in the degree of protection. On the other hand, individual level results are insignificant for caste groups. We find evidence of reductions in wage gaps consistent with Becker's model for gender groups. We find an increase in wages of unskilled women relative to unskilled men in the less protected sectors, but there is no statistically significant effect on skilled

workers. This is also consistent with predictions of traditional trade theories for an unskilled labor abundant country like India.

The chapter is organized as follows: Section 2 examines past literature; Section 3 describes the conceptual framework; Section 4 describes the data; Section 5 presents the methodology; Section 6 presents the individual level results; and Section 7 concludes.

3.2 Past Literature

Trade reforms have been a major aspect of development in several countries throughout the second half of the twentieth century and numerous studies have been devoted to examining the impact of more openness in previously protected economies.[22] The focus has been on the effects on worker productivity, growth, poverty, employment and wages, but evidence is mixed. Theory is ambiguous on expected effects on poverty since labor market rigidities might prevent or slow down the re-allocation of factors across sectors. Topalova (2005) studied the effect of the Indian trade reforms and found an adverse impact on poverty in states with inflexible labor regulations, while there was no overall effect on inequality in India as a whole. Niimi et al. (2003) examined the growth channel of the effect of trade reforms on poverty in Vietnam and found an increase in incomes of the poor employed in certain sectors where trade volume rose and prices of tradable goods increased, thereby reducing poverty. Goldberg and Pavcnik (2004b) who examined the effect on employment conditions and wages didn't find any effect of trade reforms on poverty in Colombia.

[22] See Goldberg and Pavcnik (2004a) for a detailed review on impact of trade reforms in developing countries.

While most of the East Asian economies experienced reduction in wage inequalities after trade liberalization, this has not been the case everywhere. Studies on Mexico (Feliciano 2001, Harrison and Hanson 1999, Revenga 1997) found worsening wage inequalities following trade reforms, with an increase in the relative wages of skilled workers. Evidence of skill-biased technical change was found in Colombia (Attanasio et al. 2004), in addition to evidence that trade reforms affected industry wage premiums, with sectors that became more open experiencing larger decreases in wage premiums. Mishra and Kumar (2005) found evidence that reforms increased industry wage-premiums in sectors employing relatively more unskilled labor, and therefore reduced wage inequality.

3.2.1 Effect of Trade Reforms on Discrimination

Literature on the impact of trade reforms has also focused on competitive effects expected from neo-classical theories such as Becker's (1957) model, which predicts that competition will force out discriminating employers in the long run. Black and Brainerd (2004) analyzed the impact of increased competition via increased trade on residual gender wage gaps in concentrated relative to competitive industries that were both exposed to trade. Concentrated industries are likely to have greater market power, which they can use to sustain discriminatory tastes. Thus an increase in competition should generate larger wage gap effects in concentrated relative to competitive industries under similar circumstances. They find evidence supporting Becker's theory that discriminating employers find it harder to continue being discriminatory as competition increases, with evidence of relative improvement in the gender wage gaps in concentrated relative to competitive industries in the US.

The studies on effects of trade reforms on gender wage gaps in different countries suggests that effects depend on characteristics of the reforms and the initial conditions in each country (Fontana et al., 1998). Artecona and Cunnigham (2002) examined the effect of trade reforms in the Mexican manufacturing sector on the gender wage gap. They didn't find any statistically significant effect on the gender wage gap in non-competitive industries that were exposed to trade reforms but found that relative wages increased substantially for skilled workers. Since women are mostly unskilled, the gender wage gap is expected to worsen following the reforms. They found weakly significant results that gender wage gap fell in industries that became more competitive as a result of trade reforms.

Oostendorp (2002) found evidence that openness is negatively related to the gender wage gap within occupational categories in a cross-country study. However, evidence to the contrary was found by Berik et al. (2003) in their analysis of Taiwan and South Korea. They compared the effects of trade reforms on wage discrimination in competitive and non-competitive industries. They found that increasing import shares were associated with rising wage discrimination against women in concentrated industries, contrary to implications of neoclassical theory. This is partly explained by a reduction in female employment in concentrated sectors, and therefore in their bargaining power. They inferred that equal pay and opportunity legislation needs to be enforced to achieve improvement in female wages. Joekes (1999) reviews the literature on effects of trade reforms on gender in different realms, such as employment, wages, and greater empowerment within the household. She observes a positive relation between trade expansion and women's employment and relative

wages, with the highest impact in the lowest income countries. See Fontana (2003) for a more detailed review of gender effects of trade liberalization.

3.2.2 Contribution of this Study

Clearly no generalizations can be drawn on the effects of trade liberalization on different groups. Each country produces specific results based on its unique reform experience and underlying institutional characteristics and resource endowments. While there were no significant effects on the gender wage gap in Mexico (Artecona and Cunningham, 2002), there were negative effects in Korea and Taiwan (Berik et al., 2003). Manufacturing sector reforms in these countries, for example, generated outcomes based on different channels. While the skill-premium increased in Mexico, females lost employment and bargaining power in Taiwan and Korea.

Our analysis in this chapter adds to the current empirical literature on the effects of trade reforms by extending the analysis to testing the neo-classical theory for India, and specifically examines the impact on gender and caste groups. It is plausible that taste-based employer discrimination exists in India against female and lower caste workers based on social norms. The caste system has been the basis of studies of discrimination, but not the focus of trade-reforms related studies, which this analysis aims to test directly. We use individual level data along with information on trade statistics that serve to identify the degree of openness of a sector, to examine the impact of increased competition, via more openness, on wage inequalities for gender and caste groups. Our analysis also contributes to the current policy debate on the effect of increased openness on the Indian economy since we find an improvement in

the relative wages of female workers in sectors that became more open as a result of the trade reforms.

3.3 Conceptual Framework

The primary reason for wage differences between groups is the difference in observable characteristics like educational attainment, years of work experience, training and occupational characteristics. Previous studies (Sambamoorthi, 1984; Banerjee and Knight, 1985) and our analysis in the previous chapter have found that even after controlling for differences in observable productive characteristics such as education and experience, much of the wage gap is still unexplained.

The strategy we use in this study is to test the implication of Becker's theory of discrimination. Becker's model predicts that as the economy becomes more competitive employer discrimination should decline, subject to constant returns production technology, absence of nepotism and government intervention. Since there are neither perfectly competitive labor markets, nor instances without government interventions in markets, it is hard to test this prediction. It may be more reasonable to test for a negative relation between competition and discrimination (Gersen, 2004). This would be best tested where one can identify a change in the degree of competition and the corresponding change in discrimination. Trade reforms of 1991 provide such an instance of an increase in competition affecting the manufacturing sector of the economy. Taste-based discrimination in India is plausible if higher caste employers dislike hiring lower caste workers and male employers dislike hiring female employees to conform to social norms. As mentioned earlier, results from Chapter 2 imply that much of the wage gap between low and high caste workers is

explained by productivity differences, and therefore imply lesser evidence of taste-based discrimination by higher caste employers. We however check the case of caste workers to verify that this is indeed the case.

It is generally believed that increasing international competition in a previously protected economy will force the domestic producers to become more competitive. Levinsohn (1993) terms this *imports-as-market-discipline hypothesis*. The hypothesis assumes that firms are technically efficient and predicts that price-marginal cost mark-ups for previously imperfectly competitive firms will fall as a result of trade liberalization via lower tariffs and quotas. Several formal trade theories provide different implications of an economy opening to trade subject to different conditions pertaining to factor mobility, resource endowments, etc. While there are conflicting theories on the impact of more openness, the underlying premise in this study is that trade reforms increase domestic competition and therefore create pressures on all employers to set wages competitively.[23]

We use individual level data to obtain wage and employment information which is combined with information on trade barriers to ascertain the change in the degree of competition faced by each manufacturing sector. Assuming that one of the reasons for wage gaps between groups is the existence of taste-based discrimination by employers, it is expected that the increase in competition will reduce the wage gap. In this analysis, we assume that there already exists a wage gap between groups and test for an improvement in relative wages between groups. The extent of the decrease

[23] Aghion and Howitt ("*Endogenous Growth Theory*", 1998, Cambridge, Massachusetts, MIT Press) discuss the possibility that openness in the international arena may not always lead to growth and development in a country unless accompanied by knowledge accumulation and institutions that provide incentives for technological innovations (Betancourt and Seiglie, 1999).

in the wage gap would vary according to the change in the degree of competition. We expect larger effects on the wage gap in a sector that experiences larger reductions in tariff protection, and therefore experiences a larger increase in competition relative to other sectors. We find evidence consistent with this expectation for gender groups.

3.4 Data

3.4.1 Individual Level Data

The individual level data comes from the *Employment and Unemployment Survey* (*Schedule 10*) of the *NSSO* (National Sample Survey Organization) of the Government of India. The data is described in detail in Chapter 2, Section 2.3. Data is available for the years 1983, 1987-88, 1993-94 and 1999-2000. Thus it covers periods both before and after the 1991 reforms. The data are available as a time-series of cross-sections of households.

The analysis is restricted to the urban sector, which has more reliable data on wages and more individuals in the manufacturing sector than rural areas. The tariffs were relaxed for the manufacturing sector industries post-1991 as part of the reforms. The analysis is however restricted to 72 (of 200) 3-digit manufacturing sector industries for which information on tariff and non-tariff barriers is available. The analysis is restricted to employed individuals in the age group of 15 to 65 years. The analysis is further restricted to 16 States which are typically covered in studies on India. The omitted States are in north-east India, where frequent insurgency problems may have affected data collection. Topalova (2005) reports that there is evidence of very little mobility among workers and that there are no significant spikes in

migration post-reforms. This also limits the scope for unobserved changes in the composition of the population. The list of States is available in the Data Appendix.

Wages, reported in the survey as weekly earnings, have been converted to constant rupee terms using consumer price indices from the International Financial Statistics. The top and bottom 1% of wages are trimmed to omit outliers. Education variables are defined as dummy variables for each level of education completed. There are 10 education categories: those not literate (*notlit*); literate but with no formal education (*noformal*); those with below primary education (*belowpr*); completed primary schooling (*pr*); completed middle school (*mid*); completed secondary school (*sec*); graduate with agricultural degree (*agrigr*); engineering graduate (*engr*); medicine graduate (*medgr*); graduate in other subjects (*othergrad*). The omitted category is *notlit*. Age variables are included as a fourth order polynomial. *Marital status* is a dummy that equals one for those currently married and is zero otherwise. *Household head* equals one if the individual is the head of the household, and zero otherwise. Information on social group identifies whether the individual belongs to a scheduled caste. Scheduled Castes form about 17% of the population. Information on household religion identifies whether the individual is a Muslim. Muslims are the second largest majority and comprise about 13% of the Indian population. The results presented here omit Muslims to distinguish effects more clearly for gender and caste groups without further confounding effects on females between Muslims and others. Therefore we restrict the analysis to Hindus.

Summary statistics for individuals employed in the 72 manufacturing industries are presented in Table 3.1. Summary statistics reveal that women on

average earn significantly less than male workers, and the average relative wage ratio improved from 0.27 to 0.33 over the time period under analysis. Female workers are also younger, and less educated than the male workers. Lower caste workers in fact have higher earnings than females. The relative wage ratio of lower castes to higher caste workers improved from 0.76 to 0.83 over the entire period. Lower caste and female workers have lower levels of education than their counterparts, but there is improvement over time for all the groups. The numbers also indicate the educational achievements by caste and gender. We observe that the educational levels completed have improved for each group between 1983 and 1999-2000.

Detailed information is available on the activity status of individuals, who can be self-employed, in salaried employment, or seeking work. Information is also available on the occupation within each industry, and this is used to construct occupation categories based on Banerjee and Knight's (1985) analysis. Individuals are classified into six occupations, namely professional, skilled, clerical, service, unskilled, and production. Occupation dummies are included in some of the regression specifications and the omitted category is skilled workers. The industry classification is given by the *National Industrial Classification (NIC)* of the Government of India. Concordance tables were used to convert all the industry codes into their *NIC-87* equivalents. The list of the 72 three-digit manufacturing industries is provided in the Data Appendix.

3.4.2 Trade Data

Tariff data is obtained from Das (2003). Das computes the Corden measure of the effective rate of protection (ERP) for 72 3-digit Indian manufacturing industries,

for the four phases 1980-81 to 1985-86, 1986-87 to 1990-91, 1991-92 to 1994-95, and 1995-96 to 1999-00. He also calculates the import coverage ratio and the import penetration rate for these industries for the same period.

The Corden measure of **effective rate of protection** equals the percentage excess of domestic value added over foreign value added, due to tariffs, which Das (2003) calculates as follows:

$$\mathbf{ERP_j = (VA^*_j - VA_j) / VA_j}. \qquad (1.12)$$

VA^*_j is the value added at free trade prices, and VA_j is value added at tariff distorted prices for final product j.[24] The effective rate of protection measures the protection to domestic factors of production based on tariffs on both input prices and output prices, and thus is a better measure than the nominal rate of protection which doesn't account for tariffs on inputs. Higher ERP therefore imply higher degree of protection. It must be noted however that in a country like India, which relied heavily on quantitative restrictions (and other non-tariff barriers) in addition to high tariffs, the effective rate of protection alone might not give a complete picture of the degree of protection.

The **import coverage ratio** (ICR) is a measure of the frequency of non-tariff barriers (NTBs) weighted by imports or by production. Das (2003) defines the Import Coverage Ratio as $C_i = \sum D_i M_i / \sum M_i$, where j is industry and i is a particular product line within an industry; D_i is a dummy variable which equals one if the product is listed under restricted (or banned/limited permissible/canalized) imports, and zero if

[24] The value added functions are defined as VAj = (1-∑aij) and VA*j = (1+tj) - ∑(1+aij), such that ERP can be expressed as ERPj = (Tj - ∑aijTi) / (1 - ∑aij), where j is the j'th activity or product, Tj is the nominal tariff rate for j'th product, Ti (i=1,2,…,n) are the nominal tariff rates of tradable intermediate inputs used in the j'th activity, aij (i=1,2,…,n) are the cost shares of inputs in production of j'th activity, ∑aij is sum of shares of intermediate inputs in the final value of j, and is a weighted average of input tariffs on all intermediate inputs with weights according to input shares (Das, 2003).

the product is on the open general list (and is freely importable); M_i is the value of imports of the i^{th} product category which is subject to NTBs; and $\sum M_i$ is the sum of the value of imports of all the product lines within the industry. The change in the import coverage ratio gives some idea of the trend in NTBs, but does not capture the exact price advantage to domestic producers from such protection. Higher import coverage ratio indicates greater protection.

The **import penetration rate** (IPR) measures the combined effect of both tools of protection, namely tariffs and NTBs. It is expected that lower tariffs along with lower NTBs (for example more items shifted from the restricted list to the open general list) will raise imports and vice-versa. The import penetration rate is defined as the ratio of industry imports to domestic availability, where domestic availability is defined as domestic production plus imports minus exports. Exports and imports of product lines are aggregated to obtain the industry levels. Das (2003) defined the import penetration rate as $MPR_j = M_j / (P_j + M_j - X_j)$, where j is industry; P is domestic production; M is imports; and X is exports. Higher import penetration rates imply lower levels of protection.

Summary statistics of trade measures are presented in Table 3.2. Trends in ERP in 1983 and 1999 are shown in Figure 2. The summary statistics reflect substantial declines in both tariffs and non-tariff barriers. The average effective rate of protection declined from 115% to about 40% over the period under analysis, and the average import coverage ratio declined from 97% to 25%. While in the first phase, nearly 70% of industries had ERP in the range of 50 to 150%, in the last phase of reforms, almost 80% of the industries had tariffs in the range of 0 to 50%, and

none of the industries had tariffs above 100%. Almost 92% of industries had 100% import restrictions in the first phase, which declined to 5 industries in the last phase. There weren't large changes in the import penetration rate (Das, 2003).

