

# Wage Differentials Between The Public And The Private Sector: How Comparable Are The Workers?

Josefa Ramoni-Perazzi, (Email: jramoni@ula.ve), Universidad de los Andes  
Don Bellante, (Email: dbellant@usf.edu), University of South Florida

## ABSTRACT

*This paper's purpose is to examine the importance of controlling for the comparability of the units in a sample when examining wage differentials of various types. We use U.S. CPS data for the period 1992-2000 to compare estimated wage differentials between public and private sector workers obtained using two different methodologies: Lee's two-step method that controls for the selection bias resulting from the non-randomness of the sample, and the propensity score matching method that controls also for the comparability of the workers. Lee's method suggests that federal workers are paid a premium, while state and local workers are underpaid compared to private sector workers. However the matching method indicates that this data is too heterogeneous to be used to compare wages across sectors. We conclude that, when the outcome under study is not only affected by some sort of selection but also requires comparable groups, the traditional methodology may not be enough.*

## INTRODUCTION

Since the 1970s, many studies have attempted to compare wages between the public and private sector in the United States, under the premise that similar workers should be paid similar wages, regardless the sector of employment. The majority of these studies are based on simple human capital models, expanded to account for some productivity related characteristics of the workers, although more recent papers have paid more attention to the selection bias problem. The general literature usually concludes that there exist a positive wage premium, or rent, paid to public sector workers, that increases with the level of government. Such a rent has been attributed to either the differences in the structure of the wage-setting system between the two sectors, or as the result of methodological problems.

That the wage-setting systems differ in both sectors is obvious. The public sector is not a profit maximizing agent, and therefore its wages are not determined by the forces driving the labor market as happens in the private sector. Besides, the public sector is subject to political influences, which could lead the government to pay wages higher than those paid to similar workers in the private sector. It is because this lack of information provided by the market that the public sector applies the prevailing wage principle that states that similar workers, performing similar jobs, should be paid similar wages, regardless the sector of employment. However, since not all jobs in the public sector have a similar counterpart in the private sector, the comparison should examine comparable workers. But the use of the wrong methodology to approach the problem may provide misleading results. It is known that since workers differ in their skills and taste for job attributes, they will self-allocate into the job that maximizes their utility. The resulting nonrandom allocation of workers among sectors may generate selection bias problems. In cases like this, ruling out possible sources of endogeneity that may affect the outcome is not always enough. When the analysis is based on comparing units that are assumed to be similar, one should also examine the comparability of the so considered "comparable groups." Comparability implies that workers have similar human capital, as well as other productivity-related characteristics so that they have the same likelihood of being employed in a specific sector. Using data from the Current Population Survey (CPS) March supplement for the period 1992-2000, we compare the results of models where only the endogeneity of the sector of employment is accounted for with those obtained when the

comparability of the groups also is considered, based on propensity score matching methods. The results indicate that ignoring the level of heterogeneity among the units in the sample under study may yield misleading results. Unfortunately, in spite of its wide use to compare wages across sectors, the CPS data set seems to lack substantial numbers of workers from both groups that are truly comparable.

## **BACKGROUND**

The first attempts to estimate wage differentials between the public and the private sector can be traced back to the 1970s with a series of studies conducted by Sharon Smith (1974a, 1974b, 1977, 1981), based on OLS estimates of human capital models. Her work was then followed by similar papers based on transformations of her models, to account for some variables considered as relevant, or disaggregating the data by occupation or by region. These studies were based on the idea that workers with similar productivity-related characteristics should be paid similar wages regardless of the sector of employment. However none of them actually controlled for how comparable were the workers in the sample. The general conclusion was that public sector workers, especially women, are overcompensated and that the wage premium increases with the level of government.

During the 1990's, the main concern of the studies at this regard shifted toward finding the best methodological approach to the public/private wage comparison, emphasizing the reduction of any potential source of bias. Of particular importance was the selectivity bias resulting from the nonrandomness of the sample. Most of the studies correcting for selection bias in comparing wages across sectors were based on two-step models that control for the effect of one or more of the following decisions on wages: union status, labor force participation and sector of employment [Gyourko and Tracy, (1988); Belman and Heywood, (1989); Venti, (1989); Hundley, (1991); Choudhury, (1994); Hoffnar and Greene, (1996)]. In general, these studies support the idea of a premium paid to public sector workers, but show inconclusive results regarding the effect of selection bias on the wage premium.

The two-step-type methodology these studies apply may correct for the potential bias resulting from the nonrandomness of the sample, so that the estimated parameters would not be affected by the probability of being in the sample. But this methodology does not ensure that the workers are similar, as the comparison requires. A more recent approach to dealing with selection bias in nonrandomized observational studies is based on so-called matching methods. These methods allow the investigator to limit the comparison to paired individuals on the basis of a set of observables or pre-treatment characteristics, reducing the bias and generating matched control groups that can be used as "virtual" counterfactuals. The most common matching approach bases the comparison on the propensity score or conditional probability that the individual is in the treated group, which reduces the k-dimensional comparison to one single variable [Rubin (1973); Rosenbaum and Rubin, (1983)]. In the labor economics field, matching methods have basically been used to measure the impact of job-training programs on wages [LaLonde (1986); Ichimura (1993); Heckman et. al. (1997); Dehejia and Wahba (1999, 2002); Kordas and Lehrer (2002); Abadie and Imbens (2002)], while their potential use to compare wages across sectors has not been yet exploited. Ramoni (2004) estimated the propensity score as given by the probability of choosing to work in a specific sector, and used this score to select homogeneous matched subsamples on which to base compensation comparisons among sectors in the United States.

## **DATA**

This study uses the information provided by the Current Population Survey (CPS) March Supplement, which includes individual-level information on labor force status and other demographic variables for individuals working in the federal, state, and local governments as well as in private industry.

