A Distributional Analysis of Public–Private Wage Differential in India

Mehtabul Azam[1] — Nishith Prakash[2]

Abstract. In this paper, we examine public–private wage differential among men in India across the entire wage distribution. We find that the raw wage gap between public and private sector is positive across the entire wage distribution in both urban and rural areas. A quantile regression-based decomposition reveals that that the public sector workers enjoy a positive wage premium across the entire wage distribution in both urban and rural areas, although the magnitude of wage premium is smaller at the top quantiles.

1. Introduction

In this paper, we examine public–private wage differential among men in India across the entire wage distribution using nationally representative National Sample Survey (NSS) data collected in 2009–10. We investigate whether the differences in wages exist across the entire wage distribution or concentrated in tails. We also explore whether the differences observed at each percentile is explained by differences in characteristics of workers across the two sectors. For this, we use a quantile regression-based decomposition technique proposed in Machado and Mata (2005; MM henceforth). The MM technique is a natural choice as it is well suited to capture the heterogeneity in wage gap that is explained (unexplained) by the differences in characteristics between the two sectors across the entire wage distribution.

We believe that it is important to study the wage differences across the entire wage distribution rather than concentrating on averages. Evidence from many other countries — for example Canada (Mueller, 1998), UK (Disney and Gosling, 1998), Zambia (Nielsen and Rosholm, 2001), and Germany (Melly, 2005) — suggest that the least squares estimate of the mean public sector wage premium gives an incomplete picture of the conditional wage distribution. In addition, there is a growing perception among the public sector employees in India, especially those at the top end of the wage distribution, that they are paid less than the similarly qualified counterparts in the private sector.[1] However, evidences either in favor or against this perception are lacking. The existing evidences in India suggest that on an average, public sector workers are paid a large positive wage premium. However, these are not very helpful to corroborate or contradict the prevailing perception because of two reasons. First, the existing evidences refer back to the 1990s or earlier, and there has been a significant change in the structure of the Indian economy in the last two decades. Second and most importantly, the existing evidences tell that the average differences and do not answer whether there exist any differences at the opposite tails. It is generally believed that as a good employer, the public

[1]Oklahoma State University & IZA, Stillwater, USA. E-mail: mazam@okstate.edu.
[2]University of Connecticut, IZA & CReAM, Storrs, USA.

LABOUR 29 (4) 394–414 (2015) DOI: 10.1111/labr.12068

JEL J3, J45

sector may be willing to pay lower skilled workers more, whereas the government might be reluctant to offer higher wages to high-skilled workers, as the public may not want public servants to earn more than comparably trained and experienced private sector counterparts (Katz and Krueger, 1993).

The public sector employees' wages has been an active research area in the industrialized nations. Some of the earlier works comparing the earnings of public sector employees in the United States was undertaken by Smith (1976, 1977). Smith finds that the rate of pay is higher for public sector employees than for private sector employees, and the wage premium is larger for female than for male public sector employees. Bender (1998) and Gregory and Borland (1999) provide extensive surveys of the literature on public–private wage differentials in a range of countries and report that central governments pay more on average than the private sector, even after controlling for differences in productivity characteristics. For example, federal government workers earn between 5 and 30 per cent more than their private sector counterparts in the United States. Katz and Krueger (1993) offer several institutional and political reasons for the gap in wages between public and private sector employees. Giordano *et al*. (2011) find a conditional pay gap in favor of the public sector after controlling for differences in employment characteristics between the two sectors in 10 euro-area countries. Their analysis also highlights substantial heterogeneity across countries, with Greece, Ireland, Italy, Portugal, and Spain exhibiting the highest public sector premia. Depalo *et al*. (2015) evaluate the public–private wage differential for men in 10 euro-area countries in the period 2004–07. They find that the pay gap is often decreasing over the distribution and it is mostly determined by higher endowments in the upper tail of the wage distribution and by higher returns of such endowments at the low tail, with considerable heterogeneity across countries. Lausev (2014) provides a survey of the literature in countries transitioning from communist to market-based economies. Unlike developed and developing countries, a negative public sector pay gap is commonly estimated in those transitioning countries.

In contrast, the understanding of public–private wage differentials in India has received relatively less attention despite the size of public sector. The first set of evidence refers back to the early 1980s, where a number of studies (Duraiswamy and Duraiswamy, 1995; Lakshmanasamy and Ramasamy, 1999; Madheswaran, 1998; Madheswaran and Shroff, 2000) use the survey of Degree Holders and Technical Personnel conducted along with the 1981 Census of India. They find that on an average, wages in the private sector is higher than the public sector. However, the evidence from the 1990s shows the opposite. Data on wages from the 1997 to 1998 Annual Survey of Industries (ASI), which cover the factory sector, suggest that the public sector offers higher remuneration to workers compared with the private sector (Glinskaya and Lokshin, 2007). Using the National Sample Survey (NSS) data for 1993–94 and 1999–2000, Glinskaya and Lokshin explore public–private wage differential in India using ordinary least squares (OLS) and propensity score matching methods. They find that that the public sector wage premium in 1999–2000 ranges from 62 to 102 per cent over the private-formal sector and between 165 and 259 per cent over the informal-casual sector depending on the choice of the methodology. They also find that the wage differentials tend to be higher in rural areas compared with urban areas, higher among women than among men, and higher among low-skilled workers compared with high-skilled workers.

The wage premium enjoyed by the public sector workers in India is not surprising as in most cases the public sector wants to be a good employer and may be willing to pay higher wages to its employees. Although the wages in the private sector is determined based on profit considerations, the wages in public sector continue to be decided by the government. The central government sets wages for the public sector employees through the Pay Commission, which is a central body at the federal level. The Pay Commission in India tries to maintain the ratio between the minimum pay scale and per capita income.[2]

Thus, whether the wage premium enjoyed by the public sector employees in the 1990s (as reported in Glinskaya and Lokshin, 2007) exists even in 2010, and whether this wage premium differs at the tails remain empirical questions.[3] These questions remain interesting as aligning the average rates of pay will be misleading if the wage distributions between public and private sector differ significantly at tails. Evidence from other countries suggest that the distributions do differ, as found in the double imbalance concept articulated by Schager (1993), Katz and Krueger (1993), and Elliott and Duffus (1996). This imbalance is characterized by public sector workers at the lower end of the wage distribution receiving the largest wage premium, whereas the public sector workers at the top end of distribution earn less than their private sector counterparts. If the Indian experience is also similar to this (as claimed by group A employees — top end public sector employees — in their memoranda to Sixth Pay Commission), then policies aimed at aligning the wage structures at the top of the distribution should be followed. In this paper, we seek to address these questions by examining the differences in the wage distributions of public and private sector.

Our main findings are as follows. First, public–private wage differential among male workers is positive across the entire wage distribution in both urban and rural India. Second, the observed wage gap is not entirely explained by differences in characteristics of workers (covariate effect) across most of the wage distribution. Indeed, the public sector workers enjoy wage premium across the entire wage distribution, and contribution of the wage premium (coefficient effect) in total wage gap is larger than the contribution of differences in workers' characteristics in most part of the distribution except at the upper tails. Thus, we do not find evidence of double imbalance in India, as the wage premium enjoyed by public sector workers is positive across the entire distribution. It is true that the wage premium for public sector workers is lower at the top quantiles; however, it remains positive throughout the distribution. Moreover, these wage premium estimates provide lower bounds as public sector jobs in India come with many benefits, which are generally not captured in wage data, for example, pension, housing, etc. Third, we find that the averages mask a great deal of heterogeneity not only in the observed raw wage gap, but also in the coefficient and covariate effects.

This paper is organized as follows. Section 1.1 provides a brief description of Indian labor market, Section 2 discusses the empirical strategy, Section 3 describes the data, Section 4 presents the results, and Section 5 concludes.

1.1 Indian labor market

In 2013, there were around 481 million workers in India, the second largest after China.[4] Of these, over 94 per cent work as self-employed in agriculture or nonagriculture activities or in unorganized enterprises. Various agencies collect information on employment and wages in India.[5] The National Sample Survey conducts nationally representative household surveys that collect employment information (wages, type of work, industry, occupation, etc.) of individuals in addition to many other individual characteristics. The ASI collects employment data for organized manufacturing sector by surveying firms, and thus wage data are not linked to individual characteristics. The Directorate General of Employment and Training (DGET) in the Ministry of Labour & Employment is another source of data pertaining to employment market information (EMI). Under the EMI program, collection of data is done following the establishment reporting system. The establishments are required to furnish at regular intervals details about the number of persons they employ by sex, vacancies that have occurred, and the type of persons, which are in short supply. This information relates only to 'Employers' and 'Employees'. The data collected under the EMI program covers only the organized sector of the economy which covers all establishments in the public sector irrespective of their size, and nonagricultural establishments in the private sector employing 10 or more persons (Government of India, 2012).