3.4.3 Annual Survey of Industries Data

This dataset obtained from publications of the *Economic and Political Weekly* in India is used to get information on the number of firms, and the gross fixed capital formation for each industry. The number of firms is used to generate a measure of industrial concentration to account for changes other than changes in trade protection levels for each industry. The results presented are based on using the *percentage change in the number of firms* between 1983 and 1999 relative to 1983 as the measure of change in industry concentration. The gross fixed capital formation provides information on all the new physical investments in an industry and is used to control for changes in terms of foreign direct investments and increased domestic investments as a result of trade reforms and industrial de-regulation. Data is available at the 3-digit industry level, which is the level of aggregation of the trade data as well.

3.5 Methodology

Focusing on employer discrimination, Becker's model implies that discriminating employers earn lower profits than non-discriminating employers, which cannot be sustained in the long run under perfect competition. Thus perfect competition in the long run would eliminate employer discrimination.[25]

[25] Customer and employee discrimination can however persist in the long run.

Trade liberalization in India in 1991 provides an instance of a significant change in the level of competition in the economy. It is therefore expected that firms that could indulge their preferences pre-reforms when trade protection allowed super-normal profits may be unable to continue doing so once competition eliminates those profits. The trade reforms are used to test the impact on relative wage differentials between men and women, and between low and high caste workers. To test for a negative relationship between competition and discrimination, we look at the effect on individual's wages in industries that experienced a reduction in tariffs. We expect that industries experiencing larger decreases in tariffs, since they become more competitive, would see a larger reduction in discrimination, measured in terms of relative wage differentials between the groups of interest.

3.5.1 Individual Level Analysis

The individual level analysis involves standard Mincerian earnings regressions with log of real weekly wages as the dependent variable, demographic characteristics (such as gender and caste dummies, age, education, marital status, sex of the head of the household, occupation, etc.) and measures of protection in each industry. Gender and caste dummy variables indicate whether these groups get higher or lower returns than the comparison group. The impact of trade reforms on gender (caste) groups is captured by the coefficient on the interaction term between the trade protection measure and the *female* (*caste*) dummy variable, which is expected to be negative. The negative coefficient would indicate that lowered protection raises wages of female (lower caste) workers relative to male (higher caste) workers.

The main equation used to assess the impact on wages is the following within-group regression:

$$Ln\ w_{ijtk} = \alpha + \beta(X_{ijtk}) + \gamma_0(Female_i) + \gamma_1(Protect_{jt}) + \gamma_2(Protect_{jt} \times Female_i) + \lambda_t + \phi_k + \theta_j + \psi_{jt} + \varepsilon_{ijtk}. \quad (1.13)$$

In the above equation, *ln* w_{ijtk} is the natural log of weekly earnings of individual *i* in industry *j* in time *t* in state *k*; *X* consists of the individual characteristics comprising age (in quartic), nine education dummies for each level of education completed, marital status (equals one if married and zero otherwise) , and household head dummy (equals one if head of household, zero otherwise); *Female* is a dummy variable which equals one if individual is a female and zero otherwise; *Protect* includes the tariff and non-tariff measures of protection of the industry; ϕ, λ, and θ denote state, post-liberalization and industry fixed effects, respectively. The above equation which is written for gender groups is extended by adding a *caste* dummy variable and its interaction terms to analyze the effect for caste groups. Since the tariff and non-tariff barriers are reported at the industry level, standard errors are corrected by clustering at the industry level. The regressions are also weighted using sampling-weights.

By including the *female* (*caste*) dummy variable, the method assumes that there exist wage differences between these groups. *Protection* measures indicate the overall impact of the change in protection on wages in the economy across groups. This method assumes that lower effective rate of protection and import coverage ratios, and higher import penetration rates imply increased competition within an industry. The negative of the import penetration variable is used to simplify

comparison between coefficients. The coefficients of the protection measures are expected to be negative if lower protection raises wages.

The *protection* measures interacted with *female* are the main variables of interest. The coefficient of *protection x female* indicates the impact of a change in the degree of protection on the relative wages of females across industries. Increased competition in the sectors experiencing more openness will cause employers to be less discriminatory and therefore females will experience an increase in their wages relative to male workers. We expect *protection x female* to be negative.

The regression is run for the full sample consisting of males, females, low and high caste workers, and also separately for gender and caste groups. The full sample is restricted to Hindus to simplify the analysis. In looking at the gender groups alone, the sample is restricted to the higher castes, while in examining the caste groups the sample is restricted to males.

Levels of protection vary across industries and over time. Across-industry (cross-section) differences in levels of protection indicate the correlation between wages and levels of protection, which would also be a function of other characteristics of the industry. It would not account for the effect of a change in protection within an industry. Since the aim is to analyze the effect of different magnitudes of change in protection over time rather than differences in level of protection across industries in any one year, industry fixed effects are included to focus on within-industry comparisons. While protection levels have changed in each of the four years (protection was increased between the first two years of data, and decreased after 1991) under analysis, we intend to capture the effect of the trade

reforms in 1991. Therefore *industry x post-liberalization* interaction terms are included to account for changes in the post-liberalization period. An interaction term *Industry x PostLib x Female* is also included to account for differential returns to men and women of changes in post-reform period across industries.

3.5.2 Additional Controls

We include *Industry* x *female* and *education* x *protection* measures to account for differential returns to women workers across industries, and differential effect of protection on people with different educational achievements, respectively. We also include *state* fixed effects in the regressions. States in India have significant linguistic barriers that make migration across states more difficult.[26] While states vary in terms of their levels of industrialization, the changes in trade protection were at the industrial level and applied uniformly across states. Since labor markets are considered to be differentially rigid across states, it is meaningful to include *state* and *female* (*caste*) interaction terms to account for different experiences across states.[27]

3.5.3 Limitations

The analysis here focuses on looking at the wage effect. The labor force participation rates of women have not changed much during the entire period under study. Thus there need not be concerns that the results obtained are being driven by a significant increase in the number of women entering the labor force. It is possible

[26] While seasonal migration between states might occur, this is mostly in the rural and agricultural sector than in organized manufacturing.

[27] The results with inclusion of *state x female* and *state x caste* are not presented here, but it doesn't change the results much, though the *state x female* and *state x caste* coefficients are jointly significant in each set of regressions for gender and caste groups respectively.

that the impact on employment probabilities of women are more likely than the wage impact, but the analysis in this chapter focuses on wages.

Inclusion of protection measures accounts for differences in degree of protection and therefore in the resulting degree of competition across industries. However changes in competition between industries because of industrial de-regulation may bias the estimates by overestimating the impact of lowered trade protection. While trade reforms began in mid-1991, industrial de-regulation had begun in the late 1980s. Ideally the regression should directly control for changes in the industrial structure resulting from de-regulation so that effects of trade reforms can be clearly and separately analyzed. The degree of concentration of the domestic industry would be a useful measure in this context. The domestic economy would become less concentrated over time, both due to de-regulation and due to trade liberalization. Unfortunately the data has no direct measure of industrial concentration. The *percentage change in the number of firms* within an industry over time is used as a crude measure to capture the change in degree of concentration within that industry. Traditional industry concentration measures such as the *Herfindahl Index* and the *Concentration Ratio* account for market share of each firm. While the measure constructed here is not able to capture market share, it is included nevertheless to capture changes in industrial structure. While inter-industry comparisons wouldn't be of much use in this case, including this measure would nevertheless provide additional information. Weighting the number of firms by industry gross value added can enable cross-sectional comparisons across industries,

since industry A with fewer firms and higher gross value added than industry B is likely to be more concentrated.

It is possible that political clout of bigger firms in an industry prior to reforms in 1991 could have influenced the actual level of tariffs that was fixed for that industry.[28] Since the reforms were effectively imposed in India as a condition to receiving aid from the IMF to avoid a financial crisis, the political influence of industries should not influence tariff levels decided as part of the reforms.[29] Apart from offering protection, the high tariffs were also used to generate revenue. The decision regarding which industries continued getting protection was determined on the basis of considerations of strategic importance and products which were important for mass-consumption, rather than considerations to favor the bigger firms across industries. Thus industry-level changes are expected to have been guided by concerns of economic policy more than nepotism. The *change* in the protection levels can therefore be considered exogenous and therefore OLS should generate consistent estimates. Tariff and non-tariff barriers were reduced across the board for all industries and our analysis captures the effect of the difference in the extent of decrease in protection across industries.

[28] Bertrand (2004) found evidence that increase in foreign competition in the US changed the sensitivity of current wages to current unemployment rate, with larger effects in more financially constrained industries.

[29] The study assumes that political influence of different firms doesn't affect wage-setting. The public (government controlled) sector was pre-dominant in organized sector production and therefore the government influenced wage setting not only in the public sector but also in the private sector. Collective bargaining is therefore generally believed to be less important than the role of government in wage setting in India. Minimum wage laws were imposed in both organized and unorganized sectors. Also, labor laws in India are quite rigid. The Industrial Disputes Act of 1947 made it very hard for firms to fire workers, with firms employing more than 100 workers requiring prior permission, which was seldom granted, from state governments.

An improvement in women's unobserved characteristics could also help explain their relative wage increase. But as long as there is no difference in the improvement of these characteristics (for example, women stay longer in the labor force) across industries or sectors based on the openness of those sectors, it cannot explain differences in results across sectors since uniform effects may be expected in all sectors. Given the lower social status of women (and lower castes), it is likely that the women who choose to enter the labor market are differently advantaged relative to those who choose not to work. The analysis should ideally correct for such a selection bias, which is part of the extensions we intend in the future.

3.6 Results based on Individual Level Analysis

We present the results for the manufacturing sector based on regression equation 1.13, for the entire sample and for gender and caste groups separately. The sample is restricted to Hindus.

3.6.1 Results for Entire Sample

Results are presented in Table 3.3 and Table 3.4. Each column represents a separate regression, with the same dependent variable (log of real weekly earnings). Starting with the simplest specification in column (1) of Table 3.3, which includes demographic characteristics, *female* and *caste* dummy variables, explains almost 50% of the changes in log weekly earnings. Note that the coefficient of the *caste* dummy is positive and significant in the early specifications. We also run regressions (not reported here) similar to the specification in column (1) for skilled and unskilled workers (these regressions don't include *education x female* and *education x caste*).

Skill groups are defined on the basis of education levels completed. If skilled workers are considered to be those with more than high school education (that is, have at least some college education), then earnings differentials between females and males are statistically significant only for unskilled workers. The coefficients of *PostLib* and *PostLib x Female* are both positive and statistically significant for unskilled workers while they are both insignificant for skilled workers. This provides evidence that unskilled female workers benefited post-liberalization. The earnings differential between castes is statistically significant only for skilled workers.

Column (2) of Table 3.3 includes the ERP measure of protection, *ERP x caste* and *ERP x female* to capture the effect of lowered protection on wages of each group. *ERP x female* and *ERP x caste* are both statistically significant and negative as expected. The magnitude of the effect of lower protection is larger for females relative to males, than for lower castes relative to higher castes. The coefficient of *ERP x female* is -0.197 and is -0.05 for *ERP x caste*, which implies a 1.97% increase in relative wages of females, and a 0.5% increase in relative wages of lower castes when the ERP is reduced by 10-percentage points.

Industry fixed effects are included in column (3) to focus on within-industry effects. The inclusion of industry fixed effects reduces the magnitude of the effect of lower protection slightly for females, while effects on both females and lower castes remain statistically significant. *Industry x PostLib* effects are included to account for changes across industries before and after reforms, along with *PostLib x Female* and *PostLib x Caste* to account for changes affecting the groups differently, before and after reforms. The effect of lower protection on caste groups becomes statistically

insignificant once *PostLib* is included, while the effect on females remains statistically significant.

Additional specifications in columns (5)-(8) include *Industry x Female*, *Industry x Caste* to account for differential returns to groups across industries; and *Industry x PostLib x Female* and *Industry x PostLib x Caste* to account for returns to groups across industries before and after reforms. *Occupational* dummies as well as their interactions with female and caste dummy variables are included in columns (7) and (8). Across all the specifications, there is a statistically significant effect of lower ERP on relative female wages, but there is no significant effect of lower ERP on relative wages of lower caste workers (once *PostLib* is included). The last column includes the import penetration measure, which however doesn't have a significant effect on relative wages of either group, as indicated by coefficients of *–IPR x female* and *–IPR x caste*.

Table 3.4 presents additions to specifications in Table 3.3, for the entire sample. These specifications include all three measures of protection. *Protection x PostLib* is included in some specifications to account for the overall effects of lowered protection before and after liberalization. Table 3.4 also includes the industry concentration measure of *percentage change in number of firms*, and the *gross fixed capital formation* of industries. *ERP x education* are seen to be mostly jointly insignificant. The overall conclusion from all the specifications is that there is a significant effect of lowered protection on female relative to male workers, while there is no significant effect for lower caste workers.

Protection x PostLib x Female captures the effect of lowered protection post-liberalization for women relative to men. The negative coefficient on this multiple interaction term implies improved outcomes for female workers in less protected sectors post-reforms. Employers in these sectors may experience greater competitive pressures from other sources as well in the post-reform period, one of which may be de-regulation of the domestic economy. *Protection x PostLib x Female* are statistically significant using both *ERP* and *IPR* measures and plausibly indicates that other changes (such as domestic de-regulation) in the economy in the post-liberalization period had a heightened effect in the less protected industries than others. The combined effect of lower protection on female wages relative to male wages, in terms of *ERP* is the sum of the effects on *ERP x female* and *ERP x PostLib x female*. This equals -0.747 [= -0.182 -0.565], which implies that for a 10-percentage point fall in the ERP there is a 7.5% increase in wages of females relative to male workers. *IPR x PostLib x Female* is statistically significant and the combined effect of a change in the *import penetration rate* is -0.292 [= -1.720 + 1.428], which implies that for a 10-percentage point increase in the import penetration rate, there is a 2.9% increase in the relative wages of female workers. There is however no significant effect on lower castes. An interaction term between *female* and *caste* dummies is included in the last two specifications but was statistically insignificant. The next section analyzes the effect on gender groups.

3.6.2 Results for Gender Groups

The sample used to analyze the impact on males and females is restricted to higher castes to avoid problems of differential impact on these groups across higher

and lower castes. Results are presented in Table 3.5, Table 3.6 and Table 3.7. The first column of Table 3.5 shows the simplest specification for gender groups. Including all demographic characteristics of age, education, marital status, etc. explains 47% of variations in log weekly earnings. A *female* dummy variable is included to capture relative differences in earnings between gender groups. The coefficient of *female* is negative and significant and implies 57% lower earnings for female relative to male workers. The second column includes the *ERP* measure of protection to capture the effect of lower protection on wages. The coefficient of *ERP* is -0.086 and indicates that a 10-percentage point decrease in the *ERP* increases wages for everyone by 0.8%. Next, *industry* fixed effects are included in column (3) to focus on within-industry variations. This raises the effect of lower protection to a 1% increase in wages for a 10-percentage point decrease in *ERP* as reflected by the coefficient of *ERP* (-0.110). We also calculate the impact of lowered protection in lowering the female-male wage gap reported in Chapter 2. The actual decrease in the female-male wage gap between the pre- and post-liberalization periods is 0.163. We use Blinder decomposition technique to determine the impact of lower protection post-liberalization measured by change in ERP on the wage gap and find this impact to be 0.034. Therefore lowered protection explains about 21% (equals 0.034/0.163) of the decrease in the female-male wage gap.

