

AN ECONOMETRIC ANALYSIS OF THE DEMAND
FOR LABOR IN SPANISH MANUFACTURING

Luis Servén

Documenta 87-11

1. INTRODUCTION

The record high unemployment rates and the parallel collapse of manufacturing employment that Spain has suffered since the late seventies have contributed to attract increased attention to the determinants of labor demand in the Spanish economy, and particularly in its industrial sector. Several recent papers have analyzed empirically the recent evolution of employment in Spain, both from equilibrium and disequilibrium perspectives¹. Most of these studies use aggregate time series data whose quality is subject to considerable criticism².

An alternative source of information is provided by the Central de Balances of the Bank of Spain (CBBE), which since 1981 collects data on a large cross-section of manufacturing firms. However, the use of individual cross-section data from this source in some recent studies to analyze labor demand lead to the conclusion that industrial employment shows a surprisingly low response to changes in real output even in the medium run, a finding that would seem to rule out any possibilities of a substantial improvement in industrial employment over that time horizon.

¹See for example Dolado et al (1986) and Serrvén (1986), and the references therein.

²While the aggregate employment figures lack homogeneity due to the changes of definition, methodology and frequency of the employment surveys, the wage data seem to grossly overstate the wage increases that took place in the seventies.

In this paper we use the CBBE panel to estimate and test a simple empirical model of labor demand for Spanish manufacturing firms. To arrive at a consistent estimation procedure we use a straightforward extension of the techniques recently proposed by Griliches and Hausman (1986), to deal with errors in variables in the context of panel data.

The paper is organized as follows. First, we briefly discuss some theoretical issues in Section 2. Section 3 describes the estimation procedure and reports the empirical results. Finally, our conclusions are summarized in Section 4.

2. THEORETICAL ISSUES

We shall consider a very simple specification of labor demand that relates employment to real wages and output, as well as to labor productivity (whose evolution is assumed to follow a time trend) and other firm-specific effects that are unobservable. More formally,

$$(1) \quad l_{it} = \alpha_i + \beta_1 t + \beta_2 q_{it} + \beta_3 u_{it} + \eta_{it} \quad \begin{matrix} i=1 \dots N \\ t=1 \dots T \end{matrix}$$

where l , q and u represent the logs of employment, output and the real wage respectively, t stands for time, α_i is a firm-specific disturbance and η_{it} is an iid disturbance with mean zero.

It is well known that OLS or GLS estimation of equation (1) may give biased estimates of the β 's due to the potential correlation between the regressors and the omitted firm-specific variables. While simple transformations of the data (such as differencing or the within transformation) should in principle eliminate this source of bias, in practice they often lead to estimates which are closer to zero than economic theory would predict. Good examples of this situation are Mairesse and Dormont (1985) or, for the Spanish case, Sebastian (1986), who estimate simple labor demand equations similar to (1) using differenced data, and find long-run output elasticities well below unity even after allowing for substantial

lags in the effects of the explanatory variables.

Recent work by Griliches and Hausman (1986) has shown that these counterintuitive results may well be due to the presence of errors in variables. If some (or all) of the regressors are subject to errors of measurement, then different transformations of the data aimed at eliminating the individual effects α_i will generally result in different degrees of (usually, downward) bias in the parameter estimates. They point out that a judgement on the existence of errors in variables can be made by comparing the estimates that result from these alternative transformations, and emphasize that, under appropriate assumptions on the stochastic properties of the error of measurement, consistent IV estimation may be possible due to the availability of "internal" instruments obtained from simple transformations of the original regressors.

It is clear that the empirical results obtained using equations similar to (1) (usually in first-difference form) that we mentioned above may be due to measurement errors affecting the regressors. If firms' employment decisions are based upon the anticipated "permanent" levels of output and labor costs (as will be the case as long as hiring and firing decisions involve some sort of fixed costs) then the use in (1) of the current values of these variables instead of the theoretically relevant ones amounts to a typical situation of errors in variables.