We considered five waves for the period 1992-2000<sup>1</sup> pooled into a single data set in order to gain efficiency. The final sample includes individuals 16 years of age and older, and excludes those individuals working in the agricultural sector, those self-employed or in the armed forces as well as those working less than 10 or more than 80 hours a week or who earned less than half of the minimum hourly wage in effect at the time of the interview. The final data set includes 52,792 individuals, 86.54% of them working in the private sector and the remaining 13.46% employed in the public sector. Among those in the public sector, 22.62% work at the federal level, 25.39% work at the state level, and 51.99% work for local governments. This disaggregation of the information by level of government

constitutes the one of the most important feature of this data set, as well as the relatively larger sample size. This may explain why many studies comparing wages are based on the CPS. Unfortunately, the specific data set used in this study shows a very high level of heterogeneity among workers across sectors, suggesting that it may not be suitable for comparisons purposes.

**METHODOLOGY**

When the availability of the data or the sampling method is influenced by a selection process that is related to the value of the dependent variable under study, a sample selection problem may arise, introducing correlation between the error term and the regressors, thus biasing the estimates.

In the case of wage comparisons across sectors, a potential source of bias may come from the fact that individuals self-allocate into the sector that matches their preferences. Consider the following wage model

$$W_s = X_s \beta_s + \varepsilon_s \tag{1a}$$

$$W_s^* = M_s \gamma_s + \eta_s \tag{1b}$$

where  $X_s$  and  $M_s$  are vectors of observed productivity and demographic characteristics,  $s$  is a categorical decision variable that represents the sector of employment, and  $\varepsilon_s$  is an error term with mean zero and variance  $\sigma^2_s$ . The wage  $W_s$  is observed if and only if the sector  $s$  is chosen, which happens when  $W_s^* > \max_{j \neq s} (W_j^*)$ . Estimating  $\beta_s$  and  $\gamma_s$  ignoring that  $\varepsilon_s$  and  $\eta_s$  are not independent, may yield inconsistent results.

Heckman (1974) developed a way to address the sample selection problem based on a two-step procedure. This methods is implemented by first estimating  $\gamma_s$  in equation (1b) based on a logit or probit model, in order to compute the Mills ratio as given by the ratio between the standard normal density function [ $\phi(M\Gamma)$ ] and the normal distribution function [ $\Phi(M\Gamma)$ ]. This ratio is then used in the right hand side of equation (1a) to obtain consistent estimates of  $\beta_s$ . Lee (1983) generalized Heckman’s method to the case where the selection is based on a multinomial logit model. In general, these methods are seen as yielding consistent but not asymptotically efficient estimates under the normality assumption. The main drawback of these methods is that the reliability of the results is often limited to the satisfaction of a complex set of assumptions that include functional form restrictions and exclusion restrictions regarding the variables to be included in the output equation necessary for identification purposes, making the method less robust statistically, and providing only a partial correction for the selection bias issue.

To properly compare wages between the public and the private sector, one should compare the wage a worker is paid in the public sector and the wage he would be paid in the private sector. Of course it is impossible to observe both wages for anyone, making it necessary to find an adequate counterfactual to, say, public sector workers, i.e. private sector workers who are as similar as possible to those in the public sector. Matching methods are based on the idea that the selection bias is reduced when the comparison of outcomes is performed only between treated and control units who are similar. These methods therefore involve pairing treated and control units that are comparable in terms of their observable characteristics, which are assumed to be independent from the assignment to the treatment. But matching based on a  $k$ -dimensional vector of characteristics may not be feasible, especially for a large  $k$ . In the seminal paper by Rosenbaum and Rubin (1983), the authors proposed the propensity score matching as a method to reduce the bias in the estimation of treatment effects. The main feature of the propensity score is that it reduces the dimensionality of the comparison from  $k$  variables to a single value.

Let sector of employment  $D$ , be the decision variable. The treated group is given by individuals working in a particular sector, for instance the public sector ( $D=1$ ), while the control group includes the private sector workers ( $D=0$ ). The propensity score (PS) as defined by Rosenbaum and Rubin, is the conditional probability of being in the public sector, given pre-treatment characteristics  $X$ . i.e.,  $PS = Pr \{D=1 | X\} = E \{D | X\}$

If the selection of sector of employment is random according to  $X$ , it is also random according to PS. With randomization the treated and control groups do not substantially differ from each other, so that the results are no longer conditioned to the treatment.

For the PS method to represent the best possible approximation to a random allocation of individuals across sectors such that their wages can be compared, the assumption of strong ignorability of the treatment should be satisfied. This condition combines the two following assumptions:

### **Balancing Of Pre-Treatment Variables**

Given the PS, also referred to as ignorability of treatment or unconfoundedness, which ensures that the PS provides all the necessary information regarding the determinants of the decision about sector of employment, so that the outcome (wages) is independent from D. This assumption requires observations with the same PS to have the same distribution of observable (and unobservable) variables regardless of the treatment status (sector). If this is true, for a given PS the assignment to treatment (sector) is random. Therefore treated and control units should on average be identical.

### **Common Support**

The probability of assignment to treatment is bounded between one and zero, which implies that the matching should be performed over an area of common support. If there is no overlap in the PSs in both sectors, it is not possible to match individuals, and no outcome difference can be estimated with that data. Therefore this condition rules out the possibility of perfect prediction, since workers with PS equals to 1 or to 0 cannot be matched. This would mean that only workers with a similar probability of working in the public sector could be matched.

The main drawbacks of this method lie on the estimation of the PS. As Rosenbaum and Rubin showed, under the strong ignorability assumption the PS eliminates the bias, making unnecessary any assumption about functional form, separability of outcome or exogeneity of the conditioning variables, so common in control function models. However, Heckman and Navarro (2004) showed that these matching methods break down when there are variable that perfectly predict the decision, which suggest that choosing the appropriate variables to model the decision may be difficult. In fact, the propensity score matching methods do not distinguish between the variables affecting the outcome (Z) and those affecting the decision (X), as happen in many econometric procedures, including control function models. This distinction could help to overcome the perfect prediction problem, if there are some variables X not in Z. Moreover, the propensity score matching methods do not provide any rule as to which variables to include or exclude in the conditioning set.