Since reforms were introduced in 1991, to account for changes pre- and post-reforms, a *PostLib* dummy variable is included. While the coefficient of *ERP* (-0.070) is still statistically significant, the magnitude is lower. *ERP x female* is included next in column (5) to account for the differential effect of lower protection on women

workers compared to male workers. The coefficient of *ERP x female* is negative and statistically significant and indicates that for a 10-percentage point decline in the effective rate of protection, there is a 2% increase in wages of female relative to male workers. To account for changes across industries in the post-liberalization period compared to the pre-reform period, *industry x PostLib* is included, which reduces the impact of lower protection on relative female wages to 1.67% for a 10 percentage point decrease in ERP. When state fixed effects are included (not reported here) to the specification in column (6) of Table 3.5, it doesn't change the magnitude or significance of the primary variable of interest, *ERP x female*.

Including Other Measures of Protection

Table 3.6 presents additions to specifications in Table 3.5 that include other measures of protection. Column (7) includes only the IPR measure which is statistically insignificant, but *–IPR x female* is significant and negative and implies that when the import penetration increases by 10-percentage points, the wages of female workers increase by 5.5% compared to wages of male workers. Column (8) includes both ERP and IPR measures. The *import penetration rate* is statistically insignificant in all the specifications while the *ERP* is statistically significant. However, *ERP x female* is mostly insignificant once other measures are included, while *IPR x female* is mostly significant. Thus lower protection through increased penetration seems to have a greater effect on relative female earnings than lowered tariffs. *Industry x female* is included since females may be impacted differently across industries. The coefficient of *negative IPR x female* in column (11) is statistically significant and indicates that for a 10-percentage point increase in the import

penetration rate (which indicates lowered protection) there is about 5% increase in relative wages of female workers. The *ERP x female* coefficient (which is significant in this specification) indicates a corresponding 2% increase in relative female wages for a change in the ERP. *Occupation* and *occupation x female* variables are also included. To account for differential impact on men and women post-reforms across industries, additional interaction terms *industry x PostLib x female* are included in the last two specifications. The *IPR x female* coefficient remains significant but is smaller and indicates a 4% increase in relative wages of female workers when the import penetration rate increases by 10 percentage points. Meanwhile the *ERP x female* coefficient is no longer significant. *Percentage change in number of firms*; and *real gross fixed capital formation* are included in the last specification to account for changes in industry concentration and growth. Inclusion of these industry measures make both *ERP x female* and *IPR x female* insignificant.

Table 3.7 presents additional specifications for gender groups. Column (14) includes *protection x PostLib x female*. These multiple interaction terms provide the effect of other changes in the economy, which interact with changes in protection, to further affect wages. It might be expected that an industry that is de-regulated domestically, allowing existing firms to expand and new firms to enter, will experience greater increase in competition as a result of trade reforms. The coefficient of *ERP x female* is insignificant in all but the first specification. The coefficient of *ERP x PostLib x female* is significant in all but the first specification and indicates that a 10-percentage point fall in the *ERP* raises relative female wages by more than 5% in the post-reform period. *IPR x female* is statistically significant in columns (17),

(18) and (19). The last specification includes the third measure of protection, the import coverage ratio and interactions similar to those for the other measures, but *ICR x female* and *ICR x PostLib x Female* are insignificant.

The *percentage change in the number of firms* is included in columns (16)-(19). It is significant, though small and positive across all the specifications. A positive coefficient on the concentration variable implies an increase in the wages due to increased competition. However, *concentration x female* and *concentration x caste* were insignificant (results not reported here), which implies that there was no differential effect across groups of a change in the degree of concentration of industries. The *gross fixed capital formation* variable is also very small in magnitude but positive and weakly significant.

Potential Explanations

There can be several explanations for the reduction in the wage gap between men and women, for instance higher female educational attainment; occupational changes; supply-side changes such as greater commitment to staying on longer in the labor force; or even lesser discrimination against females. The educational changes are accounted for by including the education dummies.[30] When *education x PostLib* is included, *IPR x Female* is still significant, and *education x PostLib* variables are jointly significant. This accounts for changes in education over time. *Education x female* are jointly significant implying differential returns to education for female relative to male workers. *Occupation x female* is also jointly significant.[31] The latter significant result points to relative improvements in returns for females within

[30] *Education* is a set of 9 dummy variables. *Not literate* is the omitted category.
[31] *Occupation* is a set of 5 dummy variables. *Skilled* is the omitted category.

occupations. Table 3.8 shows the relationship between the percentage of female employment by industry and the percentage change in the ERP between the first and last year. The figures are presented for 5 industries each, with the highest and lowest ERP levels in 1983. There is no systematic relationship between these changes, assuring us that the results are not the by-product of women disproportionately entering industries where the ERP declined substantially. Thus after having controlled for these changes, the significant coefficient of *protection x female* plausibly indicates the actual effect of the increased competition, rather than a spillover from the higher educational achievements of women workers, or changes in industrial composition over time.

Several studies have presented evidence on the existence of differential returns to gender in India. Sambamoorthi (1984) using data from a city in south India found evidence that about 36% of the wage gap between men and women was attributable to discrimination. Madheswaran and Lakshmanasamy (1996) found that over 98% of occupational disparity between males and females was unexplained (and could be due to discrimination by firms, or women's preferences or both) and that the predominant cause of lower female relative wages was within-occupation differences rather than disadvantageous occupational distribution of female workers. Our results for gender groups are consistent with results from the Blinder-Oaxaca decomposition of the wage gap into explained and unexplained parts. The decomposition results point towards greater relevance of a discrimination-based wage gap story for gender rather than caste groups. The wage decomposition numbers had shown that a larger part of the wage gap is explained for lower castes than for females relative to higher

castes and males respectively. This can explain the statistically significant impact of lowered protection on potential discrimination against female workers.

The common association studied between trade reforms and wages operates through differential impact on skill groups. Studies [see Banga, 2005; Dutta, 2005] examining the impact of trade reforms on wage inequalities between skilled and unskilled workers find that wage inequalities have gone up. The educational distribution of women is poorer than that of men.[32] Even after accounting for such differences, there is a significant impact of lowered protection on relative wages of female workers.

For males and females, separate regressions (not reported here) are run for those with more than high school education and for those with less than high school education, which is the usual basis of defining skilled workers. For those with more than high school education (skilled workers), none of the *protection x female* terms is statistically significant, whereas for those with less than high school education (unskilled workers), *IPR x female* is still significant. This implies a beneficial effect of lowered protection on the relative wages of females with less than high school education, while there is no statistically significant improvement in the relative wages of females with more than high school education as a result of higher import penetration. This indicates improvements in returns to the less skilled female workers which are in keeping with our expectation that in an unskilled labor abundant country the returns to unskilled labor should increase if the country has a comparative

[32] While about 40% of female workers in the sample are not literate, only 14% of male workers are not literate. Those who have completed primary schooling form the second highest proportion of female workers (about 18%), while the corresponding group for male workers is those who have completed secondary schooling (about 24%). Female graduate workers comprise only 4% of the sample while male graduate workers comprise about 9% of the sample.

advantage in the production of the unskilled labor intensive good. These results are consistent with Mishra and Kumar (2005) findings that unskilled workers (males and females) benefited from trade reforms in the urban manufacturing sector.

3.6.3 Results for Caste Groups

The sample is restricted to males for this set of regressions. While the *caste* dummy is negative and statistically significant, the interaction terms involving *caste* and other variables are not significant. The results are presented in Table 3.9. The results imply that lower castes do not benefit any differently from lowered protection than do higher castes. This result is consistent with findings from the decomposition of the wage gap into explained and unexplained parts. While wage gaps do exist between low and high caste workers, it doesn't stem from discrimination, rather from poorer educational achievements of lower castes. The individual level regressions are also run separately for skilled and unskilled workers. While *protection x caste* was insignificant for both groups of workers, *protection x caste x postlib* was significant for skilled workers, indicating that there is a significant effect of lowered protection on high skilled low caste relative to high caste workers.

3.7 Conclusion

Previous studies have established the evidence of wage gaps between females and males and between low and high castes. Even after controlling for all observable productive characteristics, the wage gap persists. While the entire wage gap is not necessarily due to discrimination, there is evidence of wage discrimination against females and lower castes in India. While direct evidence in the labor market on

discrimination is scant, there is substantial evidence of its persistence in the social realm.[33] If some of this is taste-based discrimination on the part of males and higher caste employers conforming to social norms, we expect that an increase in competition will reduce the extent of employer discrimination by reducing rents. We use the trade reforms in India as an exogenous change in the degree of competition in the domestic economy to examine the impact of increased competition on wage discrimination against female and lower caste workers.

Using individual level analysis, we find that industries which experience greater liberalization (in terms of larger reductions in the protective barriers) experience a proportionately higher reduction in the wage differential between women and men, but there is no statistically significant impact on wage differentials between lower and upper castes. Lower protection in terms of lower tariffs and greater import penetration increase the wages of female relative to male workers. The result for gender groups is consistent with the implication of Becker's theory of employer discrimination that increased competition should reduce employer discrimination. In addition, there is a heightened effect of other changes in the post-liberalization period on the industries which experience lower protection and therefore on relative wages in these sectors. However, the lack of any direct measures of concentration in the industrial sector precludes an analysis similar to other studies which compare the effect of increased competition across concentrated and less concentrated sectors.

[33] Betancourt and Gleason (2000) found evidence of unequal treatment of lower castes in allocation of publicly provided goods such as health and education in rural areas. Chattopadhyay and Duflo (2003) examined the impact of empowering women in rural areas by reserving posts for them in Gram Panchayats (village councils) to increase their representation in local governance.

After accounting for productive characteristics, we find statistically significant wage differentials between unskilled male and female workers, but no significant wage differentials between skilled male and female workers. We find an increase in wages of unskilled women relative to unskilled men in the less protected sectors, but there is no statistically significant effect on skilled workers. This indicates evidence of decrease in unskilled worker wage differentials for gender groups, in keeping with Becker's prediction. This result is also consistent with traditional trade theories that predict that opening up an unskilled labor abundant country should increase the returns to the unskilled labor input if it is used intensively in the good in which the country has a comparative advantage. There is mixed evidence on the effect of reforms on wage inequality between skill groups in India. Our result is consistent with Mishra and Kumar (2005) who used the same data and found evidence of improvement in wages of unskilled workers relative to skilled workers.

Though *a priori* we expected to find a significant impact of lower protection on lower castes, the results are statistically insignificant for lower castes relative to higher castes. This might be explained by the fact that the analysis focuses on the urban areas where the social stigma of belonging to a lower caste group might be lesser than in rural areas. Also, the affirmative action program of reservations for lower castes in government jobs has contributed to increased opportunities for lower castes in public sector jobs, which might in turn result in their lower numbers in urban manufacturing which might explain the insignificant results.

This result is relevant both as a contribution to the empirical literature on the economics of discrimination and as a policy analysis of the economic impacts of trade

reforms in India. The empirical literature provided mixed evidence with evidence of increased competition leading to a reduction in wage discrimination in the US, while studies for developing countries obtained insignificant or sometimes opposite results. This result bolsters the argument in favor of Becker's theory of discrimination. In terms of policy analysis, the welfare effects of trade liberalization in India are still subject to debate. The result that greater openness has in fact reduced discrimination against female workers therefore contributes to an argument in favor of liberalization in India.

Chapter 4 Conclusion

The analysis in this dissertation documents the pattern and evolution of discrimination against women and lower castes in the Indian labor market over the course of almost two decades. It adds to our understanding of the nature of the discrimination that prevails against these groups.

The second chapter conducts an extensive analysis of discrimination in the Indian labor market among groups that have traditionally been economically disadvantaged. While lower castes have been disadvantaged for centuries due to a well-established social hierarchy, women have been disadvantaged due to social norms which dictated that it was a taboo for women to be working outside the protected environs of her household.

We examine the wage gaps between male and female workers and between low and high caste workers in 16 Indian States over a seventeen year period. We use a time-series of cross-sections of nationally representative data on individuals and households to analyze these changes. The first step is an earnings regression analysis to determine if there is a significant difference in returns between groups that is not fully captured by productive characteristics. We include, in the list of productive characteristics, dummy variables for each level of education completed; a fourth order polynomial in age; allowed for interactions between age and education to control for qualitative differences in education across cohorts; include interactions of education with *caste* and *male* dummy variables to account for differences in educational returns across these groups. We also include *household size* and a *marital status* dummy variable to control for self-selection bias. Among the institutional factors, we

include industry, state and occupation fixed effects to control for the main factors we expect will explain wage differences between individuals. We examine the coefficients of *caste* and *male* dummy variables in wage regressions using different specifications involving the above-mentioned explanatory variables. We find the *caste* and *male* coefficients to be statistically significant in most of the specifications; implying the existence of wage differences not fully accounted for by explanatory factors. However the magnitude of the *caste* coefficient is much smaller than the magnitude of the *male* coefficient. Having established that there are unexplained wage gaps, we examine the magnitude of this unexplained component.

The extent of the wage gap is estimated using two main techniques, namely the Blinder-Oaxaca decomposition and the Brown occupational decomposition techniques. The former method decomposes the wage gap into explained and unexplained components. This method assumes that all the differences in productive characteristics are explained wage differences and the residual component is attributed to discrimination. There is however a drawback to this method if we are not able to control for all observable and unobservable productive characteristics that affect an individual's wages. This is the omitted ability problem. However if we don't expect the unobservable factors to be changing systematically over time, then we can expect it to be influencing wages in each period in the same manner, such that analysis of changes over time are not affected. Using Oaxaca-Blinder decompositions, we find that, once industries and occupations are controlled for, observable characteristics of workers and jobs explain almost the entire wage differentials between castes. The result for gender groups is however different: almost

half the wage gap between male and female workers is unexplained even after controlling for occupations, industries, workers' and jobs' characteristics.

We use the occupational decomposition technique to further examine the wage gap in the specific context of the occupations that individuals choose. Usually occupational dummy variables are included in earnings regressions to account for wage differences across occupations. But this method considers the possibility that pre-labor market discrimination might result in occupations being chosen in a discriminatory manner such that if we account for this difference in occupational structure between different groups, we might be able to better explain the continued existence of a wage gap. The occupation-based decomposition analyses show that the differences in occupational attainment are not statistically significant, and wage discrimination is the primary cause of wage gaps for female and male workers, while most of the wage gap is explained for low and high caste workers. We conclude that within-occupation discrimination in India is particularly important for women and not so important for lower castes, while discrimination in access to occupations is not significant for either group. Part of this may be explained by the affirmative action programs, some of which include providing subsidized schooling and higher education, and coaching to prepare for entrance exams, besides reservations in government sector jobs for lower castes. These wage decomposition results indicate that for those who do manage to get the education, there are clearly fewer hurdles in accessing jobs. Since a major part of the wage gap is still *explained differences* in productive characteristics, improving access to education for minorities would be a step in the right direction.

Chapter 3 of the dissertation uses the changes in trade policies observed in India during the 1990's to test the taste-based theory of discrimination due to Becker. Liberalization of the economy (reduction in tariffs and non-tariff trade barriers and deregulation) increased the degree of competition thereby reducing the scope for employer discrimination due to lower rents. We use the reduction in import tariffs and non-tariff barriers implemented in India as an exogenous shock to the degree of competition in the Indian economy. We evaluate in particular the impact of the change in the degree of competition on wage discrimination against women and lower caste workers.

