JOURNAL OF URBAN ECONOMICS 19, 156–173 (1986)

The Earnings of State Government Employees in the United States

GEORGE J. BORJAS*

Department of Economics, University of California, Santa Barbara, California 93106

Received March 27, 1984; revised July 19, 1984

I. INTRODUCTION

In 1979 there were 3.7 million state government employees (SGE's) in the United States, with a monthly payroll of $3.9 billion. This state government employment represented a sizable increase from the 1970 level of 2.8 million workers.[1] Recent work by Smith [1977] provides evidence of a wage differential between SGE's and private sector workers. She finds that (male) wage rates are about 8% higher in the public sector. However, controlling for observable socioeconomic characteristics reverses this finding and leads to a 3–11% wage differential favoring the private sector.[2] Although Smith's study presents interesting statistics, it does not provide much insight into employment policies followed by state governments which lead to their wage and employment levels. This paper provides a behavioral foundation by shifting the focus from a comparison of state government/private sector wage levels to a study of wage differentials among SGE's employed by different states.

The framework used in the analysis extends my earlier work on wage policy in the federal bureaucracy [see Borjas, 1980, 1982]. It is assumed that state governments view personnel policies as a political tool, and that the policies followed are those that have the highest political returns. This model yields implications regarding variations in sizes of state governments, the number of SGE's, and the SGE wage as a function of the relevant political factors in the state.

Table 1 summarizes state government employment statistics. The first column gives the number of SGE's in 1970 for the largest 15 states. Column

*I am grateful to Professor Edwin Mills for extensive comments which greatly improved the clarity of the paper.

[1] This data was obtained from the U.S. Department of Commerce "Statistical Abstract" [1980, pp. 320–322].

[2] See Smith [1977, p. 63]. The results discussed in the text are obtained from the male sample in the 1975 Current Population Survey.

0094-1190/86 $3.00

TABLE 1

State Government Employment Statistics in the 15 Largest States

State	1970 Employment (in 1000's)	$\ln(WAGE_g)$	$\ln\left(\frac{WAGE_g}{WAGE_p}\right)$
California	239	1.531	.172
New York	201	1.467	.089
Pennsylvania	132	1.299	.027
Texas	133	1.151	.025
Illinois	132	1.340	−.063
Ohio	107	1.223	−.111
Michigan	121	1.532	.081
New Jersey	65	1.397	−.040
Florida	79	1.190	.062
Massachusetts	67	1.431	.088
Indiana	70	1.291	−.006
North Carolina	71	1.128	.187
Missouri	64	1.172	−.079
Virginia	74	1.182	.067
Georgia	61	1.102	.041
Total U.S.	2755	1.305	.048

Note. Source: Column 1 is obtained from U.S. Department of Commerce [1980, p. 322]. The remaining columns are calculated from the 1970 U.S. Census male sample. Key to variables: $WAGE_g$ = wage rate in the state government sector; $WAGE_p$ = Wage rate in the private sector.

2 presents the average (ln) wage rate for SGE's in these states. The variation in these figures is remarkable. The average (ln) wage rate varies from 1.10 to 1.53 in these 15 states. Of course, part of the differences may be due to differences in wage levels across states unrelated to political factors. Column 3, however, reveals the persistence of strong differentials after accounting for regional wage differentials. In particular, it presents the average (ln) wage differential between SGE's and private sector workers in each of the states. It can be seen that the wage differential varies from −.079 to .187 in the 15 largest states. Table 1, therefore, suggests the existence of significant differences in the relative wage of SGE's across states.

Section II presents the theoretical framework which will provide an interpretation of the wage differentials documented in Table 1. Section III derives the variables which will proxy for the unobserved theoretical factors in the empirical analysis. Section IV presents the empirical results obtained from the 1970 Public Use Sample of the U.S. Census. Finally, Section V summarizes the study.

II. FRAMEWORK

Suppose there are n identifiable constituency groups in the state. These n groups may represent income classes, industrial categories, or blocks of voters organized along ethnic/racial lines. The proportion of the state population (N) in the ith group is q_i, where $\sum_{i=1}^{n} q_i = 1$.[3]

Suppose that voters get utility from consuming both private and public goods. Utility is translated into political support through a "vote" function, where higher levels of voter utility lead to stronger political support for the incumbent government. The level of private goods consumption is represented by disposable income available to voters. Public goods consumption (i.e., state government output) is represented by the size of the state government as measured by the number of SGE's. Assume that all type i constituents are identical so that they all have the same income level Y_i. Let t be the proportional tax rate on personal income and L be the size of the state government's labor force. Then the vote function for a representative individual in the ith constituency group is given by

$$V^i = V^i\left[L, (1-t)Y_i\right], \tag{1}$$

where $V_L^i = \partial V^i/\partial L > 0$; $V_y^i = \partial V^i/\partial(1-t)Y > 0$; and $V_{LL}^i < 0, V_{yy}^i < 0$. Throughout the analysis it will be assumed that (1) is strongly separable in its arguments ($V_{Ly}^i \equiv 0$). This separability property (along with the concavity restrictions) ensures that the second-order conditions are satisfied.