Consistent estimation may be achieved if appropriate instruments for q and w in (1) are available. In the usual case that no valid "outside" instruments exist, identification of equa-

tion (1) will have to rely exclusively upon "internal" instruments.

Griliches and Hausman focus on the construction of such instruments in the case of one single regressor subject to measurement error. Application of their technique to our case of two explanatory variables potentially measured with error is straightforward. More generally, the K -regressor errors-in-variables problem can be formulated as

$$(2) \quad Y_{it} = X_{it}^* \beta + \alpha_i + n_{it} \quad i=1, \dots, N; \quad t=1, \dots, T$$

where X_{it}^* and β are $1 \times K$ and $K \times 1$ vectors respectively, and the columns of X_{it}^* are potentially correlated with the unobserved individual effect α_i . The true X_{it}^* are not observed; only their noisy counterpart X_{it} is available to the econometrician³:

$$(3) \quad X_{it} = X_{it}^* + E_{it} \quad i=1, \dots, N; \quad t=1, \dots, T$$

where again E_{it} is a $1 \times K$ vector of measurement errors. Stacking the model in terms of observables we have

³Note that the variables in equation (2) could be reinterpreted as residuals from a preliminary regression on the subset of variables that are known to be orthogonal to the disturbance.

$$(4) \quad Y_{it} = X_{it}\beta + (\alpha_i + \eta_{it} - \epsilon_{it}\beta) \quad \begin{array}{l} i=1,\dots,N \\ t=1,\dots,T \end{array}$$

Stacking the observations for given i , the model can be rewritten

$$(5) \quad Y_i = X_i\beta + i\alpha_i + \eta_i - \epsilon_i\beta \quad i=1,\dots,N$$

where i is a $T \times 1$ vector of ones, Y_i and η_i are $T \times 1$ vectors and X_i and ϵ_i are $T \times K$ matrices. Stacking the model once more we have

$$(6) \quad Y = X\beta + \alpha + \eta - \epsilon\beta$$

where Y, α, η are $NT \times 1$ vectors and X and ϵ are $NT \times K$ matrices.

The parameter vector β can in principle be estimated using an instrumental variables estimator $\beta_{iv} = (Z'X)^{-1}Z'Y$. In the absence of external instruments, Z will be a $NT \times K$ matrix of "internal" instruments of the form $Z^k = (I_N \otimes P_k)X^k$, where Z^k and X^k denote the k -th columns of Z and X respectively, and P_k is a $T \times T$ matrix whose structure depends on the stochastic properties of the measurement errors ϵ^k .

^aThe argument is obviously unchanged if the instruments for the k -th column of X are constructed from linear combinations of several columns of X , i.e. $Z^k = \sum_j (I_N \otimes P_{kj})X^j$.

For Z to be a valid instrument matrix, two conditions must be satisfied:

$$(7) \text{plim } \frac{1}{N} \sum_{i=1}^N Z_i' X_i = A$$

where A is a nonsingular matrix, and

$$(8) \text{plim } \frac{1}{N} \sum_{i=1}^N Z_i' (\alpha_0 + \alpha_1 + \eta_i - \epsilon_i) = 0$$

The first condition requires that the instruments be correlated with the regressors, and it may be rewritten

$$(9) \text{plim } \frac{1}{N} \sum_{i=1}^N \begin{bmatrix} (X_1^1)' P_1' X_1^1 & \dots & (X_1^1)' P_1' X_1^K \\ \dots & \dots & \dots \\ (X_1^K)' P_1' X_1^1 & \dots & (X_1^K)' P_1' X_1^K \end{bmatrix} = A$$

with A nonsingular.

The second condition is the usual requirement that the instruments be uncorrelated with the composite disturbance term. In the present context, it amounts to two different requirements:

$$(10a) E \epsilon_i = 0 \quad k=1, \dots, K$$

which eliminates the individual effect α_i from the model by having the columns of each P_k matrix sum to zero; and

$$(10b) \quad \text{plim} \frac{1}{N} \sum_i (E_i^k)' P_k E_i = 0 \quad k, l=1, \dots, K$$

which requires that the instruments be uncorrelated with the multivariate measurement error.