The NSS surveys are widely used to generate estimates of various characteristics pertaining to employment, unemployment, and labor force at the national and state levels. In 2011–12, Work Force Participation Rate (WFPR) for persons of age 15–59 years was 57 per cent at the all-India level. WFPR for rural males was the highest (82 per cent) followed by urban males (78 per cent), rural females (37 per cent), and it was the lowest for urban females (21 per cent) (Government of India, 2013). In addition, many official documents also use estimates from DGET for organized sector employment size.[6]

During the planning era 1950–91, public sector was given a lot of attention and was considered as an instrument for the rapid growth. A number of steps were taken by central government to strengthen public sector that included transfer of ownership from private to public sector, covering the nationalization of commercial banks in 1969, nationalization of insurance in 1972, takeover of coal industry in public sector in 1973, and Foreign Exchange Regulation Act (1973). By 1981, public sector employment increased to about 68 per cent of the total employment in the organized public and private sectors.[7] Following fiscal crisis in 1991, the public sector has undergone some major structural changes with a greater emphasis on market forces. The structural measures include removing the restrictions on foreign direct investment, reduction in the number of strategic sectors reserved for the public sector, disinvestment of public sector units and financial sector reforms. Contrary to 1950–91, there has been a tendency of transferring the public ownership to private sector, and an enormous importance is given to private investment in core sectors of the economy. There has been a transition from the public sector to the private sector, but the volume of this transition is quite small. In 2008, the organized sector employed 27.5 million workers, of which 17.3 million (about 63 per cent) worked for government or government-owned entities.[8]

2. Empirical methodology: decomposing the wage gap

We use the MM technique to decompose the wage gap between public and private sector at each percentile into two components — one that is explained by the differences in the 'characteristics' of workers across the two sectors (called the *covariate effect*) and the other due to differences in 'prices', i.e. differences in returns to those characteristics (called the *coefficient effect*). The idea underlying the MM technique is that the conditional quantiles of w, given by $Q_\theta(w|x) = x'\beta(\theta)$, can be estimated by quantile regression. The conditional quantile process — that is, $Q_\theta(w|x)$ as a function of $\theta \in (0, 1)$ — provides a full characterization of the conditional distribution of wages given x. The estimated conditional quantile function parameters can be used to simulate the conditional distribution of w given x via an application of the probability integral transformation theorem. Thus, the MM technique combines quantile regressions with bootstrap methods to generate distributions that conform to the specified model. The decomposition involves creating three distributions of (log) wage — private, public, and counterfactual. The private (public) distribution of (log) wage is estimated using private (public) covariates and private (public) returns to those covariates, and thus conform to the specified model, whereas the (log) wage distribution for counterfactual group is estimated using covariates of one sector and returns to those covariates in other sector.

Following Albrecht *et al.* (2003) adaption of MM technique, the three distributions are generated as follows.

- For each sector, $\tau = Pub$, Pvt, we estimated 99 quantile regressions (QR) on a fine grid of quantiles, $\theta = [0.01, 0.02, \ldots 0.98, 0.99]$

$$Q_\theta(w|x;\tau) \tag{1}$$

which yields 99 estimates of the Quantile Regression (QR) coefficients $\hat{\beta}(\theta_i;\tau)$ for each sector.

- Then we draw m = 1,000 random draws with replacement from the distribution of covariates for each $\hat{\beta}(\theta_i;\tau)$'s and stack $x(\tau)'\hat{\beta}(\theta_i;\tau)$ to get the marginal distribution for sector τ which conforms to our model.[9]

$$w^* = \left\{x^*(\tau)'\hat{\beta}(\theta_i;\tau)\right\}_{i=1}^{i=999} \tag{2}$$

- To construct the counterfactual density, we take $\hat{\beta}(\theta_i;\tau)$ estimated from sample of private (public) sector workers, and instead of drawing from the distribution of covariates from private (public) sector workers, we draw it from the rows of the public (private) sector, i.e. $x(Pub)$ ($x(Pvt)$).

Let $f(w(\tau))$ denote the estimate of the marginal density of w (log of wages) in sector τ based on observed sample, i.e. the empirical density, and $f^*(w(\tau))$ is the marginal density implied by the model. Extending this notation to the counterfactual distributions, we may define $f^*(w(Pvt); x(Pub))$ as the density that would have prevailed if public sectors workers were paid the 'prices' prevailing in the private sector. Letting α be a usual summary statistic (for instance, quantile or scale measure), we may decompose differences in α as:

$$\begin{aligned}\alpha\{f(w(Pub))\} - \alpha\{f(w(Pvt))\} &= \alpha\{f^*(w(Pub))\} - \alpha\{f^*(w(Pvt))\} + residual \\ &= \underbrace{\alpha\{f^*(w(Pub))\} - \alpha\{f^*(w(Pvt); x(Pub))\}}_{\text{Coefficient effect}} \\ &\quad + \underbrace{\alpha\{f^*(w(Pvt); x(Pub))\} - \alpha\{f^*(w(Pvt))\}}_{\text{Covariate effect}} + res \end{aligned} \tag{3}$$

This decomposition will then give us the contribution of covariates, coefficients, and an unexplained part (residual).[10] The residual is the difference that is unaccounted by the estimation method (estimated as the difference between the estimates of the total changes derived by using the empirical wage density and by using the estimated marginal densities).[11]

3. Data

We use nationally representative household survey data from the Employment and Unemployment Schedule, administered by the NSS Organization (NSSO), Government of India. For this study, we use data from 66th round of the NSS conducted in 2009–10.[12] The data contain information on household size and composition, social group, religion, monthly consumption, landholdings, demographic variables (age, gender, marital status), educational participation and attainment, and a detailed employment section on principal and subsidiary activities (industry, occupation, and wages earned). The sample of households is drawn based on a stratified random sampling procedure, and all the analysis is done using survey weights.

The NSS data classify workers in three categories; self-employed, regular wage salaried, and casual workers. A regular wage-salaried worker is a person who works in others' farm or nonfarm enterprises (household and non-household), and in return, received salary or wages on a regular basis. On the other hand, a casual worker is a person who is engaged in others' farm or nonfarm enterprises, and in return, received wages according to the terms of the daily or periodic work contract. Weekly wages are reported at current prices for regular wage/salaried and casual workers, whereas there is no wage information for self-employed. One can distinguish public sector employment based on enterprise type.[13] As shown in Table A2, only few casual workers report to be working in the public sector. In addition, most of the existing wage literature in India has modeled casual and regular-salaried workers separately because of difference in structural parameters associated with these two types of workers. For example, Dutta (2006) finds that while the explanatory power of human capital variables is reasonably high for regular wage/salaried workers, which for casual workers are very low, suggesting that human capital characteristics do not explain the wage determination process for casual workers. Given that almost all of the public employees are regular-salaried workers, we compare public sector workers to private sector regular-salaried workers only and not to casual workers, i.e. we restrict our sample to regular-salaried male workers of age 18–60.[14,15] Around 2.5 per cent of regular-salaried male workers do not report their enterprise type. Those workers are dropped from the analysis. As the sample is restricted to regular-salaried male workers, we discuss the selection issues in detail in Section 3.1.[16]

3.1 Selection issues

The selection of sample raises questions about selection bias. Selection issues for the analysis of the public sector wage gap have been strongly emphasized in the literature based on mean regression. Our sample construction involves selection at two stages. At first stage, the selection into regular-salaried workers may be nonrandom. Second, there may be nonrandomness involved in the choice of sector. In the context of public–private wage differential conditional on being in regular-salaried employment, selection into regular-salaried employment may not be as problematic as long as the selection mechanism remains similar for both public and private regular-salaried jobs.[17,18]

However, conditional on being in regular-salaried work, there is a choice being made by workers whether to work in the public or private sector, which might bias the public–private wage differential. We follow Buchinsky's (1998) approach for the selection correction in a quantile regression model by using the power series expansion of the inverse Mills ratio of the normalized estimated index. Following Buchinsky (1998), we estimate:

$$Q_\theta(w|x, public=1) = x'\beta(\theta) + h_\theta(z'\gamma) \quad [4]$$

Buchinsky (1998) suggests a series estimator (a power series approximation) for this flexible selection term

$$\hat{h}_\theta(z'\gamma) = \delta_0(\theta) + \delta_1(\theta)\lambda(z'\gamma) + \delta_2(\theta)\lambda(z'\gamma)^2 + .. \quad [5]$$

where $\lambda()$ is inverse mills ratio (imr); and the δ's vary with θ. We estimate the unknown parameter γ using the Klein and Spady (1993) estimator for binary choice models, and approximate $\hat{h}_\theta(z'\gamma)$ including $\lambda(z'\gamma)$, and $\lambda(z'\gamma)^2$.[19] As reported in Table A3, we do not find strong evidence (statistically insignificant selection terms) of selection at the selected quantiles. Therefore, we proceed with our MM decomposition exercise without correcting for sample selection.[20] Henceforth, our results should be interpreted conditional on being in regular-salaried employment.