Equation (1) can be interpreted as the probability that a representative individual from constituency group i supports the state government. The government's objective is to maximize the expected number of votes:

$$\text{Max}\, \bar{V} = \sum_{i=1}^{n} p_i q_i N V^i\left[L, (1-t)Y_i\right], \tag{2}$$

where p_i is the probability that type i individuals vote.

The government is constrained by the wealth of the state. Through its power of taxation the government can "spend" the state's resources on the consumption of public goods and private goods. The budget constraint is

$$F = wL + \sum_{i=1}^{n} (1-t) q_i N Y_i, \tag{3}$$

where F is the state's wealth, and w is the SGE wage. In order to derive the state's demand function for SGE's it will initially be assumed that w is constant. The link between the exogenous variables in the model and the

[3] It is assumed that the q_i are exogenous to the model. It is clear, however, that in a more general framework state government policies may create particular coalitions leading to a different vector of q_i. For a discussion of endogenous constituency groups see Peltzman [1976] and Becker [1983].

SGE wage will be discussed below. It will also be assumed that the state's wealth, F, is independent of actions of the state government.[4]

The state government's objective is to maximize (2) subject to (3) by its choice of L and $(1 - t)$. The first-order conditions are

$$\sum_{i=1}^{n} p_i q_i N V_L^i - \lambda w = 0, \tag{4}$$

$$\sum_{i=1}^{n} \left(p_i q_i N V_y^i Y_i - \lambda q_i N Y_i \right) = 0, \tag{5}$$

where λ is the marginal vote gain due to a small increase in the state's wealth.

The comparative statics of the model imply that

$$\frac{dL}{dF} > 0. \tag{6}$$

Hence states with greater wealth have larger state governments. This income effect follows from the assumption of diminishing marginal benefits to additional private goods consumption by voters.

A more interesting problem is addressed by considering variations in the constituency vector $(q_1, \ldots, q_n)$. In particular, consider an increase in the population share of the ith group. Since $\Sigma q_i = 1$, an increase in q_i must be compensated by decreases in some remaining proportions. A natural experiment is to compensate the increase in q_i by decreases in all $q_j (j \neq i)$ such that the relative shares of the $n - 1$ groups in the vector $(q_1, \ldots, q_{i-1}, q_{i+1}, \ldots, q_n)$ remain constant. Thus, the increase in q_i does not affect the ratio q_j/q_k $(j, k \neq i)$. Such a change in q_i also causes a change in the state's level of wealth. Abstracting from this income effect, it can be shown that the direction of the change in labor demand due to a compensated change in q_i is

$$\operatorname{sign} \frac{dL}{dq_i} = \operatorname{sign} \left[\varepsilon_L^i \bar{\varepsilon}_y - \varepsilon_y^i \bar{\varepsilon}_L \right], \tag{7}$$

where

$$\varepsilon_x^i = \frac{\partial \ln V^i}{\partial \ln x}, \qquad \bar{\varepsilon}_x = \frac{1}{\bar{V}} \sum_{j=1}^{n} p_j q_j N V^j \varepsilon_x^j.$$

Equation (7) implies that a compensated increase in q_i increases the size of the state government if the "tastes" for public services are relatively stronger in group i than in other groups. Greater insight into this result is

[4] In other words, the time horizon of the model is short enough that the mobility of resources caused by the choice of relatively high or low tax rates has not had an effect on F.

provided by restricting the variation in tastes across the various constituency groups. It is reasonable to expect ε_L^i to vary among the groups since state services are likely to be directed to specific targets.[5] But there is no reason to expect differences in voters' tastes toward private consumption, and Stigler and Becker [1977] suggest it is reasonable to assume that ε_y^i is the same for all i. Then $\varepsilon_y^i = \bar{\varepsilon}_y$ and

$$\operatorname{sign}\frac{dL}{dq_i} = \operatorname{sign}\left(\varepsilon_L^i - \bar{\varepsilon}_L\right). \tag{7'}$$

Therefore, the change in q_i increases the size of state government if group i benefits from public goods relative to all other voters, and decreases it otherwise.

Abstracting from income effects, p_i and q_i enter symmetrically in the model since the number of "effective" constituents is determined by their product. This fact suggests that (7) and (7′) indicate what happens to L when p_i changes.[6]

Finally, the last comparative static results of interest concern the changes in L caused by a change in the state's income distribution. Consider first the effect of an increase in the state's income inequality. The change to be considered is a mean-preserving change in the income distribution which holds the proportions q_i constant. The latter restriction is imposed since the effect of changes in the q_i on L have already been derived in (7) and (7′). The former restriction is imposed to abstract from the aggregate income effect which arises from general changes in the income distribution. A mean-preserving increase in inequality clearly involves changing the $(Y_1, \ldots, Y_n)$ vector. The required change in this vector can be obtained by defining

$$R_i = \alpha Y_i + \beta, \tag{8}$$

where R_i is the transformed level of income for the ith constituency group. A mean-preserving change imposes $E(Y) = E(R)$, so that

$$\beta = (1 - \alpha)E(Y). \tag{9}$$

[5]The variation in ε_L^i is an alternative (and simpler) way of introducing into the model the fact that some state government services are targeted to specific groups, and that, in principle, the vote functions in (1) should depend on a vector describing the nature of state output and not on the single characteristic L.