The PS is generally estimated by parametric procedures, usually logit or probit,<sup>2</sup> on the basis of a set of conditioning variables that affect in our case the decision to work in the public sector following the general algorithm:

- Estimate the propensity score by using a logit (or probit) model starting with a parsimonious specification with linear covariates.
- Divide the observations into  $k$  equally spaced intervals of the PS, such that within each interval the difference in the PS for treated and control groups is insignificant.
- Within each interval, test the hypothesis that the average of each covariate does not differ between treated and control groups.<sup>3</sup>
- If the intervals are not balanced, re-estimate the PS by adding higher order variables and/or interactions. If there are no differences, the specification is accepted.

To emphasize the importance of controlling for the comparability of the workers among sectors when calculating wage differentials, we apply both Lee's two-step method and propensity score matching method to the CPS data set for the period considered. The selection of sector of employment is modeled as given by:

$$\text{SECTOR} = f(\text{AGE, EDUCTG, GENDER, RACE, VETERN, MARTST, CENREG}) + v \quad (3)$$

where SECTOR is a selection variable that equals one for workers in the private sector, two for workers in the federal level, three for the state level and four for the local level of government. To apply the Lee's method, equation (3) is estimated using a multinomial logit model from which to calculate hazard functions. The propensity score matching

methods, on the other hand, are based on binary decisions, so that the estimates of the multinomial logit model cannot be used.<sup>4</sup> To overcome this problem, we compare each level of government (treated group) with the private sector (control group), assuming that the probability of working in the public sector is different in each level of the government, and that these probabilities are independent from each other. In this case, SECTOR would be a dummy variable that equals one if the worker is employed in a specific level of government, and zero for private sector workers.

It is expected that individuals working for the government at any level are relatively older (AGE) and have a higher level of education (EDUCTG). More discrimination is expected in the private sector, in which case we should see a higher proportion of men (GENDER) and whites (RACE) working in this sector. The public sector tends to give more employment opportunities to veterans (VETERN). Also, assuming that public sector workers enjoy more job stability as some studies suggest, married workers (MARTST) should be more likely to work for the government. Public and private job opportunities may not be evenly spread around the country, with a greater or lesser number of public jobs in different regions (REGION). Finally,  $v$  is a stochastic error term that follows a normal distribution with zero mean and variance  $\sigma_v^2$ .

In the case of the two-step procedure, a vector of hazard functions  $\hat{m}$  is used to correct a wage equation as given by:

$$\text{LOGW} = \alpha_0 + \alpha_1\text{AGE} + \alpha_2\text{AGE}^2 + \alpha_3\text{EDUCTG} + \alpha_5\text{RACE} + \alpha_6\text{GENDER} + \alpha_7\text{MARTST} + \alpha_8\text{HEALTH} + \alpha_9\text{FULTME} + \alpha_{10}\text{T} + \xi \tag{4}$$

where LOGW is the logarithm of hourly wages; AGE is a proxy for general experience, and is expected to have a positive impact on wages ( $\alpha_1 > 0$ ) at a declining rate ( $\alpha_2 < 0$ ). The returns to formal education should also be positive. Following the literature, wages are expected to be higher for men and whites. Married workers are seen as being more responsible at work, and therefore receiving higher wages.

Health status (HEALTH) is considered here as a measure of a relevant condition affecting not just the worker's general performance, but also his tenure and therefore affecting wages. The same positive effect on wages can be expected for the full-time working condition (FULTME). Finally, T includes four dummy variables, each corresponding to a different year of the survey, assuming the year 1992 as the base, to account for potential cohort effect, while  $\varepsilon$  is a vector of stochastic errors distributed  $N(0, \sigma_\varepsilon^2)$ .<sup>5</sup>

In the matching method, on the other hand, treated units are matched following the Nearest Neighbor (NN) matching method<sup>6</sup> with replacement, so that each treated unit can be matched with several similar control units. After individuals are matched, the unmatched members of the control group are discarded. The subsamples in this way obtained, are used to estimate wage equations as given by (4) by sector of employment.

Rent is assumed to exist if the returns to the productivity-related characteristics are higher in the public sector. The wage comparison can be done considering indexes calculated by level of education as given by (Bellante and Long, 1983) based on the sum of the differences in the  $j$ th coefficient estimated in the wage equations by sector evaluated in terms of the average characteristics of the workers in the private sector ( $\bar{X}_{pj}$ ).<sup>7</sup> If these factors receive the same remuneration in all sectors, such an index should tend to zero. Otherwise, if such an index still shows that a wage differential persists, this could indicate that the returns to these factors are not the same in all sectors, and expression (5) would measure the pure sectoral differences in the payoffs in human capital and other qualities between public and private sector workers, i.e. rent.

$$\sum_{j=1}^k (\hat{\beta}_{gj} - \hat{\beta}_{pj}) \bar{X}_{pj} \tag{5}$$

**RESEARCH RESULTS**

The main statistics of the sample shown in Table 1 indicate that public sector workers are on average significantly older, especially at the federal level, compared to those in the private sector. Also as expected, a relatively higher proportion of women working for the government is observed, particularly at the state and local levels. In the same way, the participation of nonwhite workers in the private sector is relatively smaller, with a higher tendency for them to work at the federal and local levels. Veterans are more likely to work in the public sector, especially at the federal level. Married workers seem to be more likely to work for the government, especially at the federal and local levels, in line with the presumption that married workers will particularly value more stable jobs, and the public sector is perceived as offering greater job stability. Workers living in the South and West census regions are more likely to work in the federal sector.

The sample clearly shows that public sector workers tend to have higher levels of education compared to private sector workers, with a higher proportion of workers with some college or more is observed in the public sector, especially at the federal and state levels. Also, average wages are substantially higher in the public sector at any level, but especially in the federal government. While the general literature suggests a relatively higher compression of wages in the public sector [Borjas (2002); Katz and Krueger (1991); Poterba and Rueben (1994)], a smaller dispersion of wages around the mean is only observed in the state level. Only federal workers seem to work more average hours per week, than those in the private sector.