Consistent with Becker's theory of discrimination and with the results from Chapter 2, we find that industries that experience more liberalization experience a proportionately higher reduction in the wage differential between women and men. The impact on wages of caste groups was however not statistically significant. This result is relevant both as a contribution to the empirical literature on the economics of discrimination and as a policy analysis of the economic impacts of trade reforms in India. Previous empirical evidence on the impact of trade liberalization on gender discrimination has been mixed. The result from our analyses is evidence in favor of Becker's theory. Welfare effects of trade liberalization in India are still the subject of debate, so an important input into the discussion is whether increased openness was associated with reduced discrimination or not. The analysis in this dissertation provides valuable support in evidence of a beneficial effect in terms of lowered male-female wage gaps as a result of increased competition through trade reforms.

Tables & Figures

Table 1.1 Work Force Participation Rates for Social Groups in Urban India

Year	Scheduled Castes		Others (not including STs)	
	Male (1)	Female (2)	Male (3)	Female (4)
1983	490	205	514	139
1987-88	492	213	509	140
1993-94	505	199	523	145
1999-00	503	185	522	128

Source: *Sarvekshana, 87th Volume, Government of India.*

Notes: Numbers are for all workers (reported using the 'usual' status). They represent number of persons employed per 1000 persons in India as a whole. 'Usual status' includes persons who had, for a relatively longer period of the year, either worked or were looking for work, and also those from among the remaining population who had at least for some time with some regularity.

ST is scheduled tribes, i.e. backward tribes who were listed in the Constitution and provided with reservations (7.5%) in government jobs. For 1999-00, 'Others' includes Other Backward Classes (i.e. groups other than Scheduled Castes who were included in another list eligible for reservations in 1989).

Table 1.2 Work Force Participation Rates for Males and Females in Urban India

Year	Category of Work	Male	Female
1983	Principal status	500	120
1987-88	Principal status	496	118
1993-94	Principal status	513	121
1999-00	Principal status	513	117

Source: Sarvekshana, 87th Volume, GOI. [Nos. in %]

Notes: Numbers are Number of persons employed per 1000 persons according to the usual status for all of India. 'Usual status' includes persons who had, for a relatively longer period of the year, either worked or were looking for work, and also those from among the remaining population who had at least for some time with some regularity.

Table 2.1 Summary Statistics for Scheduled and Non-Scheduled Caste Workers

Variables	Non-Scheduled Castes*	Scheduled Castes	Non-Scheduled Castes	Scheduled Castes	Non-Scheduled Castes	Scheduled Castes	Non-Scheduled Castes	Scheduled Castes
	1983-84		1987-88		1993-94		1999-2000	
Real Weekly Earnings (Rs.)	247.86	179.04	394.99	270.24	318.24	229.90	401.95	308.55
Age	35.0	34.3	35.7	34.5	36.1	34.7	36.6	35.3
Proportion (in %) of each group having:								
No education‡	21.1	49.2	17.7	48.2	16.4	43.6	14.3	36.9
Below Primary	10.3	14.0	10.1	14.2	9.2	13.1	7.5	12.2
Primary Education	16.2	15.6	15.8	15.8	12.3	14.3	10.2	13.6
Middle School Education	17.2	11.8	14.8	11.7	15.9	13.6	16.8	16.6
Secondary Education	22.3	7.4	24.3	7.6	26.2	11.5	29.3	15.2
Graduate	12.9	2.1	17.4	2.6	20.1	4.0	22.1	5.6
Proportion (in %) of each group in following industries:								
Agriculture	11.6	14.1	9.4	14.9	10.3	16.3	7.8	11.8
Manufacturing	25.2	23.5	24.9	22.6	22.8	18.4	12.6	9.7
Services**	24.4	28.9	25.3	27.8	26.4	28.3	4.9	9.1
Proportion (in %) of each group in following occupations:								
Professionals	6.4	1.5	7.3	2.0	7.9	2.1	7.8	2.8
Service***	9.5	19.7	8.9	18.6	8.2	15.8	8.7	16.3
Production	13.6	25.0	12.6	25.3	12.9	28.0	14.3	28.0
N	39,621	6,425	41,971	6,560	42,459	6,782	41,711	8,070

**Non-scheduled castes and tribes are neither scheduled castes nor scheduled tribes. These figures are for urban areas in 16 states.*
***Community, Social and Personal Services. *** This includes hotel & restaurant workers, maids etc.*

Table 2.2 Summary Statistics for Male and Female Workers

Variables	Males	Females	Males	Females	Males	Females	Males	Females
	1983-84		1987-88		1993-94		1999-2000	
Real Weekly Earnings (Rs.)	257.89	145.44	398.01	288.43	327.64	217.59	406.28	300.53
Age	35.0	34.9	35.5	35.4	35.9	35.9	36.4	36.4
Proportion (in %) of each group having:								
No education‡	18.4	56.2	15.8	48.8	14.5	43.3	13.1	39.2
Below Primary	11.5	7.9	11.1	8.4	10.0	8.4	8.3	8.2
Primary Education	17.5	9.9	17.1	9.7	13.1	10.0	11.1	9.1
Middle School Education	18.6	6.0	16.1	6.3	17.4	8.2	18.2	10.4
Secondary Education	22.1	11.3	24.1	12.7	26.6	14.0	29.5	15.8
Graduate	12.0	8.6	15.7	14.1	18.4	16.1	19.9	17.3
Proportion (in %) of each group in following industries:								
Agriculture	9.4	23.8	8.0	20.0	8.6	21.4	6.7	16.4
Manufacturing	25.3	23.2	24.8	23.4	22.5	20.5	13.3	6.9
Services*	22.9	34.8	23.1	37.1	23.8	38.1	4.0	12.7
Proportion (in %) of each group in following occupations:								
Professionals	4.7	10.5	5.1	13.2	5.5	13.6	5.4	14.1
Service**	9.2	18.8	8.5	18.0	7.9	15.0	8.3	17.3
Production	16.1	10.6	15.3	9.8	16.2	10.2	18.0	9.9
N	38,007	8,039	39,722	8,809	39,594	9,647	40,613	9,168

*These figures are for urban areas in 16 states. *Community, Social and Personal Services. ** This includes hotel & restaurant workers, maids etc.*

Table 2.3 Wage Regressions for the Entire Sample

Dependent Variable: Natural log of real weekly earnings

Panel A: Without education x caste and education x male

	1983-84	1987	1993-94	1999-2000
Below primary	0.159 (.014)***	0.181 (.016)***	0.164 (.020)***	0.175 (.016)***
Primary	0.245 (.013)***	0.280 (.015)***	0.218 (.019)***	0.241 (.015)***
Middle	0.378 (.013)***	0.378 (.016)***	0.362 (.018)***	0.326 (.014)***
Secondary	0.708 (.014)***	0.712 (.015)***	0.635 (.018)***	0.593 (.014)***
Graduate	0.994 (.017)***	1.084 (.018)***	0.947 (.021)***	0.965 (.017)***
Marital status	0.160 (.010)***	0.150 (.011)***	0.154 (.014)***	0.131 (.011)***
Household Size	-0.008 (.001)***	-0.017 (.002)***	-0.016 (.002)***	-0.012 (.002)***
Caste	-0.037 (.010)***	-0.037 (.012)***	-0.000 (.014)	-0.031 (.010)***
Male	0.529 (.011)***	0.464 (.012)***	0.473 (.013)***	0.358 (.011)***
Adjusted R^2	0.5704	0.5437	0.4324	0.5828
N	27,953	29,160	28,610	28,376

Panel B: Including education x caste, education x male

	1983-84	1987	1993-94	1999-2000
Below primary	0.155 (.035)***	0.204 (.039)***	0.170 (.045)***	0.194 (.037)***
Primary	0.150 (.031)***	0.311 (.037)***	0.136 (.044)***	0.285 (.037)***
Middle	0.499 (.040)***	0.415 (.045)***	0.399 (.048)***	0.335 (.036)***
Secondary	1.206 (.028)***	1.071 (.030)***	1.044 (.035)***	0.863 (.029)***
Graduate	1.457 (.031)***	1.521 (.030)***	1.363 (.034)***	1.245 (.028)***
Marital status	0.158 (.010)***	0.144 (.011)***	0.146 (.014)***	0.124 (.011)***
Household Size	-0.008 (.001)***	-0.016 (.002)***	-0.016 (.002)***	-0.013 (.002)***
Caste	0.023 (.016)	0.046 (.019)**	0.101 (.023)***	0.036 (.018)**
Male	0.693 (.016)***	0.658 (.019)***	0.664 (.023)***	0.498 (.019)***
Adjusted R^2	0.5799	0.5500	0.4394	0.5858
N	27,953	29,160	28,610	28,376

*Note: Standard errors are reported in parentheses. * Significant at 10%, ** Significant at 5%, *** Significant at 1%. Each regression also includes cubic and quartic terms in age, state, industry and occupation fixed effects; and 3 dummy variables which indicate the quarter of the year in which individual was surveyed.*

Table 2.4 Wage Regressions for the Entire Sample: With and Without Occupation Fixed Effects

Dependent Variable: Natural log of real weekly earnings

Panel A: Without occupation in the regression

	1983-84	1987	1993-94	1999-2000
Below primary	-0.104 (.040)***	-0.077 (.047)*	-0.176 (.058)***	-0.013 (.050)
Primary	-0.098 (.036)***	-0.103 (.042)**	-0.052 (.054)	0.004 (.046)
Middle	-0.147 (.038)***	-0.205 (.045)***	-0.217 (.053)***	-0.177 (.042)***
Secondary	0.168 (.037)***	0.087 (.042)**	-0.055 (.050)	-0.152 (.040)***
Graduate	0.471 (.047)***	0.583 (.049)***	0.499 (.058)***	0.313 (.046)***
Marital status	0.117 (.010)***	0.099 (.012)***	0.105 (.014)***	0.085 (.011)***
Household Size	-0.009 (.001)***	-0.018 (.002)***	-0.016 (.002)***	-0.015 (.002)***
Caste	-0.036 (.010)***	-0.039 (.012)***	-0.005 (.014)	-0.029 (.010)***
Male	0.506 (.011)***	0.449 (.011)***	0.454 (.013)***	0.085 (.011)***
Adjusted R^2	0.5654	0.5365	0.4276	0.5736
N	27,953	29,160	28,610	28,376

Panel B: Including Occupation in the regression

Below primary	-0.108 (.040)***	-0.084 (.046)*	-0.179 (.057)***	-0.004 (.049)
Primary	-0.095 (.035)***	-0.110 (.041)***	-0.051 (.054)	0.002 (.045)
Middle	-0.134 (.037)***	-0.202 (.044)***	-0.208 (.053)***	-0.168 (.041)
Secondary	0.120 (.037)***	0.055 (.042)	-0.067 (.050)	-0.174 (.039)***
Graduate	0.404 (.048)***	0.493 (.049)***	0.388 (.058)***	0.197 (.046)***
Marital status	0.117 (.010)***	0.102 (.012)***	0.107 (.014)***	0.086 (.011)***
Household Size	-0.009 (.001)***	-0.016 (.002)***	-0.016 (.002)***	-0.012 (.002)***
Caste	-0.030 (.010)***	-0.030 (.012)**	0.006 (.014)	-0.024 (.010)**
Male	0.519 (.011)***	0.453 (.012)***	0.463 (.013)***	0.347 (.011)***
Adjusted R^2	0.5765	0.5494	0.4376	0.5914
N	27,953	29,160	28,610	28,376

*Note: The dependent variable in each regression is the natural log of real weekly earnings. Standard errors are reported in parentheses. * Significant at 10%, ** Significant at 5%, *** Significant at 1%. Each regression also includes education variables x age interactions; cubic and quartic terms in age; state and industry fixed effects; and 3 dummy variables which indicate the quarter of the year in which individual was surveyed.*

Table 2.5 Wage Regressions for Male and Female Workers Separately- Including Education x Age

Dependent Variable: Natural log of real weekly earnings

	1983-84	1987	1993-94	1999-2000
Panel A: Male Wage Earners				
Below primary	-0.107 (.042)**	-0.046 (.051)	-0.233 (.063)***	0.007 (.052)
Primary	-0.068 (.037)*	-0.090 (.045)**	-0.066 (.059)	0.016 (.048)
Middle	-0.129 (.039)***	-0.157 (.048)***	-0.222 (.058)***	-0.152 (.044)***
Secondary	0.036 (.039)	0.042 (.046)	-0.142 (.056)**	-0.133 (.043)***
Graduate	0.284 (.051)***	0.391 (.054)***	0.238 (.067)***	0.218 (.051)***
Caste	-0.051 (.011)***	-0.072 (.013)***	-0.035 (.015)**	-0.047 (.011)***
Marital status	0.133 (.012)***	0.111 (.014)***	0.150 (.017)***	0.110 (.013)***
Household Size	-0.008 (.001)***	-0.015 (.002)***	-0.015 (.002)***	-0.012 (.002)***
Adjusted R^2	0.5328	0.5120	0.3909	0.5704
N	23,001	23,757	22,974	23,042
Panel B: Female Wage Earners				
Below primary	0.081 (.114)	-0.071 (.124)	0.001 (.141)	0.080 (.0132)
Primary	0.005 (.104)	-0.044 (.118)	0.039 (.142)	0.191 (.131)
Middle	-0.104 (.138)	-0.502 (.143)***	-0.259 (.158)	-0.034 (.126)
Secondary	0.039 (.108)	-0.116 (.115)	-0.105 (.124)	-0.476 (.102)***
Graduate	0.348 (.130)***	0.429 (.121)***	0.378 (.129)***	-0.171 (.111)
Caste	0.062 (.026)**	0.132 (.029)***	0.146 (.032)***	0.055 (.025)**
Marital status	-0.023 (.023)	0.013 (.024)	-0.042 (.028)	-0.013 (.023)
Household Size	-0.001 (.004)	-0.012 (.005)**	-0.009 (.006)	-0.012 (.005)***
Adjusted R^2	0.5707	0.5736	0.4873	0.6067
N	4,952	5,403	5,636	5,334

*Note: The dependent variable in each regression is the natural log of real weekly earnings. Standard errors are reported in parentheses. * Significant at 10%, ** Significant at 5%, *** Significant at 1%. Each regression also includes cubic and quartic terms in age, education x age; state, industry and occupation fixed effects; and 3 dummy variables which indicate the quarter of the year in which individual was surveyed.*

Table 2.6 Wage Regressions for Low and High Caste Workers Separately – Including Education x Age