[6]As with the q_i, it should be noted that although p_i is assumed to be exogenous, a more general model should take into account that the decision of whether or not to vote can be modeled in terms of many of the variables in this model. Voting models, however, have not resolved the essential question of why rational individuals even bother to vote. The model in this paper ignores this issue in order to derive empirical implications of how *given* voting patterns by individuals affects the behavior of state governments. A recent empirical analysis of the voting decision is given in Ashenfelter and Kelley [1975].

Hence a mean-preserving transformation in the income distribution is defined by

$$R_i = E(Y) + \sigma Z_i, \tag{10}$$

where $Z_i = (Y_i - E(Y))/\sigma^*$; σ^* is the standard deviation of Y_i; and $\sigma = \alpha\sigma^*$, the standard deviation of R_i. An increase in α, therefore, captures a mean-preserving increase in income inequality. If all constituency groups have the same tastes toward private consumption (i.e., $\varepsilon_y^i = \varepsilon_y$ for all i), it can be shown that[7]

$$\frac{dL}{d\alpha} \propto \sum_{j>i} \sum_{i} q_i q_j (Y_i - Y_j)\left(p_j V_y^j - p_i V_y^i\right). \tag{11}$$

Consider the special case where all constituency groups have the same voting participation probability (i.e., $p_i = p$ for all i). Note that the sign of (11) then depends on the sum of terms like $(Y_i - Y_j)(V_y^j - V_y^i)$. Since all individuals are assumed to have identical tastes toward the consumption of private goods, the diminishing marginal (voting) utility of income implies that each of these terms is nonnegative, and will be positive when $Y_i \neq Y_j$. Thus (11) is positive and the mean-preserving increase in income inequality increases the size of the state government. The intuition for this result follows from the fact that a mean-preserving increase in income inequality "stretches" the values in the vector $(Y_1, \ldots, Y_n)$. This spreading out of the income levels, along with the diminishing marginal utility of income, implies that high-income groups place little value on the additional income while low-income groups place a high value on the lost income. These utility changes induce the government to increase its political support by redistributing the state's income through the production of public goods.

In addition, (11) shows that this result is affected by the correlation between voting participation rates and income levels. If high-income groups have higher voting participation probabilities than low-income groups, the effect discussed above is attenuated and may even be reversed. Thus the sign of the effect of increased income inequality on the demand for SGE's depends on a comparison of the rate of change in marginal utility with the rate of change in the voting participation probability as income increases. Increased income inequality may lead to a reduced demand for public goods if high-income voters are loud enough in their response to state government policies.

The second aspect of the income distribution which can be analyzed is its mean. It is easy to show that a change in $E(Y)$ (holding F and σ constant)

[7]To derive Eq. (11) it is necessary to assume that the elasticities ε_y are constant.

leads to

$$\frac{dL}{dE(Y)} < 0. \tag{12}$$

The intuition for this result follows from the first-order conditions in (4) and (5). The marginal rate of substitution given by $d(1 - t)/dL$, at a given level of political support, decreases when $E(Y)$ increases. This increase in the relative valuation by voters of their private goods consumption, holding F constant, reduces the demand for public goods.

Up to this point the model has focused on the determination of the optimal level of labor inputs in the state bureaucracy. The resulting demand function, of course, does not completely determine the equilibrium wage and employment in the state's public sector unless a labor supply function to state government jobs is also specified. Since most state governments are relatively large employers, it is reasonable to suspect that the supply curve of SGE's is upward sloping. This fact implies that shifts in the demand function for SGE's, due to the factors identified in this section, lead to corresponding changes in the equilibrium wage rates paid to SGE's. Hence the model derived in this section can be summarized by[8]

$$L = D(F, q, p, \sigma, E(Y), w) \tag{13}$$

$$L = S(U, w) \tag{14}$$

where D and S are the demand and supply functions, respectively; q is the vector $(q_1, \ldots, q_n)$; p is the vector $(p_1, \ldots, p_n)$; and U is the vector of variables which shift the supply function. It is the estimation of the demand function in (13) which is the objective of the empirical analysis.[9]

[8] The demand function has two properties which are worth noting. First, it is independent of N since the size of the state's population does not appear in the marginal rate of substitution, in the slope of the budget constraint, or in the budget constraint itself (which can be rewritten as $wL = tF$). Second, the use of voter characteristics as shifters of the demand function is compelled by the fact that the output of public services chosen by the state is unobserved. In a sense, therefore, the vector of voter characteristics proxy for the determinants of this level of output.