A relevant conclusion that can be obtained from Table 1 is that the CPS sample is a very heterogeneous data set, with substantial differences in mean or proportions among sectors for all relevant variables. In fact, based on the results of the test for differences in means or proportions (p-values are presented in Table 1) statistically significant differences among all sectors are observed in terms of all the variables except health status. Similar results were obtained when each specific level of government was compared to the private sector.

As will be shown in the next section, this high level of heterogeneity of the units suggests that the sample does not include comparable workers from each sector, which is a necessary condition if we want to test the premise of whether “similar” workers are paid similar wages. In addition, propensity score methods may be used to select homogenous subsamples, which may facilitate the comparison. However, these methods do not work with samples that do not already include a significant proportion of comparable workers, since the heterogeneity of the sample reduces the chances of finding good matches.

Variables Definition		Values	Private	Federal	State	Local	P-Value
CENREG	Census region division (%):Northeast (1); Midwest (2); South (3); West (4)	1	16.33	15.56	18.79	15.45	0.000
		2	26.48	21.41	28.44	25.95	
		3	34.49	34.97	26.55	29.36	
		4	22.71	28.06	26.22	29.23	
EDUCTG	Level of education (%): HS dropout (1); HS (2); some college (3); college and more (4)	1	21.54	9.90	11.45	16.83	0.000
		2	39.79	35.65	29.67	34.72	
		3	28.97	37.88	38.62	32.08	
		4	9.70	16.57	20.27	16.37	
FULTME	Fulltime working condition (%): yes (1); no (0)	0	24.71	9.02	26.83	29.96	0.000
		1	75.29	90.98	73.17	10.04	
GENDER	Gender (%): male (1); female (0)	0	52.36	48.10	61.31	61.68	0.000
		1	47.64	51.90	38.69	38.32	
HEALTH	Self-reported health status (%): excellent / good (1); otherwise (0)	0	6.39	5.71	5.83	5.66	0.588*
		1	93.61	94.29	94.17	94.34	
MARTST	Marital status (%): married (1); otherwise (0)	0	44.81	35.03	44.12	32.21	0.000
		1	55.16	64.97	55.88	67.79	
RACE	Race (%): white (1); otherwise (0)	0	14.79	26.26	19.35	20.35	0.000
		1	85.21	73.74	80.65	79.65	
UNION	Union status (%): yes (1); no (0)	0	98.85	85.66	92.84	92.23	0.000
		1	1.15	14.34	7.16	7.77	
VETERN	Veteran status (%): yes (1); no (0)	0	90.87	70.69	88.58	90.04	0.000
		1	9.13	29.31	11.42	9.96	
AGE	Age (years)	Mean	35.41	40.81	37.92	40.13	0.000
		s.d	12.10	10.34	12.19	11.12	
HRSWEK	Hours worked per week	Mean	36.81	40.09	34.97	35.46	0.000
		s.d	10.14	8.17	10.74	10.84	
PS	Propensity score	Mean	0.131	0.172	0.150	0.154	0.000
		s.d	0.054	0.060	0.059	0.055	
WAGEHR	Hourly wage	Mean	9.68	13.80	10.75	11.06	0.000
		s.d	5.99	6.62	5.67	6.90	
N	Sample size		45,686	1,607	1,804	3,695	

Note: \* denotes failure to reject the null hypothesis of equal means or proportions between sectors.

As stated before, based on the Lee's selection bias correction model, we estimate equation (3) using a multinomial logit model to obtain consistent estimators of the conditional expected values of the residuals in the multinomial decision model ( $m_j$  for  $j=1, \dots, 4$ , one for each sector of employment). These results are used to adjust wage equations as given by (4) by sector of employment, where the estimated coefficients corresponding to these  $m_j$  variables represent the covariance between  $\varepsilon$  and  $v$ .

As shown in Table 2, in all sectors wages increase with human capital investments, but the returns to these investments vary by sector. In fact, age has a positive effect on wages, at a declining rate. This marginal effect is substantially higher in the state level, and relatively smaller for workers in the local level. The returns to education at all level are higher in the federal sector, compared to all other sectors. At the local level, only for high skill workers are these returns higher than those in the private sector, while the state level shows the lowest returns at almost all levels of education.

Gender and race are positively related to wages, indicating that men and whites are paid more than women and nonwhites. However their impact is not higher in the private sector as anticipated. Being married and in good health are also associated with higher wages, except at federal level. Also as expected, fulltime status positively affect wages. Finally, the coefficients that account for the cohort effect have all positive signs, indicating a constant increase of average wages over time, usually stronger in the private sector over the period studied.