Dependent Variable: Natural log of real weekly earnings

	1983-84	1987	1993-94	1999-2000
Panel A: Low Caste Wage Earners				
Below primary	-0.219 (.091)**	0.039 (.100)	-0.314 (.116)***	0.011 (.090)
Primary	-0.189 (.088)**	-0.304 (.098)***	-0.135 (.111)	0.111 (.085)
Middle	-0.202 (.101)**	-0.239 (.112)**	-0.497 (.119)***	-0.124 (.084)
Secondary	0.142 (.136)	0.091 (.146)	-0.051 (.133)	-0.217 (.089)**
Graduate	0.289 (.287)	0.636 (.277)**	0.226 (.250)	-0.044 (.156)
Male	0.607 (.026)***	0.485 (.029)***	0.513 (.031)***	0.363 (.024)***
Marital status	0.185 (.028)***	0.121 (.031)***	0.079 (.034)**	0.037 (.025)
Household Size	0.004 (.004)	-0.007 (.004)	-0.002 (.005)	-0.012 (.004)***
Adjusted R^2	0.4303	0.4164	0.3584	0.5143
N	4,475	4,432	4,662	5,462
Panel B: High Caste Wage Earners				
Below primary	-0.083 (.045)*	-0.119 (.053)**	-0.139 (.066)**	-0.026 (.058)
Primary	-0.084 (.039)**	-0.073 (.047)	-0.025 (.063)	-0.053 (.053)
Middle	-0.126 (.041)***	-0.188 (.050)***	-0.149 (.060)**	-0.206 (.048)***
Secondary	0.115 (.040)***	0.062 (.047)	-0.051 (.057)	-0.205 (.046)***
Graduate	0.407 (.050)***	0.498 (.053)***	0.403 (.064)***	0.167 (.052)***
Male	0.497 (.012)***	0.446 (.013)***	0.449 (.015)***	0.339 (.012)***
Marital status	0.104 (.011)***	0.097 (.013)***	0.111 (.016)***	0.101 (.012)***
Household Size	-0.012 (.001)***	-0.018 (.002)***	-0.019 (.002)***	-0.014 (.002)***
Adjusted R^2	0.5887	0.5527	0.4360	0.5948
N	23,478	24,728	23,948	22,914

*Note: The dependent variable in each regression is the natural log of real weekly earnings. Standard errors are reported in parentheses. * Significant at 10%, ** Significant at 5%, *** Significant at 1%. Each regression also includes cubic and quartic terms in age, education x age; state, industry and occupation fixed effects; and 3 dummy variables which indicate the quarter of the year in which individual was surveyed.*

Table 2.7 Wage Regressions for Male and Female Workers Separately – Including Education x Age and Education x Caste

Dependent Variable: Natural log of real weekly earnings

	1983-84	1987	1993-94	1999-2000
Panel A: Male Wage Earners				
Below primary	-0.088 (.043)**	-0.036 (.052)	-202 (.065)***	0.021 (.054)
Primary	-0.066 (.038)*	-0.069 (.046)	-0.025 (.061)	0.023 (.049)
Middle	-0.111 (.039)***	-0.143 (.049)***	-0.174 (.059)***	-0.133 (.045)***
Secondary	0.055 (.040)	0.064 (.047)	-0.094 (.057)*	-0.107 (.044)**
Graduate	0.300 (.051)***	0.405 (.055)***	0.278 (.067)***	0.240 (.052)***
Caste	-0.013 (.019)	-0.032 (.023)	0.068 (.028)**	-0.003 (.022)
Marital status	0.134 (.012)***	0.112 (.014)***	0.151 (.017)***	0.111 (.013)***
Household Size	-0.008 (.001)***	-0.015 (.002)***	-0.015 (.002)***	-0.011 (.002)***
Adjusted R^2	0.5330	0.5121	0.3914	0.5706
N	23,001	23,757	22,974	23,042
Panel B: Female Wage Earners				
Below primary	0.125 (.114)	-0.059 (.125)	-0.022 (.144)	0.147 (.135)
Primary	0.030 (.106)	-0.001 (.121)	0.030 (.145)	0.216 (.135)
Middle	-0.056 (.141)	-0.474 (.149)***	-0.242 (.162)	-0.025 (.129)
Secondary	0.058 (.109)	-0.087 (.115)	-0.109 (.126)	-0.479 (.103)***
Graduate	0.368 (.130)***	0.430 (.121)***	0.384 (.129)***	-0.155 (.111)
Caste	0.106 (.030)***	0.160 (.034)***	0.144 (.039)***	0.079 (.034)**
Marital status	-0.023 (.023)	0.013 (.024)	-0.042 (.028)	-0.012 (.023)
Household Size	-0.002 (.004)	-0.012 (.005)**	-0.009 (.006)	-0.012 (.005)***
Adjusted R^2	0.5714	0.5739	0.4871	0.6070
N	4,952	5,403	5,636	5,334

*Note: The dependent variable in each regression is the natural log of real weekly earnings. Standard errors are reported in parentheses. * Significant at 10%, ** Significant at 5%, *** Significant at 1%. Each regression also includes cubic and quartic terms in age, education x age, education x caste; state, industry and occupation fixed effects; and 3 dummy variables which indicate the quarter of the year in which individual was surveyed.*

Table 2.8 Wage Regressions for Low and High Caste Workers Separately – Including Education x Age and Education x Male

Dependent Variable: Natural log of real weekly earnings

	1983-84	1987	1993-94	1999-2000
Panel A: Low Caste Wage Earners				
Below primary	-0.373 (.122)***	0.031 (.127)	-0.211 (.138)	-0.102 (.110)
Primary	-0.254 (.116)**	-0.296 (.126)**	-0.133 (.137)	0.125 (.106)
Middle	-0.142 (.147)	-0.192 (.147)	-0.424 (.149)***	-0.089 (.106)
Secondary	0.577 (.163)***	0.345 (.179)*	0.416 (.160)***	0.149 (.108)
Graduate	0.476 (.030)	1.309 (.318)***	0.569 (.276)**	0.232 (.169)
Male	0.621 (.030)***	0.518 (.034)***	0.591 (.037)***	0.424 (.031)***
Marital status	0.183 (.027)***	0.122 (.031)***	0.072 (.033)**	0.036 (.025)
Household Size	0.004 (.004)	-0.006 (.004)	-0.001 (.005)	-0.012 (.004)***
Adjusted R^2	0.4335	0.4187	0.3623	0.5182
N	4,475	4,432	4,662	5,462
Panel B: High Caste Wage Earners				
Below primary	-0.078 (.055)	-0.081 (.063)	-0.179 (.079)**	0.005 (.069)
Primary	-0.165 (.048)***	-0.040 (.058)	-0.146 (.075)*	-0.022 (.065)
Middle	-0.007 (.056)	-0.141 (.066)**	-0.138 (.076)*	-0.216 (.060)***
Secondary	0.582 (.046)***	0.432 (.053)***	0.321 (.065)***	0.047 (.052)
Graduate	0.818 (.054)***	0.900 (.057)***	0.750 (.069)***	0.422 (.056)***
Male	0.721 (.018)***	0.711 (.022)***	0.685 (.028)***	0.533 (.024)***
Marital status	0.096 (.011)***	0.088 (.012)***	0.100 (.015)***	0.090 (.012)***
Household Size	-0.011 (.001)***	-0.018 (.002)***	-0.018 (.002)***	-0.014 (.002)***
Adjusted R^2	0.6000	0.5603	0.4436	0.5984
N	23,478	24,728	23,948	22,914

*Note: The dependent variable in each regression is the natural log of real weekly earnings. Standard errors are reported in parentheses. * Significant at 10%, ** Significant at 5%, *** Significant at 1%. Each regression also includes cubic and quartic terms in age, education x age, education x male; state, industry and occupation fixed effects; and 3 dummy variables which indicate the quarter of the year in which individual was surveyed.*

Table 2.9 Blinder-Oaxaca Decomposition for Gender Groups

	Log Wage Gap	Std Error of Wage Gap	Part of the Gap Explained By			Total Unexplained 'Discrimination'	Std Error
			Total	EDU	OCC		
Panel A: Based on Males as the 'No-Discrimination' Model (not weighted)							
1983	0.825	0.016***	0.308	0.151	-0.010	0.517	0.021***
1987	0.685	0.017***	0.240	0.123	0.016	0.445	0.021***
1993	0.654	0.017***	0.174	0.088	0.013	0.480	0.023***
1999	0.550	0.017***	0.250	0.081	-0.040	0.300	0.019***
Panel B: Based on Weights from the Pooled Model							
1983	0.824	0.016***	0.413	0.147	0.038	0.364	0.010***
1987	0.685	0.017***	0.321	0.148	0.021	0.390	0.011***
1993	0.654	0.017***	0.280	0.125	0.035	0.374	0.012***
1999	0.550	0.017***	0.279	0.097	-0.022	0.271	0.009***

*Note: The model includes 5 education dummy variables; 4th order polynomial in age; state, industry and occupation, period surveyed fixed effects. EDU is education & OCC is occupation The last column represents standard error of the unexplained component. ***Significant at 1%.*

Table 2.10 Blinder-Oaxaca Decomposition for Caste Groups

	Log Wage Gap	Std Error of Wage Gap	Part of the Gap Explained By			Total Unexplained 'Discrimination'	Std Error
			Total	EDU	OCC		
Panel A: Based on Higher Castes as the 'No-Discrimination' Model (not weighted)							
1983	0.427	0.014***	0.402	0.263	0.060	0.025	0.011**
1987	0.532	0.015***	0.513	0.322	0.092	0.019	0.013
1993	0.458	0.016***	0.471	0.277	0.095	-0.013	0.014
1999	0.436	0.014***	0.416	0.249	0.076	0.021	0.011**
Panel B: Based on Weights from the Pooled Model							
1983	0.427	0.014***	0.394	0.255	0.060	0.033	0.010***
1987	0.532	0.015***	0.500	0.313	0.091	0.032	0.011***
1993	0.458	0.016***	0.458	0.263	0.097	0.000	0.012
1999	0.436	0.014***	0.409	0.238	0.078	0.027	0.009***

*Note: The model includes 5 education, gender, marital status & household size dummy variables; 4th order polynomial in age; state, industry and occupation, period surveyed fixed effects. EDU is education & OCC is occupation The last column represents standard error of the unexplained component. ***Significant at 1%.*

Table 2.11 Reference for Occupation-Based Decomposition Tables

1. Columns (1) and (2) show the actual proportion of Males (P_m), females (P_f), higher caste (P_n) and lower caste (P_{sc}) workers in the sample, while column (4) shows the predicted probability ($\hat{P}_m$) of female and lower caste ($\hat{P}_{sc}$) workers being in different occupations assuming they made occupational choices as if they were males and non-scheduled castes, respectively.

2. Column (5) thus is the observed differences in occupational difference, and is decomposed into unexplained and explained parts in columns (6) and (7).

3. Column (8) is the difference in average log wages of non-sc and sc workers in each occupational group, and this is decomposed into an explained part (column 9) and an unexplained part (column 10).

4. Columns (11)-(14) is the further decomposition of the earnings differential into explained part of wage differential (column 11), unexplained wage differential (12), difference due to job differences (col. 13), and job discrimination (col. 14).

Table 2.12 Occupation-based Decomposition of Wage Gap for Gender Groups in 1983

	Observed occupational distribution		Predicted occupational distribution		Observed Difference	Explained difference	Residual Difference
	P_m	P_f	$\hat{P}_m$	$\hat{P}_f$	$P_m - P_f$	$P_m - \hat{P}_f$	$\hat{P}_f - P_f$
Occupation	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
Skilled	0.146	0.033	0.146	0.142	0.113	0.004	0.109
Professional	0.143	0.227	0.143	0.107	-0.084	0.035	-0.119
Clerical	0.224	0.121	0.224	0.172	0.103	0.052	0.051
Service	0.098	0.233	0.098	0.110	-0.135	-0.012	-0.122
Unskilled	0.276	0.253	0.276	0.336	0.023	-0.059	0.082
Production	0.112	0.133	0.112	0.133	-0.021	-0.020	-0.001

	G= $\ln Y_n - \ln Y_{sc}$	E= $f_m(\bar{x}_m - \bar{x}_f)$	D= $\bar{x}_f(f_m - f_f)$	$P_f \times E$	$P_f \times D$	$P_m - \hat{P}_f$ x $\ln Y_n$	$(\hat{P}_f - P_f)$ x $\ln Y_n$
	(8)	**(9)**	**(10)**	**(11)**	**(12)**	**(13)**	**(14)**
Skilled	1.300	0.372	0.928	0.012	0.031	0.024	0.622
Professional	0.543	0.175	0.368	0.040	0.083	0.230	-0.778
Clerical	0.094	0.030	0.064	0.004	0.008	0.320	0.314
Service	1.039	0.287	0.753	0.067	0.175	-0.070	-0.688
Unskilled	1.038	0.204	0.834	0.052	0.211	-0.317	0.442
Production	1.212	0.315	0.897	0.042	0.119	-0.112	-0.003
Total				0.216 (.024)[a]	0.628 (.080)[a]	0.074 (.234)	-0.092 (.593)
				WE	**WD**	**JE**	**JD**

Notes: m, f refer to males and females respectively. Numbers in parentheses are standard errors.
[a] Significant at 1%. Refer to Table 2.11 for notes on interpreting numbers in each column.

Table 2.13 Occupation-based Decomposition of Wage Gap for Gender Groups in 1987

	Observed occupational distribution		Predicted occupational distribution		Observed Difference	Explained difference	Residual Difference
	P_m	P_f	$\hat{P}_m$	$\hat{P}_f$	$P_m - P_f$	$P_m - \hat{P}_f$	$\hat{P}_f - P_f$
Occupation	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
Skilled	0.143	0.043	0.143	0.135	0.099	0.007	0.092
Professional	0.169	0.300	0.169	0.154	-0.131	0.015	-0.146
Clerical	0.234	0.142	0.234	0.199	0.093	0.035	0.058
Service	0.098	0.183	0.098	0.105	-0.086	-0.007	-0.079
Unskilled	0.262	0.227	0.262	0.302	0.034	-0.041	0.075
Production	0.095	0.105	0.095	0.104	-0.010	-0.009	0.000

	G= $\ln Y_n - \ln Y_{sc}$	E= $f_m(\bar{x}_m - \bar{x}_f)$	D= $\bar{x}_f(f_m - f_f)$	$P_f \times E$	$P_f \times D$	$P_m - \hat{P}_f$ x $\ln Y_n$	$(\hat{P}_f - P_f)$ x $\ln Y_n$
	(8)	**(9)**	**(10)**	**(11)**	**(12)**	**(13)**	**(14)**
Skilled	1.321	0.455	0.866	0.020	0.037	0.043	0.539
Professional	0.491	0.114	0.377	0.034	0.113	0.101	-0.998
Clerical	0.021	0.015	0.006	0.002	0.001	0.221	0.366
Service	1.059	0.273	0.785	0.050	0.144	-0.040	-0.455
Unskilled	0.996	0.258	0.737	0.059	0.168	-0.225	0.414
Production	1.159	0.436	0.723	0.046	0.076	-0.054	-0.001
Total				0.210 (.021)[a]	0.538 (.064)[a]	0.046 (.152)	-0.135 (.599)
				WE	WD	JE	JD

Notes: m, f refer to males and females respectively. Numbers in parentheses are standard errors.
[a] Significant at 1%. Refer to Table 2.11 for notes on interpreting numbers in each column.