[9] An interesting empirical problem arises due to the monopsony power of the state government. Since the equilibrium wage rate is not given by the intersection of the supply and demand curves, it is important to determine which locus of points observable data identifies. It is easy to show that if the demand function is given by $L = Z\gamma_0 - \gamma_1 w$, and the supply function by $L = U\gamma_2 + \gamma_3 w$, the variation in supply curves across states identifies the locus of points given by $L = Z\gamma_0 - (\gamma_1 + \gamma_3)w$. Thus the regressions below correctly estimate the effects of the shift variables on the demand for state bureaucrats, but give an upward biased labor demand elasticity.

III. DATA

The data analyzed in this study are the 1/100 1970 U.S. Census Public Use Sample (5% State File). The analysis is restricted to working males aged 18–64 whose records provided the key variables used in the analysis.[10] There is one potentially important problem with using Census data for the analysis of state government wage determination. Census data identify the individual's state of residence and not his state of employment. To the extent that state employees live in one state while working in another, the study assigns the wrong state employer to the individual. Unfortunately, this problem cannot be corrected.

The Census provides information on earnings of state government and private sector workers. A more difficult task is obtaining proxies for shift variables in the demand function. The vector of independent variables which will be used in the regressions is:

1. State Wealth (INCOME). The measure of state wealth used below is the income from all sources accruing to individuals living in the state in 1969. This variable proxies for the theoretical variable F and is expected to have a positive effect in the demand function. The source for this data is U.S. Department of Commerce [1971].

2. Income Distribution (MEANWAGE and INEQUA). The proxies for the theoretical variables $E(Y)$ and σ are defined as the mean of log annual earnings and the standard deviation of log annual earnings in the state's male labor force in the private sector.

These variables are calculated from 1970 Census data. The theoretical analysis predicted that, ceteris paribus, MEANWAGE has a negative impact on demand, while INEQUA has a positive impact.

3. Proportions in Constituency Groups. The proxies for the vector q are the hardest to derive. To obtain the q_i the constituency groups must be identified. Clearly any listing of constituencies is bound to be arbitrary and incomplete. The empirical analysis identifies three constituencies:

(a) *Labor (UNION).* This is measured by the percentage of the state's labor force that was unionized in 1968. The data source is U.S. Department of Commerce [1972].

(b) *The Poor (POVERTY).* This is defined as the percentage of individuals in the state who are below the official poverty line in 1969. The data source is, again, U.S. Department of Commerce [1972].

[10] The sample is composed of *all* male observations identified as state government workers, and of a 20% random sample of male private sector workers. These restrictions led to a sample of 14,007 state government workers and 61,193 private sector workers.

(c) *The Unemployed* (*UNEMP*). This is the state's unemployment rate in 1969. The data source is U.S. Department of Commerce [1971].

The effects of these variables on the demand for SGE's cannot be signed unambiguously. Instead their effect depends on how each group benefits from state government services *relative* to the benefits received by the average voter. Thus the effects of these variables provide insights into the redistribution mechanism implied by state government policies.

4. Tastes Toward Public Goods (*VOTDEM*). The model allows vote functions to differ across states. To capture this variation the variable VOTDEM, indicating the percentage of voters who voted Democratic in the 1966–1968 Gubernatorial elections, is introduced. If heavily Democratic states favor larger governments, this variable has a positive effect in the demand function. The data source for this variable is Scammon [1968].

5. Voting Participation Rates (*VOTERP*). It is difficult to obtain the p_i associated with each of the constituencies above. In fact, available data allow the definition of only two income groups (above the median and below the median), and the voting participation probability can be calculated for each of these two groups. VOTERP is defined as the ratio of the high-income voting participation rate to the low-income voting participation rate. As with the q_i, the model predicts that the sign of this variable depends on a comparison of how high- and low-income persons benefit from state government services. The data source for this variable is the 1968 American National Election Study.[11]

6. SGE Organization (*STATEU*). Well-organized SGE's can capture a significant portion of the income being redistributed by the state government. The variable STATEU, the proportion of SGE's that is unionized, is introduced to capture this organizational effect. The data source for STATEU is the U.S. Department of Commerce, "Census of Governments" [1972].