3.2 Descriptive statistics

Weekly wages earned (both cash and in-kind) are reported at current prices for regular wage-salaried and casual workers. We define total wages as sum of weekly cash and in-kind wages from the principal activity.[21] Only 13 per cent of workers in our working sample reported in-kind wages in addition to the wages in cash, and in-kind wages constitute only 13.4 per cent of total wages for those who reported in-kind wages. The wages are self-reported weekly earnings, and since our sample is restricted to regular salaried workers for whom taxes are deducted at source, the wages are more likely to be net wages. However, from the information provided in the survey, it is difficult to ascertain whether it is a gross or a net wage.[22]

We divide the total wages by number of hours worked in the reference week to get the hourly wages. As evident from the descriptive statistics presented in Table 1, average hourly wages are higher in the public sector in both rural and urban areas: public sector workers earn 90 per cent more hourly wage than the private sector workers in rural areas, and 85 per cent more in urban areas. Figure 1 plots the kernel density of the log of hourly wages in public and private sectors for urban and rural areas, respectively. In both rural and urban areas, public sector wage distribution lies right of private sector wage distribution suggesting that the public sector workers are paid higher across the entire distribution in both rural and urban sector. Figure 2 that plots the difference in log wage distribution between public and private sector workers at each percentile confirms that the public sector workers are paid more across the entire wage distribution in both urban and rural areas, and the wage gap is very heterogeneous. There exists a positive wage gap at the lower quantiles in both urban and rural areas, and the wage gap increases as one move toward the middle of the wage distribution. The wage gap decreases in the upper half of the distribution, however, it remains positive throughout the distribution. The wage gap in urban areas is larger compared with wage gap in rural areas in lower half of the distribution, however, in the upper half of the distribution the wage gap in rural areas is larger than the wage gap observed in urban areas.

Looking at the unconditional distribution can be misleading, as there exist notable differences in the characteristics of public and private sector workers. For example, as evident from the descriptive statistics presented in Table 1, public sector workers are older than the private sector workers in both urban and rural areas. In addition, public sector workers also have higher educational attainment than the private sector workers: 35 per cent of public sector workers in rural India have graduate or postgraduate degree, whereas only 15 per cent of private sector workers in rural India have graduate or postgraduate degree. Similarly, 48 per cent of public sector workers in urban areas have graduate or postgraduate degree. In comparison, only 27 per cent of private sector workers in urban areas have graduate or postgraduate education. Public sector workers are more likely to be in professional or clerical jobs, whereas the private sector workers are more likely to be in production-related activities.

In the next section, we explore whether those advantages in endowments for public sector workers explain the observed wage gap, or public sector workers enjoy wage premium (paid higher prices for similar endowments). Our dependent variable is log of hourly wages and explanatory variables include age, age square, dummies for states, education levels, not married, Scheduled Castes (SCs)/Tribes (STs), Other Backward Castes (OBCs), Muslim,

Table 1. Descriptive statistics

	Rural		Urban	
	Public	**Private**	**Public**	**Private**
Log of hourly wage	3.71	2.81	4.08	3.23
Age	41.65	32.57	42.90	33.41
Currently not married	0.10	0.31	0.09	0.30
Social Group				
Scheduled Castes/Tribes	0.29	0.25	0.22	0.15
Other Backward Castes Muslim	0.37	0.41	0.30	0.35
	0.08	0.12	0.06	0.12
Education				
Primary school	0.06	0.14	0.04	0.10
Middle school	0.11	0.24	0.08	0.18
Secondary school	0.18	0.18	0.16	0.20
Higher Secondary	0.17	0.10	0.14	0.11
Diploma/certificate course	0.07	0.04	0.06	0.04
Graduate, bachelor	0.27	0.12	0.32	0.20
Postgraduate and above	0.09	0.03	0.16	0.07
Occupation				
Professionals	0.11	0.06	0.15	0.09
Technicians and associate professionals	0.27	0.07	0.16	0.09
Clerks	0.18	0.06	0.21	0.09
Service workers and shop and market sales workers	0.19	0.15	0.14	0.18
Craft and related trades workers	0.06	0.22	0.08	0.23
Plant and machine operators and assemblers	0.07	0.27	0.08	0.15
Elementary occupations	0.09	0.15	0.10	0.10
Industry				
Manufacturing of food, beverages	0.04	0.15	0.03	0.15
Manufacturing of chemical, equipment	0.01	0.15	0.03	0.18
Electricity, water, and sewerage	0.00	0.03	0.01	0.04
Wholesale and retail trade, repair	0.02	0.05	0.01	0.04
Transportation and storage	0.01	0.14	0.01	0.17
Accommodation and food service activities	0.00	0.03	0.00	0.04
Information and communication	0.06	0.21	0.11	0.09
Financial and insurance activities	0.06	0.04	0.10	0.07
Professional, scientific, and technical activities	0.40	0.02	0.48	0.08
Public administration, education, and other services	0.03	0.02	0.04	0.02
Others	0.32	0.13	0.15	0.09
Number of observations	6,276	4,736	7,646	10,416

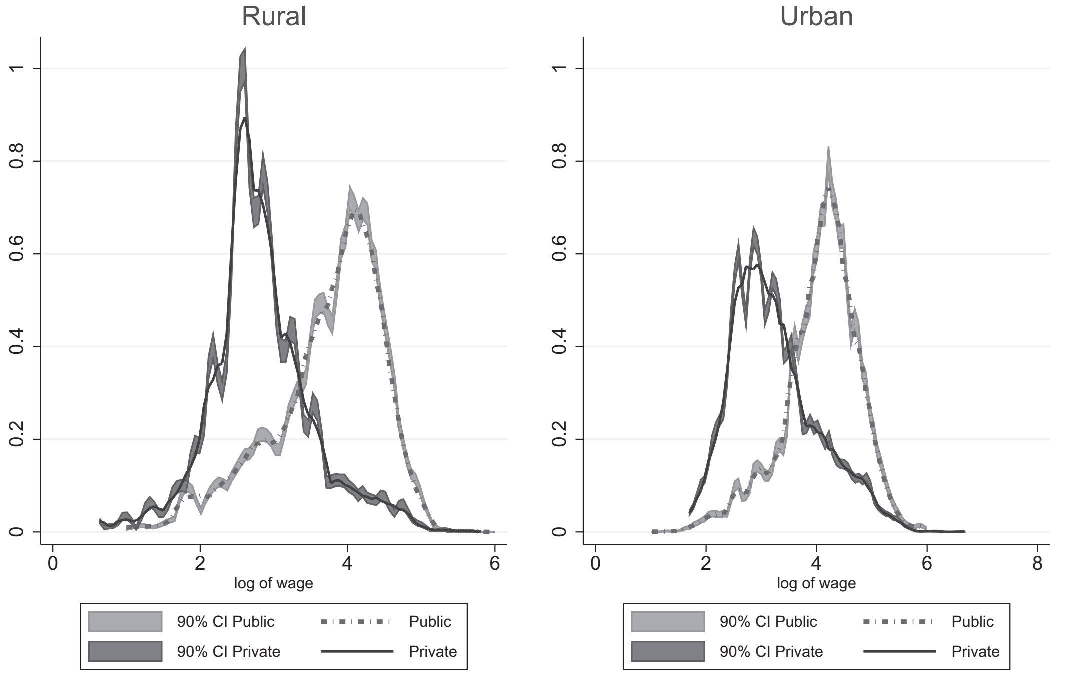
Note: Sample restricted to regular salaried male workers in age 18–60.

occupations, and industries. The SCs/STs and OBCs are the disadvantaged groups, and enjoy affirmative policies in India, whereas Muslims are the religious minority group. Table 1 presents the descriptive statistics of variables included in the covariate matrix (dummies for states are included in x matrix but are not reported in the table).

4. Results

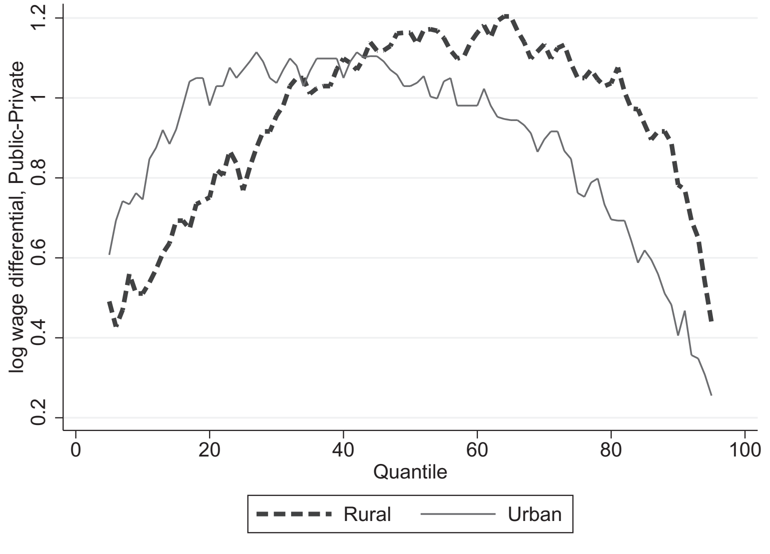
We present the results of OLS and quantile regression for public and private sector in rural and urban India in Tables A3 and A4, respectively. As expected, the OLS results show that the

Figure 1. Kernel density of log hourly wages



Note: The confidence intervals are calculated using Stata Journal asciker command. For more details see Fiorio (2004).