<b>Variables</b>	<b>Private</b>	<b>Federal</b>	<b>State</b>	<b>Local</b>
Age	0.0466 (22.46) <sup>1</sup>	0.0445 (3.71) <sup>1</sup>	0.0556 (5.00) <sup>1</sup>	0.0239 (2.66) <sup>1</sup>
Age <sup>2</sup>	-0.0005 (-23.90) <sup>1</sup>	-0.0005 (-3.85) <sup>1</sup>	-0.0006 (-5.20) <sup>1</sup>	-0.0003 (-3.30) <sup>1</sup>
Eductg (High school)	0.0614 (9.53) <sup>1</sup>	0.1465 (3.74) <sup>1</sup>	0.0120 (0.32)	0.0454 (1.78) <sup>3</sup>
Eductg (Some college)	0.0447 (4.12) <sup>1</sup>	0.1759 (3.22) <sup>1</sup>	0.0752 (1.30)	0.0021 (0.05)
Eductg (College +)	0.1350 (8.12) <sup>1</sup>	0.2903 (3.96) <sup>1</sup>	0.0407 (0.52)	0.1774 (3.04) <sup>1</sup>
Race (White)	0.1018 (12.66) <sup>1</sup>	0.1198 (3.32) <sup>1</sup>	0.0541 (1.50)	0.0725 (2.58) <sup>1</sup>
Gender (Male)	0.2305 (39.65) <sup>1</sup>	0.2147 (7.12) <sup>1</sup>	0.2095 (7.23) <sup>1</sup>	0.2548 (11.27) <sup>1</sup>
Health (Good health)	0.0877 (7.26) <sup>1</sup>	-0.0229 (-0.36)	0.0440 (0.74)	0.1243 (2.69) <sup>1</sup>
Martst (Married)	0.1253 (19.76) <sup>1</sup>	-0.0207 (-0.62)	0.1585 (4.83) <sup>1</sup>	0.0128 (0.51)
Fultme	0.1898 (35.85) <sup>1</sup>	0.2467 (6.78) <sup>1</sup>	0.2751 (10.97) <sup>1</sup>	0.1786 (10.00) <sup>1</sup>
T (1994) <sup>(a)</sup>	0.0494 (8.73) <sup>1</sup>	0.0059 (0.22)	0.0657 (2.52) <sup>2</sup>	0.0846 (4.17) <sup>1</sup>
T (1996)	0.1002 (15.23) <sup>1</sup>	0.0544 (1.72) <sup>3</sup>	0.1697 (5.28) <sup>1</sup>	0.1127 (4.68) <sup>1</sup>
T (1998)	0.1829 (28.33) <sup>1</sup>	0.0420 (1.30)	0.1626 (5.26) <sup>1</sup>	0.1934 (8.17) <sup>1</sup>
T (2000)	0.2657 (41.00) <sup>1</sup>	0.1364 (4.12)	0.1942 (6.48) <sup>1</sup>	0.2288 (10.03) <sup>1</sup>
m1	-0.6097 (-1.72) <sup>3</sup>	-0.3491 (-1.00)	0.4522 (1.32)	-0.5219 (-2.24) <sup>2</sup>
m2	-0.7515 (-2.72) <sup>1</sup>	0.0280 (0.50)	-0.1975 (-0.57)	-0.6712 (-2.67) <sup>1</sup>
m3	-3.9013 (-13.30) <sup>1</sup>	0.1915 (0.18)	-0.7340 (-4.22)	-1.9071 (-2.35) <sup>2</sup>
m4	0.0958 (0.28)	-1.5979 (-2.14) <sup>2</sup>	0.6691 (0.88)	-0.3019 (-2.09) <sup>2</sup>
Constant	0.3481 (12.18)	0.4758 (1.42)	2.3897 (5.05)	1.2283 (2.68)
R <sup>2</sup> (%)	33.08	28.25	43.07	28.89
N	38,841	1,394	1,564	3,036

Notes: t-statistics in parentheses. a) In all cases, 1992 is the year of reference. 1 denotes significant at 1 percent level; 2 denotes significant at 5 percent level; 3 denotes significant at 10 percent level.

The measure of rent received by public sector workers is obtained by comparing the marginal returns to the productivity-related characteristics in all levels of government to those of the private sector. The results of the indexes given by expression (5), evaluated at the average values of the private sector workers<sup>8</sup> indicate that federal workers at any level of education are underpaid compared to 'similar' private sector workers, while state and local level workers receive a positive wage premium that decreases with the level of education, but increases for workers with college or more (See Table 3). This would suggest that low skill workers receive relatively higher remunerations in the state and local levels of the government compared to 'similar' private sector workers, while these sectors tend to remunerate highly skilled workers less.<sup>9</sup>



Level of Education	Two-step methodology			Without controlling for selection		
	Federal	State	Local	Federal	State	Local
<b>HS dropout</b>	-0.230	0.476	0.219	0.137	0.000	0.009
<b>High School</b>	-0.145	0.427	0.203	0.164	-0.048	-0.010
<b>College</b>	-0.099	0.507	0.176	0.130	-0.046	-0.024
<b>College +</b>	-0.075	0.382	0.261	0.109	-0.010	0.065

Although the results support the hypothesis that the public sector tends to reward unskilled workers more than those in the private sector, they are in conflict with previous studies suggesting a general positive wage premium paid in the public sector, decreasing with the level of government.<sup>10</sup>

The wage differentials presented in the last section are interpreted as the premium (positive or negative) public sector workers receive compared to private sector workers. However, for these results to be considered a pure measure of the rent paid in the public sector, we should compare the returns that similar individuals working in different sectors receive, otherwise such results would also account for the differences in the productivity related characteristics themselves. But, how comparable are these workers? According to the results of the test of differences in means and proportions shown in Table 1, strong differences exist among individuals across sectors.

As stated before, the fact that we are controlling for the impact of the selection of sector of employment on wages does not ensure that all or most of the workers in the sample are similar to each other. The propensity score matching methods permit the researcher to select subsamples of individuals from each sector that are similar in terms of their PS. Following the general algorithm used to estimate the PSs that satisfy the balancing assumption, we started with a parsimonious specification of equation (3) and then tried higher order of the variables and/or interactions if the covariates were not balanced. Variables for which the balancing condition was never satisfied were dropped. As Table 4 shows, it was impossible to estimate a logit model where all the proposed covariates were balanced, without transforming some of the variables<sup>11</sup> or dropping others, with the subsequent loss of relevant information.<sup>12</sup> The state level yielded the best results in terms of the number of confounding factors that were controlled for, i.e. all of the proposed variables except region. The poorest model corresponds to the local government, so that this sector was finally excluded from the analysis.

These results can be seen as a signal of great differences among workers, not only with respect to the private sector, but also within the different levels of the public sector. In fact the need for different decision models for each level of government suggests that the structure of the decision about sector of employment varies across levels. Finally, having such remarkable dissimilarities among groups, such that the estimated PSs do not always control for all the relevant factors, is expected to increase the difficulty of finding appropriate matches.