7. Cost of Living Differentials (*LWHAT*). There are likely to be significant wage differences across states that are unrelated to political factors. These regional wage differences lead to wage differences among SGE's in

[11] The 1968 American National Election Study was conducted by the Survey Research Center of the University of Michigan. The survey contained 1673 observations, and it provided information on both socioeconomic characteristics of individuals and their electoral preferences and behavior. It should be noted that the data set is relatively small and, therefore, some states have few (and some have no) observations. In the case where VOTERP could not be ascertained for a particular state its regional mean was used. The relatively small sample size implies that VOTERP is likely to contain a high amount of measurement error and it should not be surprising if its effect hovers around zero.

different states, and will probably impart serious bias to the effect of variables like MEANWAGE. To control for this factor the variable LWHAT is included in the regressions. LWHAT is defined as the wage the average male private sector worker in the United States would earn if he were employed in the private sector of each state. Its variation, therefore, captures differences in the structural parameters of the wage determination process across states. It is calculated from the 1970 Census data by first estimating a log earnings function among male private sector workers in each of the states, and then using these estimated structures to predict the wage the average male woule earn in each state.[12]

IV. EMPIRICAL ANALYSIS

The first step in the empirical analysis is to illustrate the standardized wage differentials which exist among SGE's employed by different states. To measure these differences the following regression is estimated:

$$\ln w_{ij} = X_{ij}\beta + \sum_{j=1}^{49} \gamma_j S_j + \varepsilon_{ij}, \tag{15}$$

where w_{ij} is the wage rate of individual i in state j; X_{ij} is his vector of socioeconomic characteristics such as education, experience, etc. (see footnote 12 for a complete listing); and S_j is a dummy equal to unity if the individual resides in state j and 0 otherwise. The omitted state in the S vector is Texas. The estimated $\hat{\gamma}$'s for the 15 largest states are given in column 1 of Table 2.

The results reveal a large variance in the standardized SGE wage among the 15 largest states. For example, we find that the wage rate of an SGE in Texas is 27% lower than that of a similarly skilled SGE in California, 31% lower than that in New York, and 15% lower than that in Illinois. Further, the inclusion of the state dummies in (15) significantly increased the explanatory power of the regression. The R^2 for the wage rate equation using only the socioeconomic characteristics vector X was .218; this increased to .248 after adding the state dummies. The F statistic associated with this increase in explanatory power is 11.3, highly significant at conventional confidence levels.

An important issue which arises at this point is whether these results simply reflect interstate differences in wage levels existing in both the public and private sectors. A convincing rejection of this hypothesis is presented in column 2 of Table 2. In particular, suppose that all male observations in both the state government *and* private sectors are pooled across states, and

[12] The independent variables included in the regression are years of education; experience (age − education − 6); experience squared; whether married spouse present; whether health limits work; whether residing in metropolitan area; whether residing in central city; and race.

TABLE 2

Standardized Differences in State Employee Wages in the 15 Largest States (Relative to Texas)

	Regression	
	(1)	(2)
California	.2703	.2633
	(10.26)	(9.85)
New York	.3077	.2937
	(11.47)	(10.74)
Pennsylvania	.1614	.1538
	(5.25)	(4.93)
Texas	—	—
Illinois	.1531	.1639
	(4.82)	(5.05)
Ohio	.0684	.0718
	(2.06)	(2.12)
Michigan	.3008	.3138
	(8.82)	(9.02)
New Jersey	.2518	.2408
	(6.61)	(6.22)
Florida	.0497	.0452
	(1.34)	(1.19)
Massachusetts	.2572	.2358
	(6.91)	(6.23)
Indiana	.1323	.1532
	(3.25)	(3.69)
North Carolina	.0395	.0534
	(1.14)	(1.15)
Missouri	.0502	.0730
	(1.29)	(1.84)
Virginia	.1542	.1572
	(4.07)	(4.09)
Georgia	.0527	.0797
	(1.33)	(1.98)
Sample size	14,007	75,200

Note. The t statistics are given in parentheses. Column 1 presents the estimates of Eq. (15), while column 2 estimates Eq. (16).

that the regression

$$\ln w_{ij} = X_{ij}\beta + \delta_0 G_i + \sum_{j=1}^{49} \gamma_j (G_i S_j) + \varepsilon_{ij} \tag{16}$$

is estimated, where $G_i = 1$ if the individual is a state government employee, and zero otherwise. The vector of coefficients γ in (16) measures interstate differences in the standardized wage differential between the public and the private sector. If the variance documented in column 1 of Table 2 is also found among private sector workers, the hypothesis implies that the vector γ would be zero. The results in column 2 of Table 2, however, reveal significant variation in the standardized wage differential between public and private sector workers across states. The F statistic associated with the introduction of the $(G_i S_j)$ variables in the wage rate equation is 10.81, highly significant at conventional confidence levels. More interestingly, the numerical magnitude of the coefficients is very similar between columns 1 and 2 of Table 2. This indicates that the wage differentials documented within the sample of SGE's in column 1 cannot be explained by private sector wage differentials, but instead reflect systematic variation in the determination of wage rates across political units.