While (10a) imposes T linear restrictions on each P_k matrix, the number of restrictions implied by (10b) --and thus the number of linearly independent P_k matrices providing valid instruments for the k -th regressor-- will depend on the stochastic properties of the multivariate measurement error. Griliches and Hausman (1986, Appendix) provide a detailed account of the $K=1$ case. Obviously their results would apply also to the $K>1$ situation if measurement errors for the different variables were mutually uncorrelated at all time lags, in which case (10b) reduces to

$$(10b') \quad \text{plim} \frac{1}{N} \sum_i (E_i^k)' P_k E_i = 0 \quad k=1, \dots, K$$

which is the equation analyzed by Griliches and Hausman. In that case, (10b') and the stochastic properties of E^k alone would determine the number of available instruments for the k -th regressor (provided (9) is satisfied).

In the case of mutually correlated errors,

the restrictions that (10b) imposes on the P_k matrices may be more stringent than those implied by (10b'). For example, if measurement errors are nonstationary and follow a joint MA(1) process then condition (10b) would impose $T-1$ additional restrictions on P_k over those implied by (10b') in the case of independent MA(1) processes. Similarly, a joint nonstationary MA(2) process would impose $T-2$ restrictions on each P_k matrix in addition to the ones implied by the independent MA(2) case.

The number of linearly independent P_k matrices satisfying (10a) and (10b) is also the number of available instruments for the k -th regressor (unless (9) is violated). Obviously, if the measurement errors for the different variables follow processes of the same order (e.g., they all are MA(1) then the number of valid instruments will be the same for each regressor.

Once all the available instruments have been found, the efficient estimator can be computed using the procedures suggested by Hausen (1982) and White (1982). It is given by

$$\hat{\beta} = \left[\begin{array}{cc} \tilde{X}' & \tilde{Z}' \Omega^{-1} \tilde{Z}' \end{array} \right]^{-1} \begin{array}{c} \tilde{X}' \\ \tilde{Z}' \Omega^{-1} \tilde{Z}' \end{array} Y$$

where $\tilde{Z}$ is the matrix that has all the available instruments as columns, and Ω is the optimal weighting matrix (that allows also for conditional heteroscedasticity) defined by

$$\Omega = \frac{1}{N} \sum_{i=1}^n \tilde{Z}_i' u_i u_i' \tilde{Z}_i$$

with u_i being the residuals obtained from a preliminary consistent estimate of β .

3. EMPIRICAL RESULTS

We use data on 820 manufacturing firms for the years 1981-1985. Total employment in these firms amounted in 1985 to 360.000 workers, nearly 15% of aggregate manufacturing employment in Spain.

The data set only provides information on nominal quantities. Thus, to arrive at a measure of real output and real wages we used the INE price deflators, that consider a 17 sector disaggregation of manufacturing. This may well introduce a measurement error in both variables, to the extent that the price index relevant to specific firms differs from their sectoral one. Moreover, for each firm the CBBE only reports the total wage bill. Wage figures had to be constructed by dividing by the average number of employees, introducing an additional source of error in the wage variable.

Table 1 reports estimation results for the GLS and within specifications. Both included a time trend⁵, and the GLS regression included also 17 industry dummies⁶. While the estimates of the real wage and time coefficients are roughly simi-

⁵Preliminary regressions with year dummies lead to the acceptance of the hypothesis that a linear trend provides an adequate representation of the shifts of the labor demand equation over time.

⁶These were later excluded from the regressions since they never were jointly significant.

lar, the GLS estimate of the output elasticity appears substantially higher than its within counterpart. As expected, both point estimates are well below unity.

A Hausman (1978) test of the GLS vs. within specification yields a test statistic of 428.0. The null hypothesis that the GLS specification is correct is thus overwhelmingly rejected, a result which should not be surprising since individual firms' output and real wages are likely to be highly correlated with unobserved firm-specific characteristics such as technology or workers' ability.