Figure 2. Wage gap between public and private



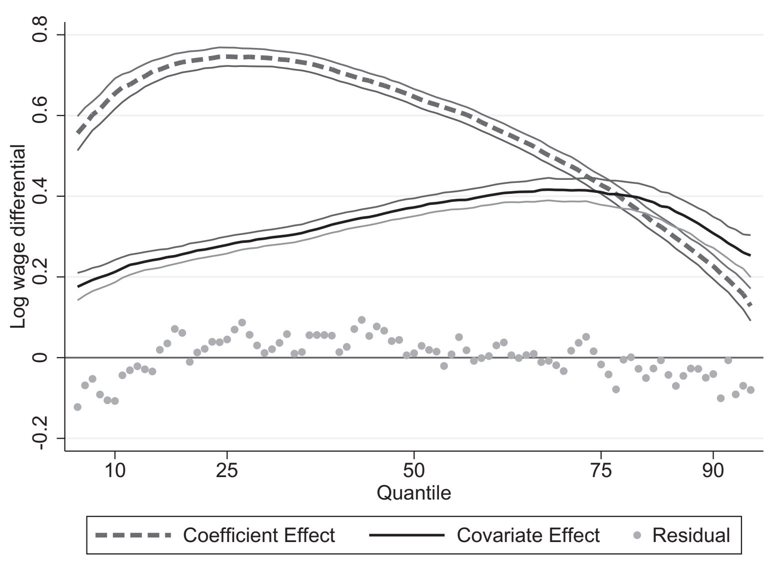
Note: The wage gap difference between actual empirical (log) wage distributions of public and private sector at each quantile α, i.e. $\alpha f(w(Pub)) - \alpha f(w(Pvt))$ where $\alpha \in (0, 1)$.

returns increase with the education levels in both public and private sectors in both rural and urban areas. However, the OLS estimates conceal the heterogeneity in returns across the distribution. Quantile regression estimates suggest very interesting pattern in returns to post-secondary education. Although the returns to post-secondary (bachelor and postgraduate degree) education increase at higher quantiles in private sector in both urban and rural areas, returns for those education groups show a declining trend at higher quantiles in public sector. This suggests that while public sector compresses the wages within higher education, private sector increases within-group wage inequality.

Since returns vary across the distribution and as discussed in the data section, the distribution of covariates differs considerably across public and private sector, we therefore move to MM decomposition that aggregates wage gap explained by the covariates into covariate effect and wage gap explained by the differences in returns to those covariates into coefficient effect. Figures 3 and 4 plot the coefficient and the covariate effects with 95 per cent confidence intervals for urban and rural areas, respectively. Table 2 reports the decomposition results for selected quantiles. The 95 per cent bootstrap confidence intervals are the quantiles 2.5 per cent and 97.5 per cent of the distribution of the relevant statistic obtained by bootstrap with 1,000 replications.

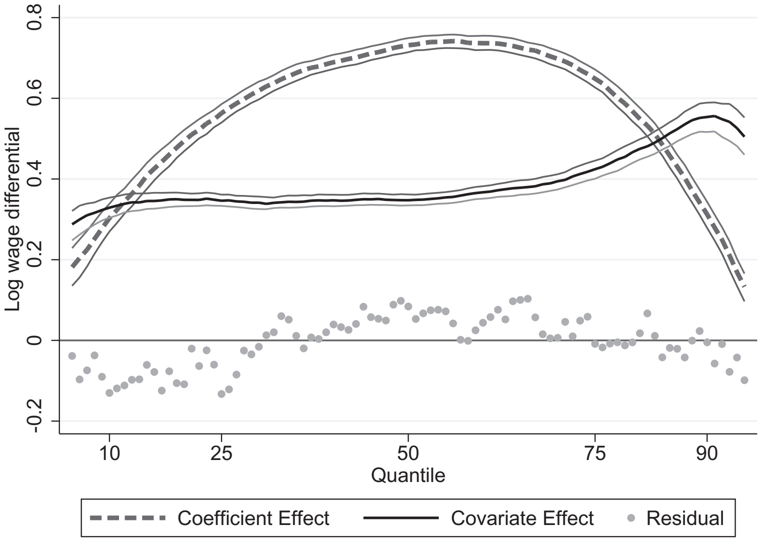
In urban India (Panel A of Table 2), public sector workers earn 85 log points more on average compared with their private sector counterparts. However, the average wage gap conceals the heterogeneity in wage gap observed across the distribution. For example, the wage gap is 76 log points at 10th percentile, 107 log points at 25th percentile, 103 log points at 50th percentile, 82 log points at 75th percentile, and only 49 log points at 90th percentile. How much of this wage gap is explained by the better characteristics of public sector workers (as discussed in Section 3.2, Table 1)? The Oaxaca Blinder (OB) decomposition, based on OLS, suggests that differences in covariates explain only about 34 log points wage gap (about

Figure 3. Decomposition of Public–Private Wage Gap, Urban



Notes: Coefficient effect refers to difference in $\alpha f^*(w(Pub)) - \alpha f^*(w(Pvt);\ x(Pub))$, and covariate effect refers to $\alpha f^*(w(Pvt));\ x(Pub) - \alpha f^*(w(Pvt))$, where $\alpha \in (0, 1)$. See text for further details.

Figure 4. Decomposition of Public–Private Wage Gap, Rural



Notes: Coefficient effect refers to difference in $\alpha f^*(w(Pub)) - \alpha f^*(w(Pvt);\ x(Pub))$, and covariate effect refers to $\alpha f^*(w(Pvt));\ x(Pub) - \alpha f^*(w(Pvt))$, where $\alpha \in (0, 1)$. See text for further details.

40 per cent of total wage gap) between public and private sector. The rest (60 per cent of the total wage gap) is the wage premium enjoyed by public sector workers, i.e. public sector pays a positive wage premium (about 51 log points) on average for a similar characteristics worker. Thus, although the public sector workers have better characteristics, it fails to explain the wage differential completely as more than half of the wage differential is explained by the price differential, i.e. workers with same characteristics are paid better in public sector compared with private sector.

MM decomposition brings out the heterogeneity in the coefficient and covariate effects, which were not captured by the OB decomposition. MM decomposition tells us that the differences in covariates only account for 21 log points (28 per cent of the wage gap) at 10th percentile, 28 log points (25 per cent of the wage gap) at 25th percentile, 37 log points (36 per cent of the wage gap) at 50th percentile, 41 log points (45 per cent of the wage gap) at 75th percentile, and 31 log points (62 per cent of the wage gap) at 90th percentile. Thus, better characteristics do not explain the entire wage gap across the distribution, although they account for increasingly higher share at higher quantiles. The fitted model does reasonably well in predicting the wage gap, as the residuals are mostly around zero except at the 10th and 90th percentile, where the predicted wage gap is more than the observed gap by about 10 per cent.[23] Importantly, the MM decomposition suggests that the public sector workers enjoy a positive wage premium across the entire wage distribution. This wage premium is quite large at lower quantiles. For example, the wage premium enjoyed by public sector workers is 66 log points (86 per cent of the wage gap) at 10th percentile, 65 log points (63 per cent of the wage gap) at 50th percentile, and 23 log points at 90th percentile (46 per cent of the wage gap). More importantly, not only the wage premium enjoyed by the public sector workers is positive throughout the wage distribution but also, the coefficient effect (wage premium) dominates the covariate effect in most part of the distribution, and

Table 2. Decomposition of public private wage gap

Panel A: Urban				
Quantile	**Actual difference**	**Coefficient effect**	**Covariate effect**	**Residual**
Mean (OLS)	0.847	0.511 *0.482; 0.541*	0.336 *0.307; 0.364*	0.000
10	0.760	0.655 *0.616; 0.692*	0.212 *0.187; 0.242*	–0.107
25	1.070	0.746 *0.723; 0.768*	0.279 *0.258; 0.300*	0.045
50	1.030	0.646 *0.625; 0.665*	0.372 *0.351; 0.395*	0.011
75	0.821	0.428 *0.406; 0.451*	0.410 *0.379; 0.442*	–0.017
90	0.494	0.227 *0.195; 0.255*	0.308 *0.272; 0.349*	–0.041
Panel B: Rural				
Quantile	**Actual difference**	**Coefficient effect**	**Covariate effect**	**Residual**
Mean (OLS)	0.895	0.482 *0.423; 0.542*	0.413 *0.357; 0.469*	0.000
10	0.502	0.303 *0.270; 0.339*	0.329 *0.304; 0.354*	–0.130
25	0.777	0.563 *0.540; 0.586*	0.346 *0.333; 0.361*	–0.133
50	1.163	0.732 *0.714; 0.749*	0.347 *0.335; 0.362*	0.084
75	1.064	0.649 *0.627; 0.671*	0.424 *0.401; 0.446*	–0.009
90	0.862	0.313 *0.279; 0.344*	0.554 *0.516; 0.589*	–0.005