The theoretical model in Section II suggests that political factors are responsible for these wage differentials. To determine the importance of these political factors, the state dummies in (15) are replaced by the political variables introduced in Section III. This leads to the regression[13]

$$\ln w_{ij} = X_{ij}\beta_0 + Z_j\beta_1 + \varepsilon_{ij} \tag{17}$$

[13] It is easy to show that the estimation of (17) indeed identifies the demand functions derived in the model. Consider the following equations:

$$\ln w = X\beta_0 + S\gamma + \varepsilon \tag{F1}$$

$$\gamma = Z\beta_1 + v, \tag{F2}$$

where X is the vector of socioeconomic characteristics, Z is the vector of political variables, and $S = [S, \ldots, S_{50}]$ is a vector of state dummy variables indicating the individual's state of employment. The parameter vector γ, therefore, measures standardized wage differentials across states for SGE's, and is predicted by the model to be endogenous and determined by Eq. (F2). Substituting (F2) in (F1) yields

$$\ln w = X\beta_0 + (SZ)\beta_1 + (Sv + \varepsilon). \tag{F3}$$

The term SZ gives the value of the political variables for state j. Further, the disturbance vector in (F3) is homoscedastic as long as ε and v are drawn from independent distributions (with standard stochastic assumptions). Thus the least-squares estimation of (17) identifies the parameters of interest. Further, it can be shown that the pooling method used by (17) is equivalent to the two-stage procedure which first estimates (F1), and then estimates (F2) using $\hat{\gamma}$. For a proof, see Amemiya [1978].

TABLE 3

Estimates of State Employee Demand Functions[a]

Variable	Regression (1)	(2)	(3)	(4)
INCOME	.0074	.0072	.0089	.0081
	(4.81)	(4.71)	(5.77)	(5.27)
MEANWAGE	.2383	−.4813	.2173	−.7736
	(2.32)	(−2.57)	(2.07)	(−7.26)
INEQUA	.2412	.2243	.2943	.2793
	(2.69)	(2.50)	(3.23)	(3.10)
VOTERP	.0070	.0149	.0203	.0288
	(.39)	(.83)	(1.11)	(1.59)
VOTDEM	−.00002	−.0007	.0004	−.0006
	(−.03)	(−1.07)	(.61)	(−.99)
UNEMP	.0198	.0325	.0157	.0346
	(2.33)	(3.65)	(1.82)	(4.05)
POVERTY	−.0045	−.0110	−.0047	−.0144
	(−1.67)	(−3.63)	(−1.74)	(−5.32)
UNION	−.0012	−.0006	−.0010	−.0002
	(−1.20)	(−.61)	(−1.03)	(−.17)
STATEU	.0021	.0015	.0023	.0014
	(6.02)	(4.29)	(6.58)	(4.18)
$\hat{L}$	−.0036	−.0032	−.0047	−.0037
	(−4.84)	(−4.72)	(−6.30)	(−5.03)
LWHAT	—	.7592	—	1.0580
		(4.60)		(40.34)
R^2	.242	.244	.199	.216

[a] The t ratios are given in parentheses. Columns 1 and 2 estimate Eq. (17), while columns 3 and 4 estimate Eq. (18).

where w_{ij} is the wage of SGE i in state j; and Z_j is the vector of political variables associated with state j. The estimated coefficients of the Z_j variables are presented in columns 1 and 2 of Table 3.

The empirical results are extremely interesting. For example, the coefficient of INCOME (state wealth) has a strong positive effect on the SGE wage. In fact, evaluated at the sample mean, the elasticity of the SGE wage with respect to state wealth is approximately .23. The empirical analysis, therefore, supports the theoretical prediction of an increase in the SGE wage when state wealth increases.

The effect of the second characteristic of the state's income distribution introduced in the regressions, MEANWAGE (the average log wage rate of

males in the private sector), depends on the specification of the equation. In column 1 it is strongly positive, but in column 2, where LWHAT (the cost of living index) is held constant, it is significantly negative as predicted by the model. The sign switch occurs because in column 1 the effect of MEANWAGE is contaminated by the fact that interstate variation in mean wage levels may proxy for regional cost of living differences. Controlling for LWHAT nets out these interstate differences in the wage rate and MEANWAGE then has a significant negative effect. Thus the predicted theoretical effect of $E(Y)$ is confirmed by the empirical analysis.

The final characteristic of the state's income distribution introduced in the regression, INEQUA (the standard deviation of log earnings among males in the private sector), has a positive effect. Recall that the model predicted the sign of INEQUA to depend on two factors: It would be positive due to the assumption of diminishing marginal utility of income, but this effect could be reversed if the voting participation of high-income constituencies was significantly greater than that of low-income groups. The results in Table 3 suggest that the latter effect is weaker than the former; hence more income redistribution takes place the more unequal the income distribution in the private sector.

The variables VOTERP (the ratio of the voting participation rate of high-income to low-income individuals) and VOTDEM (the percentage of the electorate voting for a Democratic governor) both have insignificant effects on the demand for SGE's. It is of interest to note that the positive coefficient of VOTERP suggests that high-income persons benefit relatively more from state government services than low-income persons.

The next three variables in Table 3 measure the proportions in the three constituency groups. The coefficient of UNEMP—the state's unemployment rate—is positive and strong. Its positive sign can be interpreted, in the context of the political model, as an indication that unemployed individuals receive greater than average benefits from state governments. Thus the unemployed form an important constituency in state government, and their voice has an important role in determining the size of state government, and indirectly the number and wage rate of state employees.