Table 2.14 Occupation-based Decomposition of Wage Gap for Gender Groups in 1993

	Observed occupational distribution		Predicted occupational distribution		Observed Difference	Explained difference	Residual Difference
	P_m	P_f	$\hat{P}_m$	$\hat{P}_f$	$P_m - P_f$	$P_m - \hat{P}_f$	$\hat{P}_f - P_f$
Occupation	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
Skilled	0.150	0.032	0.150	0.131	0.117	0.019	0.099
Professional	0.177	0.304	0.177	0.174	-0.127	0.003	-0.130
Clerical	0.201	0.158	0.201	0.182	0.043	0.019	0.024
Service	0.091	0.172	0.091	0.093	-0.082	-0.002	-0.080
Unskilled	0.290	0.228	0.290	0.328	0.062	-0.038	0.101
Production	0.092	0.106	0.092	0.092	-0.014	-0.001	-0.014

	G= $\ln Y_n - \ln Y_{sc}$	E= $f_m(\bar{x}_m - \bar{x}_f)$	D= $\bar{x}_f(f_m - f_f)$	$P_f \times E$	$P_f \times D$	$P_m - \hat{P}_f$ x $\ln Y_n$	$(\hat{P}_f - P_f)$ x $\ln Y_n$
	(8)	**(9)**	**(10)**	**(11)**	**(12)**	**(13)**	**(14)**
Skilled	1.033	0.347	0.686	0.011	0.022	0.110	0.586
Professional	0.478	0.147	0.331	0.045	0.101	0.020	-0.878
Clerical	0.098	0.030	0.068	0.005	0.011	0.123	0.152
Service	1.098	0.365	0.732	0.063	0.126	-0.011	-0.464
Unskilled	0.894	0.206	0.688	0.047	0.157	-0.213	0.560
Production	1.174	0.281	0.893	0.030	0.095	-0.004	-0.078
Total				0.200 (.022)[a]	0.511 (.058)[a]	0.025 (.121)	-0.122 (.578)
				WE	WD	JE	JD

Notes: m, f refer to males and females respectively. Numbers in parentheses are standard errors.
[a] Significant at 1%. Refer to Table 2.11 for notes on interpreting numbers in each column.

Table 2.15 Occupation-based Decomposition of Wage Gap for Gender Groups in 1999

	Observed occupational distribution		Predicted occupational distribution		Observed Difference	Explained difference	Residual Difference
	P_m	P_f	$\hat{P}_m$	$\hat{P}_f$	$P_m - P_f$	$P_m - \hat{P}_f$	$\hat{P}_f - P_f$
Occupation	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
Skilled	0.155	0.030	0.155	0.134	0.125	0.020	0.105
Professional	0.170	0.330	0.170	0.170	-0.159	0.001	-0.160
Clerical	0.186	0.154	0.186	0.169	0.032	0.017	0.015
Service	0.098	0.180	0.098	0.104	-0.082	-0.007	-0.075
Unskilled	0.314	0.220	0.314	0.345	0.093	-0.032	0.125
Production	0.078	0.087	0.078	0.077	-0.009	0.001	-0.010

	G= $\ln Y_n - \ln Y_{sc}$	**E=** $f_m(\bar{x}_m - \bar{x}_f)$	**D=** $\bar{x}_f(f_m - f_f)$	$P_f \times E$	$P_f \times D$	$P_m - \hat{P}_f$ **x $\ln Y_n$**	$(\hat{P}_f - P_f)$ **x $\ln Y_n$**
	(8)	**(9)**	**(10)**	**(11)**	**(12)**	**(13)**	**(14)**
Skilled	0.721	0.291	0.430	0.009	0.013	0.124	0.639
Professional	0.539	0.161	0.378	0.053	0.124	0.004	-1.129
Clerical	0.091	0.071	0.020	0.011	0.003	0.113	0.100
Service	0.974	0.393	0.581	0.071	0.104	-0.040	-0.456
Unskilled	0.797	0.208	0.589	0.046	0.130	-0.184	0.722
Production	0.958	0.202	0.755	0.018	0.066	0.005	-0.059
Total				0.207 (.026)[a]	0.440 (.056)[a]	0.021 (.113)	-0.184 (.696)
				WE	WD	JE	JD

Notes: m, f refer to males and females respectively. Numbers in parentheses are standard errors.
[a] Significant at 1%. Refer to Table 2.11 for notes on interpreting numbers in each column.

Table 2.16 Occupation-based Decomposition of Wage Gap for Caste Groups in 1983

	Observed occupational distribution		Predicted occupational distribution		Observed Difference	Explained difference	Residual Difference
	P_n	P_{sc}	$\hat{P}_n$	$\hat{P}_{sc}$	$P_n - P_{sc}$	$P_n - \hat{P}_{sc}$	$\hat{P}_{sc} - P_{sc}$
Occupation	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
Skilled	0.146	0.114	0.146	0.166	0.033	-0.020	0.053
Professional	0.143	0.032	0.143	0.051	0.110	0.091	0.019
Clerical	0.224	0.131	0.224	0.126	0.094	0.098	-0.005
Service	0.098	0.177	0.098	0.121	-0.080	-0.023	-0.056
Unskilled	0.276	0.397	0.276	0.388	-0.120	-0.112	-0.009
Production	0.112	0.149	0.112	0.147	-0.036	-0.035	-0.002

	G= $\ln Y_n - \ln Y_{sc}$	**E=** $f_n(\bar{x}_n - \bar{x}_{sc})$	**D=** $\bar{x}_{sc}(f_n - f_{sc})$	P_{sc} * **E**	P_{sc} * **D**	$P_n - \hat{P}_{sc}$ **x $\ln Y_n$**	$(\hat{P}_{sc} - P_{sc})$ ***$\ln Y_n$**
	(8)	**(9)**	**(10)**	**(11)**	**(12)**	**(13)**	**(14)**
Skilled	0.324	0.252	0.072	0.029	0.008	-0.115	0.301
Professional	0.439	0.268	0.171	0.009	0.006	0.594	0.124
Clerical	0.215	0.165	0.051	0.022	0.007	0.601	-0.028
Service	0.041	0.038	0.003	0.007	0.000	-0.131	-0.317
Unskilled	0.263	0.118	0.145	0.047	0.058	-0.598	-0.046
Production	0.026	0.002	0.024	0.000	0.004	-0.193	-0.011
Total				0.113 (.017)[a]	0.082 (.022)[a]	0.159 (.477)	0.022 (.205)
				WE	WD	JE	JD

Notes: n =non-scheduled (higher) castes & sc =scheduled (lower) castes. Numbers in parentheses are standard errors.
[a] Significant at 1%. Refer to Table 2.11 for additional notes.

Table 2.17 Occupation-based Decomposition of Wage Gap for Caste Groups in 1987

	Observed occupational distribution		Predicted occupational distribution		Observed Difference	Explained difference	Residual Difference
	P_n	P_{sc}	$\hat{P}_n$	$\hat{P}_{sc}$	$P_n - P_{sc}$	$P_n - \hat{P}_{sc}$	$\hat{P}_{sc} - P_{sc}$
Occupation	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
Skilled	0.143	0.145	0.143	0.163	-0.002	-0.020	0.018
Professional	0.169	0.043	0.169	0.052	0.126	0.117	0.009
Clerical	0.234	0.124	0.234	0.133	0.110	0.101	0.009
Service	0.098	0.170	0.098	0.126	-0.072	-0.028	-0.044
Unskilled	0.262	0.395	0.262	0.392	-0.134	-0.131	-0.003
Production	0.095	0.123	0.095	0.133	-0.028	-0.038	0.010

	G= $\ln Y_n - \ln Y_{sc}$	**E=** $f_n(\bar{x}_n - \bar{x}_{sc})$	**D=** $\bar{x}_{sc}(f_n - f_{sc})$	P_{sc} * **E**	P_{sc} * **D**	$P_n - \hat{P}_{sc}$ **x $\ln Y_n$**	$(\hat{P}_{sc} - P_{sc})$ ***$\ln Y_n$**
	(8)	**(9)**	**(10)**	**(11)**	**(12)**	**(13)**	**(14)**
Skilled	0.347	0.208	0.140	0.030	0.020	-0.120	0.107
Professional	0.566	0.400	0.166	0.017	0.007	0.796	0.063
Clerical	0.204	0.159	0.045	0.020	0.006	0.640	0.060
Service	0.112	0.164	-0.052	0.028	-0.009	-0.163	-0.256
Unskilled	0.301	0.101	0.200	0.040	0.079	-0.722	-0.016
Production	0.014	0.056	-0.042	0.007	-0.005	-0.219	0.059
Total				0.142 (.012)[a]	0.098 (.032)[a]	0.211 (.574)	0.016 (.133)
				WE	WD	JE	JD

Notes: n =non-scheduled (higher) castes & sc =scheduled (lower) castes. Numbers in parentheses are standard errors.
[a] Significant at 1%. Refer to Table 2.11 for additional notes.

Table 2.18 Occupation-based Decomposition of Wage Gap for Caste Groups in 1993

	Observed occupational distribution		Predicted occupational distribution		Observed Difference	Explained difference	Residual Difference
	P_n	P_{sc}	$\hat{P}_n$	$\hat{P}_{sc}$	$P_n - P_{sc}$	$P_n - \hat{P}_{sc}$	$\hat{P}_{sc} - P_{sc}$
Occupation	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
Skilled	0.150	0.122	0.150	0.161	0.028	-0.011	0.039
Professional	0.177	0.047	0.177	0.062	0.130	0.115	0.015
Clerical	0.201	0.109	0.201	0.113	0.092	0.088	0.004
Service	0.091	0.145	0.091	0.115	-0.055	-0.024	-0.031
Unskilled	0.290	0.478	0.290	0.418	-0.188	-0.128	-0.059
Production	0.092	0.099	0.092	0.131	-0.007	-0.039	0.032

	G= $\ln Y_n - \ln Y_{sc}$	**E=** $f_n(\bar{x}_n - \bar{x}_{sc})$	**D=** $\bar{x}_{sc}(f_n - f_{sc})$	P_{sc} * **E**	P_{sc} * **D**	$P_n - \hat{P}_{sc}$ **x $\ln Y_n$**	$(\hat{P}_{sc} - P_{sc})$ ***$\ln Y_n$**
	(8)	**(9)**	**(10)**	**(11)**	**(12)**	**(13)**	**(14)**
Skilled	0.265	0.158	0.107	0.019	0.013	-0.066	0.233
Professional	0.376	0.227	0.149	0.011	0.007	0.773	0.100
Clerical	0.124	0.121	0.004	0.013	0.000	0.558	0.024
Service	0.103	0.137	-0.034	0.020	-0.005	-0.140	-0.180
Unskilled	0.218	0.133	0.085	0.063	0.041	-0.713	-0.330
Production	-0.012	0.036	-0.048	0.004	-0.005	-0.226	0.186
Total				0.130 (.021)[a]	0.051 (.017)[a]	0.187 (.545)	0.032 (.219)
				WE	WD	JE	JD

Notes: n =non-scheduled (higher) castes & sc =scheduled (lower) castes. Numbers in parentheses are standard errors.
[a] Significant at 1%. Refer to Table 2.11 for additional notes.

Table 2.19 Occupation-based Decomposition of Wage Gap for Caste Groups in 1999

	Observed occupational distribution		Predicted occupational distribution		Observed Difference	Explained difference	Residual Difference
	P_n	P_{sc}	$\hat{P}_n$	$\hat{P}_{sc}$	$P_n - P_{sc}$	$P_n - \hat{P}_{sc}$	$\hat{P}_{sc} - P_{sc}$
Occupation	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
Skilled	0.155	0.140	0.155	0.162	0.015	-0.008	0.022
Professional	0.170	0.052	0.170	0.072	0.119	0.098	0.020
Clerical	0.186	0.124	0.186	0.113	0.062	0.073	-0.011
Service	0.098	0.151	0.098	0.112	-0.053	-0.014	-0.039
Unskilled	0.314	0.451	0.314	0.442	-0.138	-0.129	-0.009
Production	0.078	0.082	0.078	0.099	-0.004	-0.021	0.017

	G= $\ln Y_n - \ln Y_{sc}$	**E=** $f_n(\bar{x}_n - \bar{x}_{sc})$	**D=** $\bar{x}_{sc}(f_n - f_{sc})$	P_{sc} * **E**	P_{sc} * **D**	$P_n - \hat{P}_{sc}$ **x $\ln Y_n$**	$(\hat{P}_{sc} - P_{sc})$ ***$\ln Y_n$**
	(8)	**(9)**	**(10)**	**(11)**	**(12)**	**(13)**	**(14)**
Skilled	0.255	0.227	0.028	0.032	0.004	-0.046	0.134
Professional	0.214	0.119	0.096	0.006	0.005	0.695	0.143
Clerical	0.107	0.122	-0.014	0.015	-0.002	0.486	-0.074
Service	0.061	0.151	-0.090	0.023	-0.014	-0.086	-0.235
Unskilled	0.215	0.117	0.099	0.053	0.044	-0.742	-0.052
Production	0.077	0.089	-0.012	0.007	-0.001	-0.125	0.099
Total				0.136 (.018)[a]	0.037 (.020)[b]	0.183 (.508)	0.015 (.150)
				WE	WD	JE	JD

Notes: n =non-scheduled (higher) castes & sc =scheduled (lower) castes. Numbers in parentheses are standard errors.
[a] Significant at 1%, [b] Significant at 5% in a 1-tail t-test. Refer to Table 2.11 for additional notes.

Figure 1: Profits of Discriminating and Non-Discriminating Firms

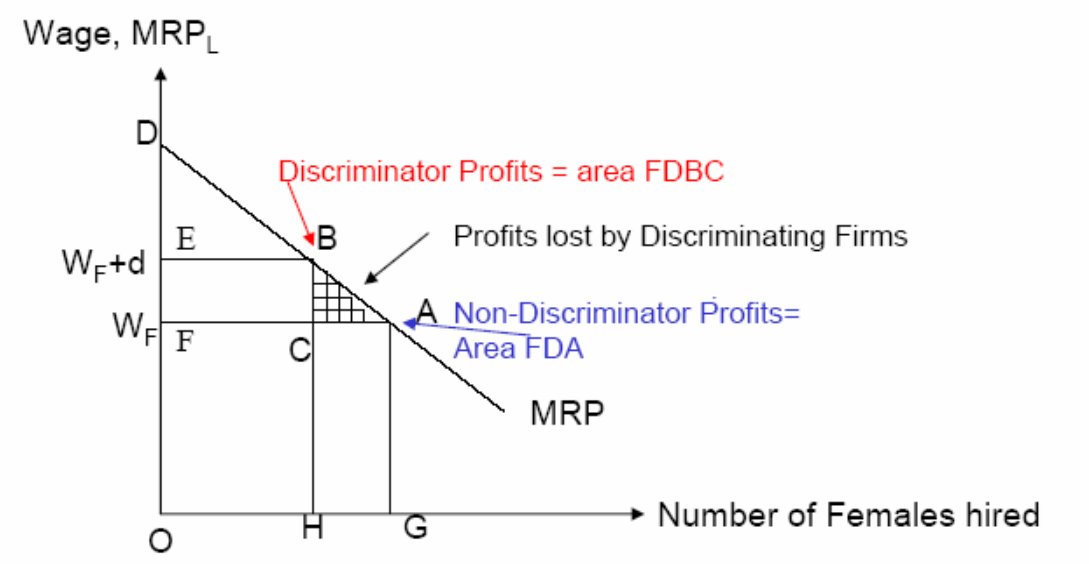


Table 3.1 Summary Statistics by Gender and Caste in 72 Industries

	1983		1999-00		1983		1999-00	
	Female	Male	Female	Male	SC	Non-SC	SC	Non-SC
Real Weekly Earnings (Rs.)	80.1	296.3	133.6	401.2	208.1	274.1	308.5	369.9
Earnings Ratio	0.27		0.33		0.76		0.83	
Age	31.6	33.3	31.6	34.5	32.7	33.1	32.9	34.3
Proportion (%)								
No formal	2.2	2.7	0.9	0.8	2.9	2.6	2.3	0.6
Below Primary	9.6	12.4	10.7	10.2	18.3	11.3	15.4	9.5
Primary	17.8	20.3	17.8	13.8	19.9	20.0	16.1	14.1
Middle	11.8	19.4	17.2	19.6	15.6	18.7	22.5	18.7
Secondary	9.7	20.1	15.2	30.0	9.7	19.8	18.1	29.3
Other Grad	2.2	6.1	5.5	11.2	0.5	6.2	2.8	11.5
Married	56.7	70.6	55.7	71.5	73.7	68.1	69.7	69.1
N	1,043	6,591	888	5,157	840	6,794	794	5,251

The sample is restricted to those employed in 72 three-digit industries.