Note: The actual difference is the difference between actual empirical (log) wage distributions of public and private sector at each quantile *a*, i.e. $\alpha\{f(w(Pub))\} - \alpha\{f(w(Pvt))\}$, where $\alpha \in (0,1)$. Coefficient effect refers to difference in $\alpha\{f^*(w(Pub))\} - \alpha\{f^*(w(Pvt);\ x(Pub))\}$, and covariate effect refers to $\alpha\{f^*(w(Pvt);\ x(Pub))\} - \alpha\{f^*(w(Pvt))\}$. See text for further details. The decomposition at mean is derived using the Oaxaca Blinder decomposition using ordinary least squares (OLS).

only at the top end of the distribution (after 75th quantile), the covariate effect becomes larger than the coefficient effect.

We find qualitatively similar results for rural India. Public sector workers in rural India are paid on average 90 log points more compared with their private sector workers. Thus, the average wage gap in rural India is marginally higher than urban India. Similar to urban areas, the average wage gap conceals the heterogeneity in the observed wage gap in rural areas also: the observed wage gap is 50 log points at 10th percentile, 77 log points at 25th percentile, 116 log points at 50th percentile, 106 log points at 75th percentile, and 86 log points at 90th percentile. OB decomposition tells us that better covariates of public sector workers only account for 41 log points wage gap (about 46 per cent of the wage gap), 48 log points wage gap (about 54 per cent of the wage gap) is accounted by better 'prices' paid to the public sector workers.

Moving beyond the mean, we find that although covariate effect explains a part of the wage gap, it is the coefficient effect that dominates the covariate effect throughout the distribution except at the tails of the distribution. Unlike urban India, the wage gap explained by covariates remains the same until the 50th percentile (although the share of covariate effect in total wage gap changes as total wage gap is very heterogeneous). After the 50th percentile, there is an increase in the wage gap explained by the covariates. In contrast to the urban India where the coefficient effect decreases at higher quantiles, in rural India, the coefficient effect increases in the lower half of the wage distribution and decreases in upper half of the wage distribution.[24]

Overall, the findings suggest that although public sector workers are better endowed, this only fails to explain the entire observed wage gap between public and private sector workers. For a worker with similar characteristics, public sector pays a higher wage, i.e. a wage premium. Although this wage premium is smaller at higher quantiles, it remains positive throughout the distribution. Hence, the perception regarding public sector workers being paid less than private sector workers (after controlling for characteristics) at the top of the distribution is misplaced as the empirical evidences do not corroborate.[25]

5. Conclusion

In this paper, we examine public–private wage differential among male workers in regular-salaried jobs in India across the entire wage distribution using the nationally representative NSS data collected in 2009–10. The existing evidences about public–private wage gap in India refer to the 1990s or earlier, and concentrate on averages. The existing evidences suggest that the average wage in public sector is higher than private sector in both urban and rural areas. However, there has been an increasing concern among the public sector employees in India, especially those at the top end of the wage distribution that they are paid less than the similarly qualified counterparts in the private sector.

By examining the wage differences across the entire distribution, we find that the wage gap between public and private sector (male) workers is positive across the entire wage distribution in both urban and rural India. However, the observed wage gap is not entirely explained by differences in characteristics of workers (covariate effect) across most of the wage distribution. The public sector workers enjoy a positive wage premium across the entire wage distribution, and contribution of the wage premium (coefficient effect) in total wage gap is larger than the contribution of differences in workers' characteristics in most part of the distribution except at the top quantiles.

Our findings do not corroborate the perception of the government employees that they are underpaid compared with comparable private sector employees, especially at top end of the distribution. The findings also question the rationale for adjusting the public sector wages upwards. Any upward revision of wages in public sector not only will lead to more wage gap compared with private sector but also will have a substantial impact on fiscal balance and inflation.

There are several qualifications and caveats to our results. First, although the survey asks the wages-in-kind, it is possible that the survey will have missed other benefits. Some of the private sector employees may have additional perquisites, but consensus is that it is generally less than the public sector (Glinskaya and Lokshin, 2007). In the absence of hard evidences, one can only speculate that the public–private wage differential may be larger once we quantify the perquisites. Second, it is extremely difficult to quantify the job security and status associated with public sector jobs. Job security would unambiguously add a lot to the benefits of public sector. Moreover, research suggests that employees of public and nonprofit entities derive satisfaction from contributing to a social cause (Borzaga and Tortia, 2006; DeSantis and Durst, 1996), which is not captured by the wages.

Appendix

Table A1. Selection in wage work (regular salaried and casual work)

		Male		Female	
		Rural	**Urban**	**Rural**	**Urban**
Wages reported	Regular salaried	11,519	18,283	2,459	4,197
	Casual workers	18,154	7,784	6,691	1,838
Wages not reported	Self-employed	40,393	19,835	14,510	3,836
	Non-worker	9,685	9,108	56,228	42,911
	Total sample size	79,751	55,010	79,888	52,782

Note: Age restricted to age 18–60.

Table A2. Selection into public and private enterprise

	Male				Female			
	Rural		**Urban**		**Rural**		**Urban**	
	Regular salaried	**Casual workers**	**Regular salaried**	**Casual workers**	**Regular salaried**	**Casual workers**	**Regular salaried**	**Casual workers**
Private	4,736	10,145	10,416	6,586	800	1,485	2,294	1,184
Public	6,276	694	7,646	281	1,515	363	1,858	62
No information	507	7,315	221	917	144	4,843	45	592
Total Sample	11,519	18,154	18,283	7,784	2,459	6,691	4,197	1,838

Note: Age restricted to age 18–60.