The matching methods consist of taking each treated unit and finding a control unit with a PS close enough to consider it as a similar counterpart. The more restrictive the matching conditions, the smaller the differences between units and the better the quality of the matches, but at the cost of less efficiency due to the difficulty of finding appropriate matches. Among the different matching methods tried, the nearest neighbor (NN) matching method ended up showing the highest percentage of average bias reduction, but also with the smallest sample size compared to the other matching methods. However, the heterogeneity already observed in this data set and the subsequent difficulties in estimating decision models that satisfy the balancing assumption also affected the ability to find appropriate matches. After matching the individuals based on the estimated PSs for the federal/private and state/private comparisons, we found that almost no worker in the private sector can be considered as a reliable counterpart for workers in those two levels of government. In fact, among the 45,686 private sector workers in the sample, only 71 workers were selected for the federal/private comparison and 154 for the state/private comparison. These results support the idea that the workers included in the CPS are so different with respect to each other, that any comparison of their returns to human capital investments, as the one obtained in the previous section, may yield misleading results if such differences are not accounted for.

<b>Variables</b>	<b>Federal / Private</b>	<b>State / Private</b>	<b>Local / Private</b>
Gender (Male)	--	-0.4163 (-7.75) <sup>1</sup>	--
Race (White)	0.7755 (11.57) <sup>1</sup>	-0.2962 (-4.79) <sup>1</sup>	--
Agectg (22-34 years)	1.2286 (7.30) <sup>1</sup>	0.0518 (0.57)	0.9822 (10.79) <sup>1</sup>
Agectg (35-49 years)	1.7697 (10.67) <sup>1</sup>	0.3331 (3.63) <sup>1</sup>	1.5292 (17.13) <sup>1</sup>
Agectg (50+ years)	1.6428 (9.53) <sup>1</sup>	0.4340 (4.28) <sup>1</sup>	1.6195 (17.19) <sup>1</sup>
Martst (Married)	--	-0.0839 (-1.59)	--
Eductg (College+) <sup>(a)</sup>	0.5881 (7.86) <sup>1</sup>	0.8370 (12.90) <sup>1</sup>	0.4542 (8.82) <sup>1</sup>
Vetern	1.2705 (21.19) <sup>1</sup>	0.4113 (4.85) <sup>1</sup>	-0.1359 (-2.32) <sup>2</sup>
Constant	-5.2428 (-32.48)	-3.1195 (-33.41)	-3.7911 (-44.51)
N	47,293	47,490	49,381
LR Chi <sup>2</sup>	956.29	307.34	692.20

Notes: t-statistics in parentheses. a) Eductg recoded as having or not a college degree. 1 denotes significant at 1 percent level; 2 denotes significant at 5 percent level; 3 denotes significant at 10 percent level.

Wage equations as given by equation (4) estimated on the NN matched subsamples would allow the comparison of the returns to productivity-related characteristics among truly similar workers. However, we have to keep in mind that the results in this way obtained might not be reliable due to the small sample sizes and the relevant information missed.

The wage estimates based on OLS applied to the matched subsamples are summarized in Table 5.<sup>13</sup> When comparing the federal and the private sector (subsample based on PS1) the results indicate that federal workers are paid higher returns to experience (AGE) and to high levels of education (some college and more) compared to similar private sector workers. The positive effect of race on wages seems to be stronger in the private sector, while the impact of gender is higher in the federal sector. Health status, marital status and full-time working condition have a stronger impact on wages in the private sector, with the first one showing an unexpected negative sign in the federal sector. The cohort effect generally shows a positive impact on wages, but usually stronger in the private sector.

The comparison between the state level and its matched private sector (based on PS2) suggest that both sectors pay almost similar returns to experience and to high and low levels of education, but lower returns for workers with some college in the private sector. As expected, the effect of race on wages is higher in the private sector, while the gender variable seems to have greater impact in the state level. As for the other productivity-related variables, their positive effect on wages is always higher in the private sector.

Based on these results, wage differentials such as those obtained in the previous section based on expression (5) were calculated. The estimated wage differentials shown in Table 6 suggest that both federal and state workers at any level of education are underpaid, compared to similar private sector workers.

Variables	Federal-private comparison based on PS1 <sup>(c)</sup>		Federal-private comparison based on PS2 <sup>(d)</sup>	
	Federal	Private	State	Private
Age	0.5737 (6.96) <sup>1</sup>	0.0437 (2.13) <sup>2</sup>	0.0540 (6.79) <sup>1</sup>	0.0655 (4.29) <sup>1</sup>
Age <sup>2</sup>	-0.0005 (-5.71) <sup>1</sup>	-0.0004 (-1.44)	-0.0006 (-5.30) <sup>1</sup>	-0.0007 (-3.67) <sup>1</sup>
Eductg (High school)	0.1432 (2.89) <sup>2</sup>	0.1650 (0.04)	0.0691 (1.53)	0.0688 (0.44)
Eductg (Some college)	0.2014 (4.56) <sup>1</sup>	0.1840 (0.33)	0.1622 (4.58) <sup>1</sup>	0.0965 (0.92)
Eductg (College +)	0.3625 (8.04) <sup>1</sup>	0.3184 (1.48)	0.3806 (9.24) <sup>1</sup>	0.3837 (3.00) <sup>1</sup>
Race (White)	0.0817 (2.86) <sup>2</sup>	0.0942 (0.22)	0.0087 (0.29)	0.0905 (1.14)
Gender (Male)	0.1586 (6.74) <sup>1</sup>	0.0098 (0.64)	0.1548 (5.00) <sup>1</sup>	0.0920 (0.63)
Health (Good health)	-0.0205 (-0.77)	0.3373 (1.08)	0.0459 (0.14)	0.1001 (1.00)
Martst (Married)	0.0170 (0.26)	0.1178 (0.41)	0.0940 (4.56) <sup>1</sup>	0.2401 (2.57) <sup>2</sup>
Fultme	0.2471 (5.42) <sup>1</sup>	0.4602 (2.38) <sup>2</sup>	0.2798 (8.73) <sup>1</sup>	0.3738 (2.25) <sup>2</sup>
T (1994) <sup>(a)</sup>	0.0037 (0.18)	-0.0149 (-0.55)	0.0659 (2.40) <sup>2</sup>	0.1220 (0.58)
T (1996)	0.0521 (0.17)	0.2354 (1.44)	0.1559 (4.55) <sup>1</sup>	0.0805 (0.17)
T (1998)	0.0417 (0.46)	0.0644 (0.12)	0.1437 (5.01) <sup>1</sup>	0.1311 (0.37)
T (2000)	0.1352 (2.88) <sup>2</sup>	0.4410 (2.02) <sup>2</sup>	0.1772 (6.72) <sup>1</sup>	0.1362 (0.94)
Constant	0.6713 (3.35)	0.0050 (0.10)	0.4595 (4.45)	0.071 (0.13)
R <sup>2</sup> (%)	33.70	96.87	45.67	96.22
N	855	54	1,014	124