Table 3.2 Summary statistics of Trade measures

	Phase 1 (1980-85)	Phase 2 (1986-90)	Phase 3 (1991-95)	Phase 4 (1996-00)
Average Effective Rate of Protection	115.1	125.9	80.2	40.4
Average Import Coverage Ratio	97.6	91.6	37.9	24.8
Average Import Penetration Ratio	0.10	0.11	0.12	0.16

Source: Das (2003)

Table 3.3 Individual Level Results for Entire Sample

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Female	**-0.772***** (.083)	**-0.590***** (.124)	**-0.522***** (.110)	**-0.382**** (.147)	**-0.336***** (.122)	-0.163 (.184)	-0.049 (.231)	-0.128 (.269)
Low Caste	**0.087*** (.044)	**0.136***** (.049)	**0.117***** (.044)	0.050 (.056)	-0.088 (.075)	**-0.291**** (.129)	**-0.293***** (.104)	**-0.314***** (.113)
ERP		-0.086*** (.021)	-0.107*** (.026)	-0.152*** (.052)	-0.132*** (.044)	-0.134*** (.047)	-0.139*** (.043)	-0.137*** (.043)
ERP x Female		**-0.197**** (.079)	**-0.180**** (.074)	**-0.234*** (.125)	**-0.282***** (.102)	**-0.389**** (.165)	**-0.441**** (.178)	**-0.392*** (.206)
ERP x Caste		**-0.050***** (.017)	**-0.053***** (.018)	-0.031 (.024)	-0.036 (.062)	0.077 (.113)	0.079 (.103)	0.088 (.105)
- IPR								-0.065 (.121)
-IPR x Female								-0.246 (.241)
-IPR x Caste								-0.120 (.271)
PostLib				-0.042 (.042)	-0.045 (.038)	-0.004 (.028)	-0.005 (.027)	-0.002 (.028)
PostLib*Female				-0.148 (.123)	-0.154 (.109)	-0.534*** (.089)	-0.554*** (.094)	-0.517*** (.109)
Post-Lib * Caste				0.066 (.047)	0.103* (.056)	0.653*** (.072)	0.652*** (.081)	0.660*** (.089)
Industry			Yes	Yes	Yes	Yes	Yes	Yes
Industry x PostLib				Yes	Yes	Yes	Yes	Yes
Industry x Female, Industry x Caste					Yes	Yes	Yes	Yes
Industry x PostLib x Female, Industry x PostLib x Caste						Yes	Yes	Yes
Occupation dummies							Yes	Yes
Occupation * Female Occupation * Caste							Yes	Yes
Demographic characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R squared	0.4989	0.5060	0.5442	0.5538	0.5708	0.5777	0.5864	0.5865
No. of Observations	14844	14844	14844	14844	14844	14844	14844	14844

*Notes: Dependent variable is log of real weekly earnings of each individual. Numbers in parentheses are standard errors. Sample is restricted to Hindus. All regressions include age (quartic), education, married, head, state, female, caste, education*female, education*caste.* * *denotes significance at the 10% level,* ** *denotes significance at the 5% level,* *** *denotes significance at the 1% level. The standard errors are clustered by industry.*

Table 3.4 Individual Level Results for Entire Sample – Additional Specifications with *PostLib*

	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Female	-0.338**	-0.572***	-0.132	-0.559**	-0.689**	-0.677**	-0.606***
	(.165)	(.094)	(.269)	(.242)	(.264)	(.262)	(.187)
Caste	-0.079	-0.010	-0.311***	-0.428**	-.400**	-0.413**	-0.076
	(.088)	(.042)	(.112)	(.183)	(.187)	(.190)	(.088)
ERP	-0.173***	-0.103**	-0.176***	-0.149**	-0.147**	-0.146**	-0.114**
	(.048)	(.045)	(.049)	(.058)	(.058)	(.058)	(.042)
ERP x Female	**-0.267*****	-0.111	**-0.388***	-0.009	-0.003	-0.011	**-0.182****
	(.102)	(.069)	(.206)	(.123)	(.126)	(.128)	(.070)
ERP x Caste	-0.035	**-0.099*****	-0.090	0.187	0.187	0.192	-0.063
	(.057)	(.031)	(.102)	(.139)	(.137)	(.140)	(.050)
Negative IPR	-0.095	-0.291	-0.094	-0.079	-0.273	0.267	-0.249
	(.115)	(.208)	(.136)	(.143)	(.195)	(.195)	(.196)
-IPR x Female	**-0.416*****	**-0.365*****	-0.223	0.192	-1.177	-1.042	**-1.720****
	(.123)	(.118)	(.246)	(.329)	(.894)	(.916)	(.731)
-IPR x Caste	-0.071	0.023	-0.095	-0.082	0.234	0.205	0.194
	(.177)	(.158)	(.280)	(.268)	(.608)	(.599)	(.234)
PostLib	-0.028	0.079	-0.003	0.042	0.067	0.048	0.025
	(.045)	(.080)	(.030)	(.067)	(.076)	(.083)	(.074)
PostLib x Female	-0.177*		-0.529***	0.155	0.292	0.193	0.442***
	(.106)		(.108)	(.245)	(.259)	(.311)	(.182)
PostLib x Caste	0.095*		0.656***	0.795***	0.755***	0.702	0.166*
	(.056)		(.088)	(.187)	(.210)	(.214)	(.089)
ERP x PostLib		-0.130		-0.052	-0.060	-0.032	-0.089
		(.099)		(.080)	(.087)	(.096)	(.090)
ERP x PostLib x Female				**-0.823****	**-0.846****	**-0.715***	**-0.565****
				(.351)	(.348)	(.411)	(.271)
ERP x PostLib x Caste				-0.142	-0.134	-0.046	0.055
				(.211)	(.219)	(.216)	(.113)
-IPR x PostLib		0.305			0.308	0.284	0.378
		(.270)			(.247)	(.257)	(.269)
-IPR x PostLib x Female					1.336	1.126	**1.428***
					(.882)	(.900)	(.790)
-IPR x PostLib x Caste					-0.429		-0.257
					(.667)		(.223)
ICR, ICR x Female, ICR x Caste						Yes	Yes
ICR x Post, ICR x Post x Female, ICR x Post x Caste							Yes
Ind x Post x Female; Ind x Post x Caste			Yes	Yes	Yes	Yes	
Occupation; Occupation x Female; Occupation x Caste	Yes			Yes	Yes	Yes	Yes
R squared	0.5801	0.5717	0.5868	0.5875	0.5876	0.5878	0.5881
No. of Observations	14844	14844	14844	14844	14844	14794	14794

Notes: Dependent variable is log of weekly earnings of each individual. Numbers in parentheses are standard errors. The standard errors are clustered by industry. Sample is restricted to Hindus. All regressions include age (quartic), education, married, head; state, education x female, education x caste, education x ERP, industry, industry x PostLib, industry x female, industry x caste; Concentration measured by Percentage Change in Number of Firms between 1983 and 1999 relative to 1983; and real gross fixed capital formation.
** denotes significance at the 10% level, ** denotes significance at the 5% level, *** denotes significance at the 1% level.*

Table 3.5 Individual Level Results for Gender Groups (Higher Caste Hindus Only)

	(1)	(2)	(3)	(4)	(5)	(6)
ERP		-0.086** (.034)	-0.110*** (.032)	-0.070** (.034)	-0.058* (.032)	-0.162*** (.058)
ERP*Female					**-0.205**** (.086)	**-0.167***** (.083)
Female	-0.577*** (.040)	-0.586*** (.067)	-0.511*** (.057)	-0.515*** (.056)	-0.324*** (.094)	-0.353*** (.088)
Demographic Variables	Yes	Yes	Yes	Yes	Yes	Yes
Industry			Yes	Yes	Yes	Yes
PostLib				Yes	Yes	Yes
Industry x PostLib						Yes
R squared	0.4742	0.4779	0.5284	0.5289	0.5300	0.5396
No. of Observations	12,823	12,823	12,823	12,823	12,823	12,823

Notes: Dependent variable is log of weekly earnings of each individual. Numbers in parentheses are standard errors. Standard errors are clustered by industry. Sample is restricted to higher castes. Demographic variables include age(quartic), education dummies, married, head.
** denotes significance at the 10% level, ** denotes significance at the 5% level, *** denotes significance at the 1% level.*

Table 3.6 Individual Level Results for Gender Groups- Additional Specifications

	(7)	(8)	(9)	(10)	(11)	(12)	(13)
ERP		-0.153*** (.054)	-0.159*** (.050)	-0.164*** (.052)	-0.148*** (.046)	-0.142*** (.044)	-0.146*** (.045)
ERP x Female		-0.130 (.090)	-0.111 (.089)	-0.091 (.080)	**-0.203*** (.104)	-0.231 (.191)	-0.233 (.191)
Negative IPR	-0.194 (.116)	-0.072 (.122)	-0.082 (.123)	-0.084 (.124)	-0.055 (.117)	-0.075 (.127)	-0.092 (.139)
Negative IPR x Female	**-0.549***** (.187)	**-0.436**** (.213)	**-0.430*** (.229)	-0.374 (.291)	**-0.512***** (.154)	**-0.414*** (.233)	-0.396 (.240)
Female	-0.542*** (.047)	-0.407*** (.096)	-0.428*** (.098)	-0.533*** (.103)	-0.347* (.204)	-0.257 (.297)	-0.255 (.297)
Occupation			Yes	Yes	Yes	Yes	Yes
Occupation x Female				Yes	Yes	Yes	Yes
PostLib x Female					-0.113 (.111)	-0.394*** (.115)	-0.394*** (.115)
Industry x Female					Yes	Yes	Yes
Industry x PostLib x Female						Yes	Yes
% Change in No. Firms							Yes
Real Gross Fixed Capital Formation							Yes
R squared	0.5551	0.5571	0.5652	0.5681	0.5799	0.5852	0.5853
No. of Observations	12823	12823	12823	12823	12823	12823	12823

*Notes: Dependent variable is log of real weekly earnings. Numbers in parentheses are standard errors. Each regression includes demographic characteristics; industry, industry x PostLib, PostLib, and state dummies. Standard errors are clustered by industry. * denotes significance at the 10% level, ** denotes significance at the 5% level, *** denotes significance at the 1% level.*

Table 3.7 Individual Level Results for Gender Groups – Additional Specifications with *ERP x PostLib*

	(14)	(15)	(16)	(17)	(18)	(19)#
ERP*Female	**-0.229****	-0.090	-0.092	-0.084	-0.117	-0.126
	(.111)	(.085)	(.084)	(.081)	(.096)	(.095)
-IPR*Female	-0.923	-1.119	-1.112	**-1.229***	**-1.248***	**-1.439***
	(.687)	(.735)	(.736)	(.730)	(.707)	(.739)
Female	-0.522***	-0.713***	-0.710***	-0.703***	-0.589***	-0.626***
	(.132)	(.110)	(.110)	(.108)	(.185)	(.193)
ERP*PostLib*Female	-0.250	**-0.603****	**-0.598****	**-0.584***	**-0.581***	**-0.511***
	(.182)	(.301)	(.301)	(.305)	(.295)	(.292)
-IPR * PostLib* Female	0.645	0.950	0.954	1.058	1.036	-1.088
	(.748)	(.822)	(.821)	(.815)	(.794)	(.798)
Post-liberalization	0.046	0.005	-0.009	0.016	0.015	-0.019
	(.043)	(.078)	(.079)	(.078)	(.079)	(.085)
Post-Lib * Female		0.388**	0.386**	0.377**	0.359*	0.426**
		(.176)	(.176)	(.178)	(.182)	(.191)
Concentration, GFKF			Yes	Yes	Yes	Yes
Occupation dummies				Yes	Yes	Yes
Occupation * Female					Yes	Yes
R squared	0.5732	0.5735	0.5736	0.5814	0.5824	0.5826
No. of Observations	12823	12823	12823	12823	12823	12786

Notes: Dependent variable is log of weekly earnings of each individual. Numbers in parentheses are standard errors. Sample is restricted to higher castes Standard errors are clustered by industry.

*All regressions include demographic characteristics, state, education x female, education x ERP, industry, industry x PostLib, industry x female, ERP, IPR, ERP x PostLib, IPR x PostLib. #column 19 includes ICR, ICR x female, ICR x postlib, ICR x postlib x female. * denotes significance at 10% level, ** denotes significance at 5% level, *** denotes significance at 1% level.*

Table 3.8 Relation between Percentage of Females Employed and the Percentage Change in ERP (12)

Industry	ERP (In %) 1983	ERP (In %) 1999	%Female 1983	% Female 1999	% change ERP
Industries with highest ERP in 1983					
Fabricated Structural Metal Products	428.7	50.6	8.20	10.04	88
Iron & Steel	225.2	51.7	8.11	9.36	77
Cells & Batteries	199.9	61.8	0.00	8.26	69
Synthetic Rubber	173.1	40.6	8.92	10.52	77
Paints, Varnishes	171.7	39.2	6.28	2.46	77
Industries with lowest ERP in 1983					
Agricultural Machinery, Equipments etc	30.4	27.9	3.30	0.00	8.0
Locomotives and Parts	47.1	28.8	4.36	5.57	39
Food and Textile Machinery	48.7	29.3	0.00	4.95	40
Fertilizers and Pesticides	50.8	28.7	2.93	5.57	44
Wires and Cables	51.5	66.5	5.60	1.26	-29

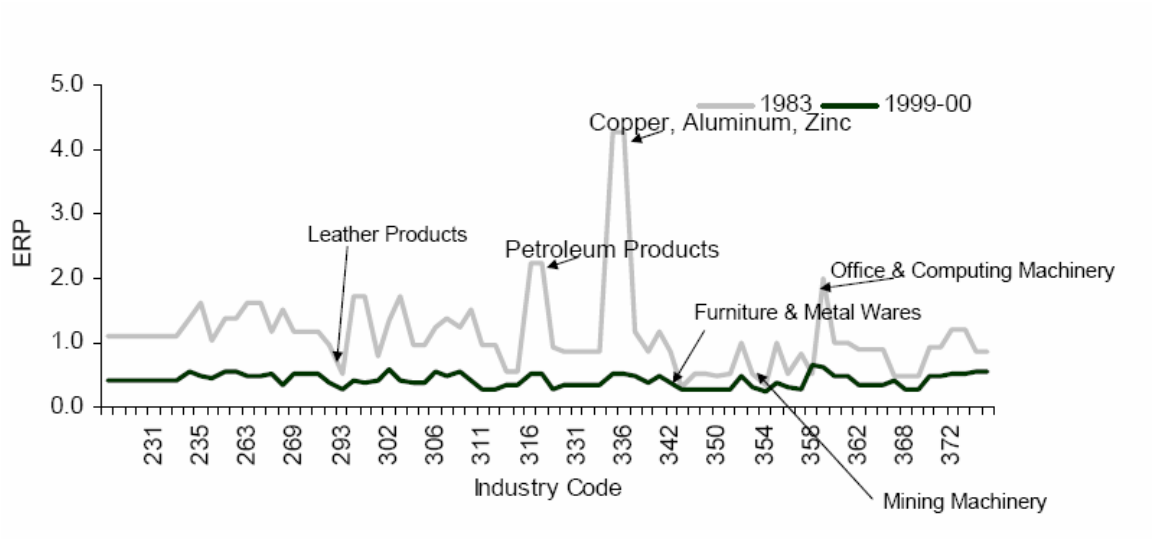
Notes: ERP denotes Effective Rate of Protection. % Female denotes percentage of female workers in each industry.