Table A3. Quantile regression (QR), public sector

Method	Urban Mean OLS	Urban Q10 QR	Urban Q25 QR	Urban Q50 QR	Urban Q75 QR	Urban Q90 QR	Rural Mean OLS	Rural Q10 QR	Rural Q25 QR	Rural Q50 QR	Rural Q75 QR	Rural Q90 QR
Dependent variable: Log of wage												
Age	0.054***	0.058**	0.086***	0.058***	0.048***	0.023**	0.055***	0.087***	0.090***	0.028***	0.032***	0.041***
	(0.008)	(0.029)	(0.011)	(0.016)	(0.009)	(0.010)	(0.008)	(0.022)	(0.015)	(0.008)	(0.008)	(0.011)
Age Square/100	−0.034***	−0.034	−0.064***	−0.040**	−0.033***	−0.011	−0.040***	−0.072***	−0.069***	−0.011	−0.016*	−0.028**
	(0.009)	(0.033)	(0.014)	(0.018)	(0.010)	(0.012)	(0.008)	(0.025)	(0.016)	(0.008)	(0.009)	(0.012)
Currently not married	0.046	−0.077	0.070	0.033	0.054	−0.019	−0.186***	−0.427***	−0.388***	−0.210*	−0.094**	−0.028
	(0.033)	(0.137)	(0.049)	(0.069)	(0.039)	(0.048)	(0.031)	(0.107)	(0.081)	(0.113)	(0.043)	(0.050)
Scheduled Castes/Tribes	−0.075***	−0.233***	−0.077***	−0.041	−0.024	−0.052**	0.007	0.084	0.124**	0.013	0.015	−0.013
	(0.022)	(0.090)	(0.026)	(0.035)	(0.023)	(0.025)	(0.024)	(0.094)	(0.049)	(0.039)	(0.033)	(0.050)
Other Backward Castes	−0.021	−0.138**	−0.024	0.019	0.042**	−0.008	−0.007	0.058	0.055**	−0.026	0.002	−0.055
	(0.021)	(0.068)	(0.032)	(0.030)	(0.020)	(0.025)	(0.019)	(0.057)	(0.022)	(0.026)	(0.023)	(0.040)
Muslim	−0.029	−0.168	−0.032	−0.038	−0.012	−0.029	−0.122***	−0.160**	−0.126	−0.102***	−0.060	−0.154***
	(0.033)	(0.165)	(0.085)	(0.054)	(0.032)	(0.031)	(0.030)	(0.069)	(0.081)	(0.023)	(0.043)	(0.040)
Primary school	0.127***	0.201	−0.004	0.011	0.202***	0.134*	0.043	−0.024	0.132	0.024	−0.101	−0.023
	(0.048)	(0.139)	(0.157)	(0.158)	(0.075)	(0.075)	(0.050)	(0.091)	(0.130)	(0.105)	(0.078)	(0.112)
Middle school	0.179***	0.363***	0.154	0.202*	0.206***	0.135***	0.284***	0.175*	0.435***	0.301***	0.144**	0.171***
	(0.042)	(0.133)	(0.109)	(0.120)	(0.073)	(0.050)	(0.045)	(0.107)	(0.138)	(0.064)	(0.073)	(0.063)
Secondary school	0.466***	0.742***	0.475***	0.404***	0.383***	0.280***	0.378***	0.186	0.591***	0.372***	0.263***	0.216***
	(0.041)	(0.107)	(0.072)	(0.113)	(0.074)	(0.051)	(0.047)	(0.135)	(0.137)	(0.073)	(0.077)	(0.066)
Higher Secondary	0.599***	1.012***	0.690***	0.496***	0.433***	0.291***	0.550***	0.432***	0.737***	0.586***	0.446***	0.459***
	(0.044)	(0.147)	(0.078)	(0.114)	(0.077)	(0.055)	(0.050)	(0.137)	(0.156)	(0.076)	(0.095)	(0.082)
Diploma/certificate course	0.699***	1.156***	0.884***	0.595***	0.474***	0.353***	0.615***	0.627***	0.848***	0.621***	0.515***	0.479***
	(0.055)	(0.186)	(0.094)	(0.125)	(0.082)	(0.061)	(0.062)	(0.176)	(0.151)	(0.103)	(0.096)	(0.110)
Graduate, bachelor	0.742***	1.108***	0.799***	0.648***	0.608***	0.451***	0.679***	0.638***	0.903***	0.675***	0.535***	0.514***
	(0.045)	(0.133)	(0.074)	(0.115)	(0.077)	(0.057)	(0.053)	(0.167)	(0.145)	(0.084)	(0.089)	(0.095)
Postgraduate and above	0.775***	0.975***	0.923***	0.742***	0.608***	0.424***	0.794***	0.738***	1.000***	0.808***	0.682***	0.711***
	(0.052)	(0.232)	(0.074)	(0.122)	(0.082)	(0.060)	(0.059)	(0.176)	(0.155)	(0.088)	(0.097)	(0.113)
IMR (inverse mills ratio)	−0.058	−0.277	−0.150	0.025	0.106	0.008	0.090	0.039	0.446**	0.144	0.103	0.098
	(0.056)	(0.447)	(0.227)	(0.202)	(0.133)	(0.105)	(0.057)	(0.350)	(0.210)	(0.196)	(0.165)	(0.225)
IMR square		0.050	0.016	−0.039	−0.024	−0.024		0.063	−0.040	−0.066	−0.025	−0.017
		(0.080)	(0.055)	(0.066)	(0.043)	(0.051)		(0.202)	(0.112)	(0.091)	(0.067)	(0.033)
Observations	5,609	5,609	5,609	5,609	5,609	5,609	5,739	5,739	5,739	5,739	5,739	5,739
F-test (IMR = IMR square = 0)	2.498	1.913	4.537	0.621	0.196	0.118	1.064	0.428	1.115	0.178	0.630	0.546
Prob > *F*	0.114	0.148	0.0107	0.538	0.822	0.889	0.302	0.652	0.328	0.837	0.533	0.579

Note: The model also includes dummies for occupation, industry, and states. Standard errors in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. OLS, ordinary least squares; QR, Quantile Regression.

Table A4. Quantile regression (QR), private sector

	Dependent variable: Log of wage											
	Urban						Rural					
Method	Mean OLS	Q10 QR	Q25 QR	Q50 QR	Q75 QR	Q90 QR	Mean OLS	Q10 QR	Q25 QR	Q50 QR	Q75 QR	Q90 QR
Age	0.047***	0.052***	0.046***	0.052***	0.043***	0.029***	0.041***	0.092***	0.048***	0.027***	0.019**	–0.002
	(0.006)	(0.008)	(0.009)	(0.008)	(0.009)	(0.011)	(0.007)	(0.024)	(0.008)	(0.007)	(0.009)	(0.017)
Age Square/100	–0.050***	–0.059***	–0.049***	–0.062***	–0.045***	–0.024*	–0.044***	–0.113***	–0.050***	–0.027***	–0.021	0.012
	(0.008)	(0.010)	(0.012)	(0.011)	(0.013)	(0.014)	(0.009)	(0.036)	(0.012)	(0.009)	(0.014)	(0.023)
Currently not married	–0.121***	–0.171***	–0.167***	–0.152***	–0.095***	–0.102***	–0.135***	–0.103*	–0.136***	–0.134***	–0.152***	–0.168***
	(0.022)	(0.037)	(0.033)	(0.033)	(0.031)	(0.031)	(0.027)	(0.059)	(0.036)	(0.026)	(0.032)	(0.058)
Scheduled Castes/Tribes	–0.201***	–0.132***	–0.125***	–0.205***	–0.219***	–0.238***	–0.075***	–0.133***	–0.104***	–0.110***	–0.061*	–0.107
	(0.023)	(0.031)	(0.029)	(0.027)	(0.027)	(0.037)	(0.025)	(0.044)	(0.033)	(0.024)	(0.034)	(0.065)
Other Backward Castes	–0.180***	–0.159***	–0.120***	–0.179***	–0.208***	–0.239***	–0.098***	–0.193***	–0.098***	–0.051**	–0.103***	–0.126***
	(0.018)	(0.026)	(0.026)	(0.026)	(0.026)	(0.027)	(0.023)	(0.053)	(0.025)	(0.021)	(0.029)	(0.047)
Muslim	–0.082***	–0.040	–0.027	–0.094***	–0.109***	–0.111***	–0.023	–0.084	–0.013	–0.019	–0.023	0.038
	(0.025)	(0.040)	(0.024)	(0.026)	(0.025)	(0.025)	(0.030)	(0.114)	(0.022)	(0.032)	(0.037)	(0.065)
Primary school	0.004	0.137***	0.050*	–0.050	–0.083**	–0.038	–0.012	0.099	0.110***	0.040	0.011	–0.093
	(0.034)	(0.035)	(0.030)	(0.052)	(0.037)	(0.048)	(0.034)	(0.062)	(0.025)	(0.040)	(0.042)	(0.069)
Middle school	0.050*	0.163***	0.087**	0.015	–0.059**	–0.029	0.108***	0.002	0.136***	0.109***	0.096**	0.045
	(0.030)	(0.033)	(0.035)	(0.040)	(0.029)	(0.038)	(0.031)	(0.069)	(0.038)	(0.034)	(0.040)	(0.069)
Secondary school	0.202***	0.253***	0.233***	0.146***	0.116***	0.172***	0.194***	0.070	0.220***	0.191***	0.181***	0.151**
	(0.031)	(0.037)	(0.036)	(0.041)	(0.031)	(0.045)	(0.034)	(0.072)	(0.046)	(0.039)	(0.042)	(0.073)
Higher Secondary	0.291***	0.389***	0.309***	0.157***	0.197***	0.321***	0.141***	0.044	0.145***	0.105	0.133**	0.230***
	(0.035)	(0.045)	(0.040)	(0.048)	(0.052)	(0.069)	(0.041)	(0.094)	(0.047)	(0.066)	(0.054)	(0.084)
Diploma/certificate course	0.502***	0.393***	0.528***	0.405***	0.388***	0.695***	0.250***	0.299**	0.287***	0.281***	0.145**	0.195
	(0.046)	(0.139)	(0.081)	(0.137)	(0.092)	(0.066)	(0.058)	(0.128)	(0.091)	(0.062)	(0.073)	(0.219)
Graduate, Bachelor	0.621***	0.511***	0.525***	0.593***	0.647***	0.738***	0.356***	0.092	0.333***	0.372***	0.309***	0.487***
	(0.037)	(0.056)	(0.058)	(0.051)	(0.054)	(0.051)	(0.044)	(0.114)	(0.064)	(0.064)	(0.076)	(0.122)
Postgraduate and above	0.836***	0.747***	0.767***	0.747***	0.852***	0.928***	0.704***	0.253	0.517***	0.697**	0.940***	0.757***
	(0.046)	(0.115)	(0.078)	(0.079)	(0.053)	(0.061)	(0.061)	(0.178)	(0.122)	(0.307)	(0.092)	(0.126)
Observations	6,885	6,885	6,885	6,885	6,885	6,885	4,227	4,227	4,227	4,227	4,227	4,227

Note: The model also includes dummies for occupation, industry, and states. Standard errors in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.
OLS, ordinary least squares; QR, Quantile Regression.