The second constituency measure introduced in the equations is POVERTY, the percentage of the state's population below the poverty level. It is of great interest to find that POVERTY has a strongly negative effect in the equation. In terms of the political model in Section II, the results indicate that the poor do *not* receive relatively more state government benefits than other constituency groups. In fact, the strong negative effect of POVERTY indicates that the poor receive fewer benefits, on the average, than other constituency groups.

The third constituency group introduced in the equation is UNION, the percentage of the labor force unionized in the state. The coefficient of

UNION is negative, but not significant. The nonpositive effect of UNION indicates that organized labor is not one of the groups receiving relatively large benefits from the state government.

The last two political variables included in the regression are STATEU and $\hat{L}$. The first variable, STATEU, gives the percentage of SGE's that is unionized. It is strongly positive and highly significant. To the extent that a higher degree of unionization reflects an underlying cohesion of interests within the state government's labor force, its positive effect indicates the importance of homogeneity of interests in raising SGE wages. The numerical value of its coefficient (when controlling for LWHAT) indicates that if nonunion SGE's were to unionize, their wage rate would increase approximately 13–15%.

The variable $\hat{L}$ is the predicted value of the number of state government employees in the state. It has been instrumented because it is an endogenous variable likely to be correlated with the disturbance in the wage function.[14] The effect of $\hat{L}$ on all the regressions is negative and significant. The implied elasticity of the wage rate with respect to L is approximately .297, so that the demand curve is somewhat elastic. It is difficult to understand the reasons for this result since studies such as Ehrenberg [1973] have found relatively inelastic demand functions in the public sector.[15]

An important criticism of the results in the first two columns of Table 3 is that the political variables could be proxies for factors which create wage level differentials across states in both the public and the private sectors. In other words, the political variables are capturing regional wage differentials which remain even after controlling for LWHAT. To provide convincing evidence that this is not the case, the regressions were reestimated using the public/private sector wage differential, rather than the level of the public wage rate, as the dependent variable. This evidence is obtained by estimating the regression

$$\ln w_{ij} = X_i\beta_0 + \delta G_i + \beta_1(G_i^* Z_j) + \varepsilon_{ij}, \tag{18}$$

which interacts the public sector dummy (G_i) with the state's political variables vector (Z_j). The estimated coefficients in β_1 are presented in

[14] The variables used in the first stage employment regression are the state's population; per capita income; UNION, STATEU; INEQUA; MEANWAGE; mean education; VOTDEM; VOTERP; UNEMP; POVERTY; standard deviation of education in the private sector; standard deviation of education in the public sector. The data source for the variable L is the U.S. Department of Commerce "Census of Governments" [1972].

[15] Recall that the estimated elasticity of the wage rate with respect to L is upward biased due to the monopsony characteristic of the market. Hence the demand curve is even more elastic than implied by the results.

columns 3 and 4 of Table 3. Column 3 estimates (18), while column 4 adds LWHAT to the vector X_i. The similarity between the estimates of (17) and (18) indicates that the vector Z explains public/private wage differentials not because it explains interstate differences in the private wage, but because Z explains interstate differences in SGE wages. The difference between these two hypotheses is substantive, and the strong rejection of the former hypothesis indicates that the empirical analysis in this paper has identified a set of factors which systematically determine the size of the state government's labor force and the wage rate of SGE's across states.

The empirical results also shed light on the importance of political variables in explaining the interstate variance in the state government wage level. The simple wage model estimated in the sample of SGE's with X (the socioeconomic characteristics) alone has an R^2 of .2183; adding Z increases it to .2423; and replacing Z with the 49 state dummies increases R^2 to its maximum value of .2480. Thus the vector Z explains 80% of the possible variation in the standardized wage rate of SGE's across states. The vector Z, therefore, explains an important portion of the interstate variance in the wage rates of SGE's.

Furthermore, it is easy to show that this conclusion is not affected by differences in the occupational structure of SGE's across states. The tasks carried out by state governments differ significantly across states. This suggests that it is important to control for the jobs that SGE's do. Table 4 presents estimates of the wage determination functions within each of four samples defined along specific government tasks: construction, health, education, and public administration. These four government functions account for 81% of all state government employment in the United States.[16] The results in Table 4 reveal that, with minor exceptions, the functional breakdown is not an important determinant of the effects of the political variables on the state bureaucratic wage rate. For example, the unemployment rate and the poverty rate remain significantly positive and negative, respectively, in practically all the regressions. Similarly INEQUA still has a positive (or zero) effect, while MEANWAGE has a negative (or zero) coefficient. The results in Table 4, therefore, suggest that the political vector Z is not simply explaining wage differentials reflecting the different functional components of the 50 state governments. Instead, it is capturing interstate dispersion in real SGE wages due to differences in the relative political power of constituents and bureaucrats across states.