Table 3.9 Individual Level Results for Caste Groups (Male Hindus Only)

	(1)	(2)	(3)	(4)	(5)
ERP*Caste	**-0.068***	-0.050	-0.051	-0.045	-0.047
	(.039)	(.040)	(.039)	(.041)	(.041)
-IPR*Caste	**-0.425****	-0.376	-0.375	-0.305	-0.302
	(.211)	(.229)	(.228)	(.227)	(.229)
ICR*Caste	0.001	0.002	0.002	0.002	0.002
	(.001)	(.001)	(.001)	(.001)	(.001)
ERP*PostLib*Caste	0.140	0.083	0.080	0.099	0.096
	(.089)	(.095)	(.095)	(.093)	(.092)
-IPR * PostLib* Caste	0.451**	**0.377***	**0.376***	0.304	0.296
	(.205)	(.217)	(.217)	(.211)	(.211)
ICR * PostLib* Caste	-0.000	-0.001	-0.001	-0.001	-0.001
	(.001)	(.001)	(.001)	(.001)	(.001)
Caste	-0.186***	-0.212***	-0.211***	-0.194***	-0.201**
	(.053)	(.061)	(.061)	(.061)	(.080)
Post-liberalization	0.080	0.074	0.083	0.088	0.089
	(.078)	(.076)	(.078)	(.077)	(.077)
Post-Lib * Caste		0.143	0.144	0.142	0.142
		(.115)	(.115)	(.116)	(.117)
Concentration, GFKF			Yes	Yes	Yes
Occupation dummies				Yes	Yes
Occupation * Caste					Yes
R squared	0.5235	0.5236	0.5237	0.5322	0.5323
No. of Observations	13,322	13,322	13,322	13,322	13,322

Notes: Dependent variable is log of weekly earnings of each individual. Numbers in parentheses are standard errors. Sample is restricted to males. Standard errors are clustered by industry.
*All regressions include demographic characteristics, state, education x caste, education x ERP, industry, industry x PostLib, industry x caste, ERP, IPR, ICR, ERP x PostLib, IPR*PostLib, ICR*PostLib.*
denotes significance at 10% level, ** denotes significance at the 5% level, * denotes significance at the 1% level.*

Figure 2: Trends in Effective Rate of Protection in 1983 and 1999



Appendices

Appendix A. Data

A.1 List of States

Table A1: List of States in India

Included States				Other States & Union Territories		
1	Andhra Pradesh	9	Madhya Pradesh	17	Arunachal Pradesh	Andaman & Nicobar Islands
2	Bihar	10	Maharashtra	18	Assam	Chandigarh
3	Delhi	11	Orissa	19	Manipur	Dadra & Nagar Haveli
4	Gujarat	12	Punjab	20	Meghalaya	Daman & Diu
5	Haryana	13	Rajasthan	21	Mizoram	Lakshdweep
6	Himachal Pradesh	14	Tamil Nadu	22	Nagaland	Pondicherry
7	Karnataka	15	Uttar Pradesh	23	Sikkim	
8	Kerala	16	West Bengal	24	Tripura	
				25	Goa	
				26	Jammu & Kashmir	

States 1 to 16 is included in the analysis. The Union Territories and States listed in 17-26 are not included. Post 2000, Bihar was further divided into Bihar and Jharkhand; Madhya Pradesh was divided into Chattisgarh and Madhya Pradesh; Uttar Pradesh was divided into Uttaranchal and Uttar Pradesh.

A.2 List of Industries with Tariff Information

Table A2 List of Industries

Code NIC-87	Three-Digit Classification Description	Code NIC-87	Three-Digit Classification Description
230, 231, 232, 235	W&F cotton Khadi	331	Iron And Steel in SF form
233	W&F of Cotton- Handloom	332	Ferro Alloys
234	W&F of Cotton-Powerloom	333	Copper Manufacturing
236 , 260	Printing of Cotton Textiles	335	Aluminum manufacturing
262, 267	Knitted or Crocheted Textiles	336	Zinc Manufacturing
263	Blankets,Shawl,Carpets& Rugs	338+ 339	Metal scraps & Non Ferrous
265	Textile Garments &Accessories	340	Fab Structural Metal Prods
268,269	Textile Products nec	341	Fab Structural Metal nec
290	Tanning and Curing of Leather	342	Furniture & Fixtures
291	Leather Footwear		
292	Apparel of Leather & Subs	346	Metal Kitchen Ware
293	Leather Products & Substitutes	343+349	Hand-tools, Weights etc
299	Leather & Fur Products nec	350	Agricultural machinery, Equipments & Parts
		351	Construction /Mining Machinery
300	Organic & Inorganic Chemicals	352	Prime Movers & Boilers
		353	Food & Textile Machinery
301	Fertilizer & Pesticides	354	Other Machinery
		355	Refrigerators & Air conditioners
		356	General Purpose Machinery
302+306	Synthetic Rubber & Manmade Fiber	357	Machine-Tools & Accessories
303	Paints, Varnishes etc	358	Office & Computing Machinery
304	Drugs & Medicines	359	Special Purpose Machinery
305	Perfumes, Cosmetics & lotions	360	Electrical Industrial Machinery
308	Explosives etc	361	Wires & Cables
309	Chemical Products nec	362	Cells & Batteries
310	Tires & Tubes	365+366	Radio & TV
		363	Lamps & Domestic Appliances
311	Rubber & Plastic Footwear	368	Electronic Valves & Tubes
312	Rubber Products nec	369	X-ray Machinery
		370	Ships and Boats
313	Plastic Products nec	371	Locomotives & Parts
		372	Wagons & Coaches
314	Refined Petroleum Products	373 + 374	Motor Vehicles, Cars & products.
316	Refined Petro Products, nec	377	Aircraft & related products
318	Coke Oven Products	379	Transport equipment nec
319	Other Coal Tar Products	375	Motorcycles And parts
330	Iron and Steel in P/SF form	376	Bicycles and Parts

Bibliography

Aigner, Dennis J., and Glen G. Cain, "Statistical Theories of Discrimination in Labor Markets", *Industrial and Labor Relations Review*, Vol. 30, No.2 (Jan., 1977), 175-187.

Appleton, Simon, John Hoddinott and Pramila Krishnan, "The Gender Wage Gap in Three African Countries", *Economic Development and Cultural Change*, Vol. 47, No.2, 289-312, January 1999.

Arrow, Kenneth J., " The Theory of Discrimination" in Orley Ashenfelter and Albert Rees, eds., *Discrimination in Labor Markets* (Princeton, N.J., Princeton University Press 1978), 3-33.

Artecona, R. and W. Cunningham, "Effects of Trade Liberalization on the Gender Gap in Mexico", Policy Research Report on Gender and Development, Working Paper 21, World Bank, Washington D.C., 2002.

Attanasio, Orazio, Pinelopi K. Goldberg, and Nina Pavcnik, "Trade Reforms and Wage Inequality in Colombia", *Journal of Development Economics*, 74 (2004), 331-366.

Banerjee, Biswajit, and J.B. Knight, "Caste Discrimination in the Indian Urban Labor Market", *Journal of Development Economics*, Vol. 17 (1985), North Holland, 277-307.

Banga, Rashmi, "Liberalization and Wage Inequality in India", ICRIER Working Paper No.156, March 2005.

Becker, Gary S., "*The Economics of Discrimination*", (Chicago: University of Chicago Press), 1957.

Bergmann, Barbara R., "The effect on white incomes of discrimination in employment", *Journal of Political Economy*, 79(2), March/April 1971, 294-313.

Berik, Guinseli, Yana van der Meulen Rodgers, and Joseph E. Zveglich, Jr., "International Trade and Wage Discrimination: Evidence from East Asia", World Bank, Poverty Reduction and Economic Management Network, Policy Research Working Paper 3111, August 2003.

Bertrand, Marianne, "From the Invisible Handshake to the Invisible Hand? How Import Competition Changes the Employment Relationship", *Journal of Labor Economics*, Vol. 22, No. 4, 2004.

Betancourt, Roger, and Carlos Seiglie, "Economic Development in the Tropics: Fiction or Possibility", in Cuba in Transition: Papers and Proceedings of the Eighth Annual Meeting of ASCE, Vol. 8, 1999.

Betancourt, Roger, and Suzanne Gleason, "The Allocation of Publicly Provided Goods to Rural Households in India: On Some Consequences of Caste, Religion and Democracy", *World Development,* Vol. 28 (2000), No.12, 2169-2182.

Bhattacherjee, Debashish, "A Note on Caste Discrimination in a Bombay Automobile Firm", *Industrial Relations,* Vol.24, No.1, (Winter 1985).

Birdsall, Nancy, and Richard Sabot, ed. "*Unfair Advantage: Labor Market Discrimination in Developing Countries*", (The World Bank, Washington D.C.), 1991.

Black, Sandra E., and Elizabeth Brainerd, "Importing Equality? The Effects of Globalization on Gender Discrimination", Industrial and Labor Relations Review, 57(4), 540-559, July 2004.

Blinder, Alan S., "Wage Discrimination: Reduced Form and Structural estimates", *Journal of Human Resources*, (1973), 8: 436-455.

Brown,Randall S., Marilyn Moon, and Barbara S. Zoloth, "Incorporating Occupational Attainment in Studies of Male-Female Earnings Differentials", *The Journal of Human Resources*, Vol.15, No.1 (Winter 1980), 3-28.

Chattopadhyay R., and Esther Duflo, "Women as Policy Makers: Evidence from a Randomized Policy Experiment in India", Poverty Action Lab Paper No.1, March 2003.

Coate, Stephen and Glenn C Loury, The American Economic Review, Vol. 83, No.5 (December 1993), 1220-1240.

Das, Deb Kusum, "Quantifying Trade Barriers: Has Protection Declined Substantially in Indian Manufacturing?", ICRIER Working Paper No. 105, July 2003.

Das, Gurcharan, '*India Unbound: The Social and Economic Revolution from Independence to the Global Information Age*', Anchor Books, New York, 2000.

Deshpande, Ashwini, "Does Caste Still Define Disparity? A Look at Inequality in Kerala, India"; *AEA Papers and Proceedings*, Vol.90, 2000, No.2, 322-325.

Doeringer, Peter B. "Internal Labor Markets and Non-competing Groups", *The American Economic Review*, Vol. 76, No.2, Papers and Proceedings of the Ninety-Eighth Annual Meeting of the American Economic Association, (May 1986), 48-52.

Dutta, Puja Vasudeva, "Accounting for Wage Inequality in India", Poverty Research Unit at Sussex, Working Paper No. 29, January 2005.

Feliciano, Zadia M., "Workers and Trade Liberalization: The Impact of Trade Reforms in Mexico on Wages and Employment", *Industrial and Labor Relations Review*, Vol.55 (2001), No.1, 95-115.

Fontana, Marzia, "The Gender Effects of Trade Liberalization in Developing Countries: A Review of the Literature", University of Sussex at Brighton, Discussion Papers in Economics, 2003.

Fontana, Marzia, Susan Joekes and Rachel Masika, "Global Trade Expansion and Liberalization: Gender Issues and Impacts", *BRIDGE* Report No. 42, Brighton: Institute of Development Studies, 1998.

Gersen, Jacob E. "Markets and Discrimination: A Theoretical and Empirical Analysis", 2004.

Goldberg, Pinelopi K., and Nina Pavcnik, "Trade, Inequality and Poverty: What Do We Know? Evidence from Recent Trade Liberalization Episodes in Developing Countries", 2004a.

Goldberg, Pinelopi K., and Nina Pavcnik, "The Effects of the Colombian Trade Liberalization on Urban Poverty", NBER, 2004b.

Harrison, Ann, and Gordon Hanson, "Who gains from trade reform? Some remaining puzzle", *Journal of Development Economics*, Vol. 59, 125-154, 1999.

Heckman, James J. "Sample Selection Bias as a Specification Error", *Econometrica*, Vol.47, No.1, (Jan. 1979), 153-162.

Joekes, Susan "A Gender Analytical Perspective on Trade and Sustainable Development" in Trade, Sustainable development and Gender, UNCTAD, Papers prepared in support of the themes discussed at the Pre-UNCTAD X Expert Workshop on Trade, Sustainable Development and Gender, UN New York and Geneva, July 1999, 33-59.

Jolliffe, Dean and Nauro F. Campos. "Does Market Liberalization Reduce Gender Discrimination? Econometric Evidence from Hungary, 1986-1998", World Bank Research Paper, May 2003.

Levinsohn, James. "Testing the imports-as-market-discipline hypothesis", *Journal of International Economics* 35, pp 1-22, 1993.

Madheswaran, S., and T. Lakshmanasamy, "Occupational Segregation and Earnings Differentials by Sex: Evidence from India", *Artha Vijnana*, Vol. 38, No. 4, pp 372-386, Dec 1996.

Mincer, Jacob, "*Schooling, Experience and Earnings*", New York, Columbia University Press, 1974.

Mishra, Prachi, and Utsav Kumar, "Trade Liberalization and Wage Inequality: Evidence from India", IMF Working Paper, January 2005.

Neumark, David, "Employers' Discriminatory Behavior and the Estimation of Wage Discrimination," *Journal of Human Resources*, 23 (1988), 279-295.

New Economic Policies, Bureau of Parliamentary Studies and Training, Lok Sabha Secretariat, Orientation Program for New Members of Twelfth Lok Sabha, New Delhi, India, July 1998.

Niimi, Yoko, Puja Vasudeva-Dutta and Alan Winters, "Trade Liberalization and Poverty Dynamics in Vietnam", PRUS Working Paper No.17, April 2003.

Nordås, Hildegunn Kyvik, "Is Trade Liberalization a Window of Opportunity for Women?", World Trade Organization, Economic Research and Statistics Division, Staff Working Paper ERSD – 2003.

Oaxaca, Ronald, "Male-Female Wage Differentials in Urban Labor Markets", *International Economic Review*, 14(3), October (1973), 693-709.

Oaxaca, R. and M. Ransom, "Discrimination and Wage Decomposition," *Journal of Econometrics,* 61 (1994), 5-21.

Oostendorp, Remco H., "Globalization and the Gender Wage Gap", World Bank Policy Research Working Paper 3256, 2004.

Phelps, Edmund S., "The Statistical Theory of Racism and Sexism", *American Economic Review*, Vol. 62, No.4, (Sept 1972), 659-661.

Reich, Michael J., David M. Gordon, and Richard C. Edwards, "A Theory of Labor Market Segmentation", *American Economic Review*, Vol. 63, May 1973, 359-365.

Revenga, Ana, "Employment and Wage Effects of Trade Liberalization: The Case of Mexican Manufacturing", *Journal of Labor Economics*, Vol.15, No.3, Part 2: Labor Market Flexibility in Developing Countries, S20-S43, 1997.

Sambamoorthi, Usha, "Labor Market Discrimination Against Women in India", Institute for Social and Economic Change, Working Paper No. 58, Bangalore, India, May 1984.

Sarvekshana, 87th Volume, Government of India publication, March 2002.

Scoville ed. "*Status Influences in Third World Labor Markets: Caste, Gender and Custom*", (Walter de Gruyter, Berlin, New York), 1991.

Topalova, Petia, "Factor Immobility and Regional Impacts of Trade Liberalization: Evidence on Poverty and Inequality from India", NBER Working Paper, 2005.