Table A5. Decomposition of public private wage gap, other counterfactual

Urban				
Quantile Mean	**Actual difference**	**Coefficient effect**	**Covariate effect**	**Residual**
(OLS)	0.847	0.367	0.479	0.000
		0.327 0.408	*0.440 0.519*	
10	0.760	0.182	0.685	–0.107
		0.142 0.221	*0.642 0.730*	
25	1.070	0.438	0.587	0.045
		0.413 0.464	*0.561 0.614*	
50	1.030	0.571	0.448	0.011
		0.549 0.593	*0.427 0.466*	
75	0.821	0.500	0.338	–0.017
		0.468 0.528	*0.317 0.359*	
90	0.494	0.269	0.266	–0.041
		0.233 0.309	*0.240 0.291*	

Rural				
Quantile Mean	**Actual Difference**	**Coefficient Effect**	**Covariate Effect**	**Residual**
(OLS)	0.895	0.359	0.535	0.000
		0.302 0.416	0.481 0.590	
10	0.502	–0.139	0.772	–0.130
		–0.177 –0.094	*0.723 0.811*	
25	0.777	0.165	0.744	–0.133
		0.141 0.193	*0.714 0.774*	
50	1.163	0.485	0.594	0.084
		0.465 0.506	*0.571 0.617*	
75	1.064	0.669	0.404	–0.009
		0.646 0.691	*0.383 0.425*	
90	0.862	0.616	0.250	–0.005
		0.583 0.651	*0.225 0.273*	

Note: The actual difference is difference between actual empirical (log) wage distributions of public and private sector at each quantile, i.e. $\alpha\{f(w(Pub))\} - \alpha\{f(w(Pvt))\}$, where $\alpha \in (0,1)$. Covariate effect refers to difference in $\alpha\{f^*(w(Pub))\} - \alpha\{f^*(w(Pub); x(Pvt))\}$, Coefficient effect refer to $\alpha\{f^*(w(Pub); x(Pvt))\} - \alpha\{f^*(w(Pvt))\}$. See text for further details. The decomposition at mean is derived using the Oaxaca Blinder decomposition using ordinary least squares (OLS).

Notes

[1] A large number of memoranda to Sixth Pay Commission (a panel set up by the Union Cabinet of India for revising the salaries of central government's employees), particularly those pertaining to group A employees (higher skill jobs), have mentioned the disparities between the private sector salaries and salaries in the government, citing this as a reason for the reduced attractiveness of the government jobs as a career option and for the decline in the quality of intake (Government of India, 2008).

[2] The Sixth Central Pay Commission was set up by the Union Cabinet of India on October 5, 2006 for revising the salaries of central government's employees with a time frame of 18 months. The implementation led to an additional Rs. 200,000 million (about 4,400 million USD) burden on the central government.

[3] It is undeniable that public sector jobs provide unparalleled job security, pension benefits, work-life balance, and status. The prestige involved in working for the government and the opportunity of making a contribution to national policy or its implementation are other aspects which add an unquantifiable value to government jobs (Government of India, 2008).

[4] World Development Indicators, The World Bank.

[5] See Government of India (2012) for details.

[6] National Commission for Enterprises in the Unorganised Sector has recommended discontinuation of the system of approximating the organized sector employment with the estimates of DGET and to use NSS employment–unemployment surveys to estimate the same (Government of India, 2009, p. 11).

[7] Source: DGET.

[8] Source: DGET, Government of India, 2011a, p. A52.

[9] To take account of household survey weights, unequal probability sampling with replacement is implemented. We end up with 999,000 observations for w^*.

[10] There is another possible counterfactual that can be used in the decomposition, i.e. $f^*(w(Pub); x(Pvt))$ the density that would have prevailed if private sector workers were paid the 'prices' prevailing in the public sector. In this case, covariate effect will be $\alpha\{f^*(w(Pub))\} - a\{f^*(w(Pub); x(Pvt))\}$ and coefficient effect will be $\alpha\{f^*(w(Pub); x(Pvt))\} - a\{f^*(w(Pvt))\}$. It is well known that the decomposition results may not be invariant with respect to the choice of the counterfactual. Therefore, the choice of a counterfactual should be guided by the question of economic interest. Since wages in public sector is set by the government, and our interest is to find out the wage premium enjoyed by the public sector, we use the 'prices' paid to private sector as reference as those 'prices' are determined in the market.

[11] The residual term comprises the simulation errors which disappears with more simulations, the sampling errors which disappears with more observations, and the specification error induced by estimating linear quantile regression (Melly, 2005).

[12] NSSO conducts large-scale nationally representative survey at 5-year intervals called as quinquennial rounds.

[13] The public sector is identified based on type of enterprise question in NSS data. The survey asks all the nonagricultural workers to categorize their place of main work activity in the following categories: proprietary male, proprietary female, partnership with members from same household, partnership with members from different household, government/public sector, public/private limited company, cooperative societies/trust/other nonprofit institutions, employer's households(i.e. private households employing maid servant, watchman, cook, etc.), and others. Government/public sector was treated as public sector, whereas the others are grouped into private sector.

[14] Labor Force Participation Rate (LFPR) for women remains low in India. For example, in 2009–10, for age group 15–59, LFPR was 85 per cent for rural male, 40 per cent for rural female, 81 per cent for urban male, and 21 per cent for urban female (Government of India, 2011b). As shown in Table A2, the sample sizes for women is small to capture the heterogeneity across India, which is the main objective of the paper.

[15] A significant proportion of private regular-salaried workers (43 and 56 per cent in urban and rural areas, respectively) work in enterprises that employ less than 10 employees, and according to DGET definition of organized sector, they are excluded. Note that the public sector workers are always included in organized sector according to DGET definition. About 12 (37) per cent of the public sector regular-salaried workers work in enterprises that employ less than 10 employees in urban (rural) areas. There has been a considerable ongoing debate about how to define organized sector (see Government of India, 2009). National Commission for Enterprises in the Unorganised Sector use alternative definition for formal/organized workers and according to their estimates about 62.6 million workers were in organized sector in 2004–05 (as opposed to the DGET estimates of 26.6 million workers in 2006) (Government of India, 2009, pp. 11–13). Organized work force in Indian case needs to be understood as those workers who have regular, contractual hired employment, and enjoy a relatively high rate of wages, which are sufficient to provide social security, emanating from sustained productivity per worker (Tendulkar, 2003).

[16] Tables A1 and Table A2 provide a detailed successive steps to arrive at the final sample, i.e. male regular-salaried workers in age group 18–60.

[17] The issue of correction for sample selection although settled for mean regression, is still evolving for quantile regression. To our best knowledge, the sample correction in quantile is limited to a dichotomous variable only as proposed in Buchinsky (1998). Huber and Melly (2015) warn against the danger of importing one-to-one mean recipes to quantile. Moreover, they also finds that the estimator developed in Buchinsky (1998) is only consistent when the slope coefficients are equal for all quantiles or when selection is randomly determined.

[18] We do not claim that the selection mechanism in regular-salaried job is same across public or private sector regular-salaried jobs. In the case of different selection mechanism, our results will be biased. As we carry out our analysis separately for urban and rural areas, there is a possibility that the workers self-select into urban and rural areas according to the unobserved factors. We thank an anonymous referee for pointing this out. As we only know the current area of residence, it is not possible to determine how many workers moved across areas and controlling for that. This will potentially bias the result if the mechanism of urban/rural choice differs across public regular and private regular employment.

[19] We use owned land size and household size as exclusion variables for rural areas, and only household size for urban areas.

[20] Insignificant selection terms at selected important quantiles do not guarantee insignificant selection at each percentile. However, it supports that the wage differentials are less likely to be driven by selection between public private conditional on regular-salaried employment if there is any. Albrecht *et al.* (2009) incorporate quantile regression selection based on Buchinsky (1998) into MM technique. However, including sample correction in MM technique leads to an added complication that the constant term in the quantile wage regression is not identified as it is conflated with the constant term of the higher order power series used to control for sample selection. Along the lines suggested by Andrews and Schafgans (1998), one can identify a subsample of workers whose observable characteristics are such that their probability of working in public sector is arbitrarily close to one, this subsample can be used to estimate the intercept term without adjusting for sample selection. However, identifying such subsample of workers in our data seems extremely difficult.

[21] The NSSO data report a wage total variable that is basically sum of wages in cash and kind, and wages reported/used in most of the literature is the total wages.

[22] The survey asks a question 'How much you earn during the last week?'

[23] In case of overprediction of gap, the coefficient and covariate effect are marginally overestimated.

[24] If one were to assume that the wage ranking of workers follow roughly the educational ranking, public sector workers have better educational distribution in both urban and rural areas (Table 1). As we have used private sector prices as reference, the premium paid increases with education levels in urban areas (Table A4). This potentially explains higher covariate effect at the higher quantiles in urban areas. Similarly, the premium paid in public sector for middle levels of education (middle, secondary, higher secondary) is larger than the premium paid in private sector in urban areas (Tables A3 and A4), this potentially explain larger coefficient effect in middle of the distribution compared with top part of distribution. In rural areas, premium paid in private sector for middle or higher secondary levels are not very different (Table A4). This potentially explains similar covariate effect in the lower half of the distribution. As college level premium is larger in private sector, the covariate effect is larger in upper half of the distribution. Similarly, the premium paid to middle, secondary, and higher secondary levels in public sector is higher than the premium paid in private sector in rural areas. This potentially explains a larger coefficient effect in rural areas in middle part of the distribution.