[16]The breakdown into these four categories uses the industry variable in the Public Use Sample and is defined by Construction, Census Code 068 (General Contractors, except building); Health, Census Code 838 (Hospitals); Education, Census Codes 857, 858 (Elementary and Secondary Schools, Colleges and Universities); Administration, Census Code 927 (State Public Administration).

TABLE 4

Effects of Political Variables on Bureaucratic Wage Rate by Function[a]

Variable	Government function			
	Construction	Health	Education	Administration
INCOME	.0059	.0158	.0095	.0043
	(1.71)	(3.17)	(3.27)	(1.63)
MEANWAGE	.2351	−1.0780	−.7772	−.3251
	(.61)	(−1.58)	(−2.32)	(−.94)
INEQUA	.5319	−.1159	.1594	.3330
	(2.66)	(−.33)	(1.00)	(2.19)
VOTERP	−.0067	.0429	.0157	.0248
	(−.18)	(.73)	(.47)	(.75)
VOTDEM	−.0017	.0012	−.0002	−.0013
	(−1.27)	(.54)	(−.13)	(−1.10)
UNEMP	.0345	.0277	.0187	.0428
	(1.85)	(.82)	(1.14)	(2.75)
POVERTY	−.0101	−.0214	−.0092	−.0109
	(−1.55)	(−1.94)	(−1.70)	(−1.98)
UNION	−.0039	−.0066	.0021	.0006
	(−1.94)	(−1.78)	(1.16)	(.36)
STATEU	.0015	.0020	.0009	.0021
	(2.01)	(1.52)	(1.24)	(3.35)
$\hat{L}$	−.0028	−.0063	−.0043	−.0020
	(−1.68)	(−2.55)	(−3.04)	(−1.47)
R^2	.288	.309	.213	.242
% Employed in type of job	14.5	7.3	36.9	22.3

[a] The t ratios are given in parentheses.

V. SUMMARY

This paper has presented a theoretical and empirical analysis of wage determination in the state government sector. The theoretical model is based on the assumption that the state government's objective is to maximize its political support. This objective function leads to predictable behavior regarding how state government employment policies react to changes in the state's wealth, income inequality, composition of constituency groups, and voting behavior.

The empirical analysis, conducted on the 1970 U.S. Census Public Use Sample, showed that the wage of state government workers shifts as a result of changes in these political factors. In fact, the results indicated that over three-fourths of the interstate variation in the state government wage rate

could be explained by the relatively small vector of political variables introduced in this paper.

It should be clear, however, that the analysis conducted in this paper is but a first step in the systematic study of state government behavior. There are several areas where further research may lead to useful results. For example, an important determinant of state government behavior is the composition and power of the various constituency groups. In this paper, the empirical analysis concentrated on three obvious constituencies and ignored the power coalitions in the rest of the population. A more complete accounting of all constituencies may uncover additional results regarding the nature of the income redistribution policies conducted by state governments. Such an accounting could, of course, use the large volumes of available data which divide the state budgets into particular functions. Thus a more general framework could match the data available on the composition of the various constituencies with the data available on expenditures in specific functions. It is not unreasonable to expect that such generalizations of the analysis conducted in this paper will lead to important insights into the labor economics of political markets.

REFERENCES

T. Amemiya, A note on a random coefficients model, *Internat. Econ. Rev.*, **19** (3), 793–796 (1978).

O. Ashenfelter and S. Kelley, Jr., Determinants of participation in presidential elections, *J. Law Econ.*, **28** (3), 695–733 (1975).

G. S. Becker, A theory of competition among pressure groups for political influence, *Quart. J. Econ.*, **98** (3), 371–398 (1983).

G. J. Borjas, Wage determination in the federal government: The role of constituents and bureaucrats, *J. Polit. Econ.* **88** (6), 1110–1147 (1980).

G. J. Borjas, The politics of employment discrimination in the federal bureaucracy, *J. Law Econ.*, **25** (2), 271–299 (1982).

R. G. Ehrenberg, The demand for state and local government employees, *Amer. Econ. Rev.*, **63** (3), 366–379 (1973).

S. Peltzman, Toward a more general theory of regulation, *J. Law Econ.*, **19** (2), 211–240 (1976).

R. M. Scammon (Ed.), "America Votes," Vol. 8, Congressional Quarterly, Washington, D.C. (1968).

S. P. Smith, "Equal Pay in the Public Sector: Fact or Fantasy," Princeton Univ. Press, Princeton, N.J. (1977).

G. J. Stigler and G. S. Becker, De gustibus non est disputandum, *Amer. Econ. Rev.* **67** (2), 76–90 (1977).

U.S. Department of Commerce, "Statistical Abstract of the U.S.," Government Printing Office, Washington, D.C. (various issues).

U.S. Department of Commerce, Bureau of the Census, "Census of Governments," Vol. 3, No. 3, Government Printing Office, Washington, D.C. (1972).