Notes: t-statistics in parentheses. a) In all cases, 1992 is the year of reference. b) Comparison with respect the local government not included. c) PS1 refers to PS estimated based on the federal/private comparison. d) PS2 refers to the PS estimated based on the state/private comparison. 1 denotes significant at 1per cent; 2 denotes significant at 5 percent; 3 denotes significant at 10 percent.

Level Of Education	Federal	State
HS dropout	-0.695	-0.494
High School	-0.716	-0.494
Some College	-0.677	-0.428
College +	-0.650	-0.497

Two things are worth noting at this point. It is true that this last comparison is based on individuals from different sectors that are comparable, at least in terms of their PS. However, it is also true that this PS lacks relevant information as the result of estimation problems due to the high level of heterogeneity of the sample. This problem, together with the small sample size finally obtained may lead one to conclude that the results of the propensity score

matching method are not reliable, or even more that such a method is not applicable to the wage comparison across sectors. Whatever conclusion is reached, it is still true that the individuals in this CPS sample are too heterogeneous both between sectors and within sectors to allow any comparison based on similar workers. Therefore, any wage differential calculated based on such a sample is not a reliable measure of how much more or less a public sector workers are paid compared to similar private sector workers, unless the differences among workers are accounted for.

The poor results obtained in the PS matching method were expected. We knew since the beginning that a fundamental condition for this method to work is to have a significant proportion of comparable units in all groups. Otherwise it would be difficult to find appropriate matches, substantially reducing the efficiency of estimates. However when the whole analysis is based on the premise that units are as similar as possible, traditional selection bias correction models also may be not enough, since they do not control for this required homogeneity. Unfortunately, the CPS data set seems to be unsuitable for use in comparing wages among similar workers, due to the strong dissimilarities observed among workers in this data set.

## **CONCLUSIONS**

This study approaches the public/ private wage comparison issue by comparing the results obtained from the application of two different methodologies that focus on correcting the selection bias problems that result from the nonrandom allocation of workers across sectors. We show that accounting for the endogeneity of the sector of employment does not ensure that the groups are comparable. If in addition the analysis assumes that the workers from different sectors are similar to each other, we also need to find reliable counterparts.

When controlling for the factors affecting the individual's decision about sector of employment, the results suggest that state and local workers are paid a positive wage premium compared to private sector workers. This premium generally decreases with the level of education. Federal workers, on the other hand, tend to be under compensated at all levels of education. However, a look at the main statistics of the variables indicated that the data set includes individuals that show little similarities among them, dissimilarities that are not accounted for by simply modeling the effect of the sector of employment decision on the wage comparison.

In applying the propensity score matching method, we found that deep differences exist among workers both between and within sectors, and that only a small and not significant proportion of them were similar in terms of their probability of working in the public sector. In fact, the problems faced when applying this method reinforced the initial conclusion that the CPS data set was too heterogeneous to be used to compare wages across sectors. Such heterogeneity affected the possibilities of finding appropriate matches, so that the results from this method on this data set lack efficiency.

These suggest that in spite of its wide use in comparing wages among sectors, the CPS March supplement -at least for the period considered, shows such a high level of heterogeneity that it makes it difficult to find appropriate matches. This finding illustrates the non-appropriateness of this data to measure rent under the premise that similar workers should be paid similar wages, unless some restrictions are introduced in order to make the sample more homogenous. An appropriate transformation of the sample would be to divide it into groups by age, limiting the analysis to a specific group or set of groups. One would expect more homogeneity within older age groups, compared to younger ones since the former tend to include workers that already have completed their education and have stabilized their position in the labor force. However, the price of this higher level of homogeneity again will be less efficiency.

## **REFERENCES**

1. Abadie, Alberto and Guido Imbens. (2002). Simple and Bias-Corrected Matching Estimators for Average Treatment Effects. Mimeo, University of California at Berkeley.
2. Bellante, Don and James E. Long. (1981). The Political Economy of the Rent-Seeking Society: The Case of Public Employees and Their Unions. *Journal of Labor Research* Volume 2, Number 1 (Spring), 1-14.