[25] In Table A5, we also report the results using the other counterfactual, i.e. $f^*(w(Pub); x(Pvt))$, the wage density that would have prevailed if private sector workers were paid the 'prices' prevailing in the public sector. Although our main conclusion of positive wage premium throughout the distribution remains valid, the magnitude of the coefficient and the covariate effect are not invariant to choice of the counterfactual. For example, OB decomposition suggests that covariate effect is larger than the coefficient effect at the mean. This is not surprising given the difference in characteristics across the two sectors. In the case of OB decomposition, when we use counterfactual $f^*(w(Pvt); x(Pub))$ ($f^*(w(Pub); x(Pvt))$), the coefficient effect is captured by $\bar{x}^{Pub}\left[\hat{\beta}^{Pub} - \hat{\beta}^{Pvt}\right]$ $\left(\bar{x}^{Pvt}\left[\hat{\beta}^{pub} - \hat{\beta}^{pvt}\right]\right)$. Hence, the same difference in prices is evaluated at two different set of characteristics. Similarly, the covariate effect for

counterfactual $f^*(w(Pvt); x(Pub))$ $(f^*(w(Pub); x(Pvt)))$ is captured by $\hat{\beta}^{Pvt}(\bar{x}^{Pub} - \bar{x}^{Pvt})$ $\left(\hat{\beta}^{Pub}(\bar{x}^{Pub} - \bar{x}^{Pvt})\right)$, i.e. the same difference in characteristics is evaluated at two different prices. This issue is well known in the literature. For example, Fortin *et al.* (2011) show that the gender gap decomposition is substantially different when either the female wage structure or the weighted sum of the male and female wage structure is used as the reference wage structure (table 3 of Fortin *et al.*). To avoid this arbitrary dependence in the traditional OB decomposition, few studies use a nondiscriminatory (in case of male/female or black/white differences) structure estimated using a pooled regression. However, in our case, we believe that private sector prices are the true reference as it is determined by the market forces. Importantly, the conclusion of positive wage premium enjoyed by the public sector workers will stand even if we use a weighted average of the prices prevailing in both sectors. Moreover, with public sector prices as counterfactual, the wage premium enjoyed by public sector workers is much larger at the top quantiles.

References

Albrecht J., Bjorklund A. and Vroman S. (2003) 'Is There a Glass Ceiling in Sweden?', *Journal of Labor Economics* 21: 145–177.

Albrecht J., Vuuren A. and Vroman S. (2009) 'Counterfactual Distributions with Sample Selection Adjustments: Econometric Theory and an Application to the Netherlands', *Labour Economics* 16(4): 383–396.

Andrews D. W. K. and Schafgans M. M. A. (1998) 'Semiparametric Estimation of the Intercept of a Sample Selection Model', *Review of Economic Studies* 65: 497–518.

Bender K. A. (1998) 'The Central Government-Private Sector Wage Differential', *Journal of Economic Surveys* 12: 177–220.

Borzaga C. and Tortia E. (2006) 'Worker Motivations, Job Satisfaction, and Loyalty in Public and Nonprofit Social Services', *Nonprofit and Voluntary Sector Quarterly* 35(2): 225–248.

Buchinsky M. (1998) 'The Dynamics of Changes in the Female Wage Distribution in the USA: A Quantile Regression Approach', *Journal of Applied Econometrics* 13: 1–30.

Depalo D., Giordano R. and Papapetrou E. (2015) 'Public–Private Wage Differentials in Euro-Area Countries: Evidence from Quantile Decomposition Analysis', *Empirical Economics* 49(3): 985–1015.

DeSantis V. S. and Durst S. L. (1996) 'Comparing Job Satisfaction among Public and Private Sector Employees', *American Review of Public Administration* 26(3): 327–343.

Disney R. and Gosling A. (1998) 'Does It Pay to Work in the Public Sector?', *Fiscal Studies, Institute for Fiscal Studies* 19(4): 347–374.

Duraiswamy P. and Duraiswamy M. (1995) 'Implications for Structural Reforms Public–Private Sector Wage Differential in India', *Indian Journal of Labour Economics* 38(4): 597–602.

Dutta P. V. (2006) 'Returns to Education: New Evidence for India, 1983–1999', *Education Economics* 14: 431–451.

Elliott R. F. and Duffus K. (1996) 'What Has Been Happening to Pay in the Public Sector of the British Economy? Developments over the Period 1970–1992', *British Journal of Industrial Relations* 34: 51–85.

Fiorio C. V. (2004) 'Confidence Intervals for Kernel Density Estimation', *The Stata Journal* 4(2): 168–179.

Fortin N., Lemieux T. and Firpo S. (2011) 'Decomposition Methods in Economics' chapter 1 in Ashenfelter O. and Card D. (eds.) *Handbook of Labor Economics*, Vol. 4, A, Elsevier: 1–772.

Giordano R. *et al.* (2011) 'The Public Sector Pay Gap in a Selection of Euro Area Countries'. Working Paper Series 1406, European Central Bank.

Glinskaya E. and Lokshin M. (2007) 'Wage Differentials between the Public and Private Sectors in India', *Journal of International Development* 19: 333–355.

Government of India. (2008) 'Report of Sixth Pay Commission'.

Government of India. (2009) 'The Challenge of Employment in India: An Informal Economy Perspective', National Commission for Enterprises in the Unorganized Sector.

Government of India. (2011a) 'Economic Survey 2010–11'.

Government of India. (2011b) 'Key Indicators of Employment and Unemployment in India, 2009–10'.

Government of India. (2012) 'Manual on Labour Statistics (I)-2012', Ministry of Statistics and Programme Implementation.

Government of India. (2013) 'Key Indicators of Employment and Unemployment in India, 2011–12'.

Gregory R. and Borland J. (1999) 'Recent Developments in Public Sector Labor Markets' in Ashenfelter O. and Card D. (eds.) *Handbook of Labor Economics*, Vol. 3c, Amsterdam: North-Holland: 3573–3630.

Huber B. and Melly B. (2015) 'A Test of the Conditional Independence Assumption in Sample Selection Models', *Journal of Applied Econometrics*, doi: 10.1002/jae.2431.

Katz L. and Krueger A. (1993) 'Public Sector Pay Flexibility: Labor Market and Budgetary Considerations', in *Pay Flexibility in the Public Sector*. Paris: OECD: 43–77.

Klein R. W. and Spady R. H. (1993) 'An Efficient Semiparametric Estimator for Binary Response Models', *Econometrica: Journal of the Econometric Society* 61: 387–421.

Lakshmanasamy T. and Ramasamy S. (1999) 'An Econometric Analysis of the Worker Choice between Public and Private Sector', *Indian Journal of Labour Economics* 42(1): 71–83.

Lausev J. (2014) 'What Has 20 Years of Public–Private Pay Gap Literature Told Us? Eastern European Transitioning versus Developed Economies', *Journal of Economic Surveys* 28(3): 516–550.

Machado J. and Mata J. (2005) 'Counterfactual Decomposition of Changes in Wage Distributions using Quantile Regression', *Journal of Applied Econometrics* 20: 445–465.

Madheswaran S. (1998) 'Earning Differentials between Public and Private Sector in India', *Indian Journal of Labour Economics* 41(4): 945–970.

Madheswaran S. and Shroff S. (2000) 'Education, Employment and Earnings for Scientific and Technical Workforce in India', *Indian Journal of Labour Economics* 43(1): 121–137.

Melly B. (2005) 'Public–Private Sector Wage Differentials in Germany: Evidence from Quantile Regression', *Empirical Economics* 30: 505–520.

Mueller R. E. (1998) 'Public–Private Sector Wage Differentials in Canada: Evidence from Quantile Regressions', *Economics Letters* 60(2): 229–235.

Nielsen H. and Rosholm M. (2001) 'The Public–Private Sector Wage Gap in Zambia? A Quantile Regression Approach', *Empirical Economics* 26: 169–182.

Schager N. (1993) 'An Overview and Evaluation of Flexible Pay Policies in the Swedish Public Sector', in *Pay Flexibility in Public Sector*. Paris: OECD: 113–124.

Smith S. (1976) 'Pay Differentials between Federal Government and Private Sectors Workers', *Industrial and Labour Relations Review* 29: 233–257.

Smith S. (1977) 'Equal Pay in the Public Sector: Fact or Fantasy', *Industrial Relations Section*, Princeton.

Tendulkar S. D. (2003) 'Organised Labour Market in India Pre and Post Reform', Paper prepared for presentation at the conference on 'Anti poverty and Social Policy in India', January 2–4, 2004.