3. Belman, Dale and John S. Heywood. (1989). Government Wage Differentials: A Sample Selection Approach. *Applied Economics* Volume 21, 427-438.
4. Borjas, George J. (2002). The Wage Structure and the Sorting of Workers into the Public Sector. Working Paper, no. 9313. NBER, (October).
5. Choudhury, Sharmila. (1994). New Evidence on Public Sector Wage Differentials. *Applied Economics* (March), 259-271.
6. Dehejia, Rajeev H. and Sadek Wahba. (1999). Causal Effects in Non-Experimental Studies: Re-evaluating the Evaluation of Training Programs. *Journal of the American Statistical Association* Volume 94, Number 448, 1053-1062.
7. \_\_\_\_\_. Propensity Score-Matching Methods for Non-experimental Causal Studies. (2002). *Review of Economics and Statistics* Volume 84, 151-161.
8. Gyourko, Joseph and Joseph Tracy. (1988). An Analysis of Public- and Private- Sector Wages Allowing for Endogenous Choices of both Government and Union Status. *Journal of Labor Economics* Volume 6, Number 2, 229-251.
9. Heckman, James. (1974). Shadow Prices, Market Wages, and Labor Supply. *Econometrica* Volume 42, 679-693.
10. \_\_\_\_\_, Hidehiko Ichimura, Jeffrey Smith, and Petra Tood. (1997). Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Program *Review of Economic Studies* Volume 64, 605-654.
11. \_\_\_\_\_, and Salvador Navarro-Lozano. (2004). Using Matching, Instrumental Variables, and Control Functions to Estimate Economic Choice Models. *The Review of Economics and Statistics* Volume 86, Number 1, 30-57
12. Hoffnar, Emily and Michael Greene, (1996): Gender Discrimination in the Public and Private Sectors: A Sample Selectivity Approach. *Journal of Socio-Economics* Volume 21, 1105-114.
13. Hundley, Greg. (1991). Public- and Private-Sector Occupational Pay Structures. *Industrial Relations* Volume 30, Number 3 (Fall), 417-434.
14. Ichimura, Hidehiko. (1993): Semiparametric Least Squares (SLS) and Weighted (SLS) Estimation of Single-Index Models. *Journal of Econometrics* Volume 58, 71-120.
15. Imbens, Guido W. (2000). The Role of the Propensity Score in Estimating Dose-response Functions. *Biometrika* Volume 87, 706-710.
16. Katz, Lawrence F. and Alan B. Krueger. (1991). Changes in the Structure of Wages in the Public and Private Sectors. Working Paper, no. 282. Princeton University, (March).
17. Kordas, Gregory and Steven F. Lehrer. (2002). Matching using Semiparametric Propensity Scores. Mimeo, University of Pennsylvania.
18. LaLonde, Robert. (1986). Evaluating the Econometric Evaluations of Training Programs. *American Economic Review* Volume 76, Number 4, 604-620.
19. Lee, L.F. (1983). Generalized Econometric Models with Selectivity. *Econometrica* Volume 51, 507-512.
20. Poterba, James M. and Kim S. Rueben. (1994). The Distribution of Public Sector Wage Premia: New Evidence Using Quantile Regression Methods. Working Paper Number 4734. NBER, (May).
21. Ramoni, Josefa. (2004). Compensation Comparisons Between the Public and Private Sectors in the United States: A Propensity Score Matching Approach. Unpublished doctoral dissertation, University of South Florida.
22. Rosenbaum, Paul and D. Rubin, (1983). The Central Role of the Propensity Score in Observational Studies for Causal Effect. *Biometrika*, Volume 70, Number 1, 41-55.
23. Rubin, D. (1973). Matching to Remove Bias in Observational Studies. *Biometrics* Volume 29, 159-183.
24. Smith, Sharon P. (1974a). Pay Differentials Between Federal Government and Private Sector Workers. Working Paper Number 53. Industrial Relations Section, Princeton University, (September).
25. \_\_\_\_\_. (1974b). Government Wage Differentials by Region and Sex. Working Paper Number 58. Industrial Relations Section, Princeton University, (December).
26. \_\_\_\_\_. (1977). Government Wage Differentials. *Journal of Urban Economics* Volume 4, 248-271.
27. \_\_\_\_\_. (1981). Public/private Wage Differentials in Metropolitan Areas. In *Public Sector Labor Markets*, edited by P. Mieszkowski and G. E. Peterson, Washington D.C.

28. Venti, Steven. (1989). Wages in the Federal and Private Sectors. In *Public Sector Payrolls*, edited by D. Wise, University of Chicago Press.

#### ENDNOTES

1. Actually, the final sample includes the CPS data for years 1992, 1994, 1996, 1998 and 2000.
2. Notice that the PS can be used as a control function when inserted in the right hand side of the wage equation to control for the probability of choosing a specific sector of employment. However, this still would not ensure that the workers are comparable.
3. Notice that testing the mean of each characteristic is a necessary but not sufficient condition for the balancing assumption to hold, since higher order moments of the distribution of these variables may differ across sectors. However, the state of the art of the software programs available allows just a weak test of the balancing condition of the pre-treatment variables limited to the comparison of the first moment of the covariates.
4. Imbens (2000) proposed an extension to the methodology suggested by Rosenbaum and Rubin that allows for estimation of average causal effects with multi-valued treatments. However, software programs for this so called Generalized Propensity Score Method (GPSM) are not available yet. The GPSM follows three steps: first, multinomial discrete response models are estimated for the set of distinct treatments; second, the conditional expectation of the outcome given the level of the treatment and the probability of receiving the treatment is calculated. Finally, the average response at treatment level is estimated.
5. The model does not control for union status, as traditional wage models do, since unions may help to provide rent for workers. Models controlling for union status may underestimate the true magnitude of the rent (Ramoni 2004).
6. Other matching methods were tried, but the NN is the one for which the matched subsamples showed the highest percentage of bias reduction.
7. For the propensity score matching method the index should be evaluated based on the average characteristics of the matched private sector workers.
8. For the binomial variables a value of one was always assumed; the comparison is done at year 2000.
9. Notice that without controlling for the selection of sector of employment, the results are totally the opposite, suggesting that a rent is consistently paid only in the federal level, while workers in the state and local levels are generally underpaid.
10. Smith (1974a) found a wage premium that ranges from 29 percent for workers at the federal level, to 5.5 percent at the state level, but is usually negative (for men) at the local level. Bellante and Long (1981) found a positive wage premium of about 20 percent at the federal level, going down to 2 percent at the state level, and becoming negative for workers at the local level. Gyourko and Tracy's estimates range between 18.9 and 1.2 percent.
11. EDUCTG was recoded as having or not having a college degree. AGE was divided into 4 categories.
12. Heckman and Navarro (2004) define what is the minimal relevant information required to properly estimate the PS.
13. Heteroscedasticity problems are more likely to be present when working with individual data. To solve it, robust equations were estimated using the Huber-White estimator of variance. In general equations supported the assumptions of non multicollinearity and correct specification, except for the private sector based on PS1.

**ENDNOTES**

---

- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9
- 10
- 11
- 12
- 13