At this point, a possible interpretation of the results would be to accept the presence of important firm-specific effects and conclude that the output elasticity of labor demand is below .5, as shown by the within regression, thus pointing towards the existence of substantial economies of scale.

The alternative view goes along the lines described earlier: firms' hiring decisions are based upon their anticipated 'permanent' levels of output and labor costs, which are unobservable. The use of current output and real wages in equation (1) creates an errors in variables problem that biases the resulting estimates (in particular, the output elasticity of labor demand) towards zero. In addition, the use of possibly inadequate price deflators would reinforce this effect⁷.

⁷Specifications including lagged employment as a regressor—in the spirit of the adjustment cost model—were also estimated using instrumental

This question can be verified by computing alternative estimators using differenced data. As Griliches and Hausman (1986) emphasize, if errors in variables are not present then OLS on the differenced data should yield similar estimates regardless of the length of the differencing period.

The resulting estimates are reported in Table 2. As expected, the first difference estimates are closer to zero than the within ones. In particular, the estimate of the output coefficient is similar to that obtained by Sebastian (1986) using cross-section data. As the differencing period becomes longer, both the technical progress coefficient and the output elasticity estimates rise in absolute value, while the real wage coefficient shows a more erratic behavior. For example, the long-difference estimate of the output elasticity is more than 40% higher than the first-difference estimate.

A Wald test of the equality of the first and long difference estimates leads to a χ^2 variable with 12 degrees of freedom. Its computed value is 42.1, thus rejecting overwhelmingly the null hypothesis of equality of coefficients for these alternative differencing periods.

In view of these results, it seems appropriate to reestimate the employment equation using an instrumental variables procedure, allowing both for the presence of measurement error in the ob-

variables procedures of the type discussed by Hsiao (1986). The coefficient on lagged employment turned out to be always insignificantly different from zero.

servation of the relevant output and real wage variables and for nonzero correlation between the explanatory variables and the firm-specific effects. Our data set only provides us with some additional financial variables (e.g., firms' indebtedness or total assets) whose validity as instruments would be highly questionable due to the likely simultaneity between firms' real and financial decisions, and also because firms' financial structure may well be systematically related to their unobserved specific characteristics.

We therefore turn to the use of "internal" instruments, of the type described by Griliches and Hausman (1986) and Hsiao (1985) in the context of panel data. Under appropriate assumptions, different linear transformations of the right-hand side variables provide enough instruments to identify all the parameters of interest.

As noted above, the number of available "internal" instruments depends upon the specific assumptions that one is willing to make on the stochastic properties of the measurement errors. We shall begin by assuming that the errors of measurement involved in using "transitory" output and wages instead of their "permanent" values are nonstationary and serially uncorrelated, although they may be contemporaneously correlated^a. Given our sample size ($T=5$), such a hypothesis yields thirty orthogonality conditions (fifteen alterna-

^aNonstationarity requires that the diagonal elements of each F_k matrix be zero (see 10b). Contemporaneous correlation between the measurement errors for the different variables would have the same implication. Hence, in this particular case (serially uncorrelated errors) the number of instruments will be the same whether errors are or are not mutually correlated.

tive transformations for each of y and w/p), so that the model is heavily overidentified. Allowing for a joint first-order moving average process in the measurement errors reduces the number of internal instruments to fourteen, while if the MA(1) processes were mutually independent there would be four additional instruments for each of y and w/p . A joint second-order moving average process would yield only four orthogonality conditions, that would become sixteen if the MA(2) processes were independent. A higher-order moving average process for the measurement errors would leave the model unidentified without additional assumptions (e.g., stationarity) on their stochastic properties⁹.

For each of these specifications, the asymptotically efficient IV estimator was computed using the Generalized Method of Moments. The optimal weighting matrix was constructed using the residuals from a preliminary 2SLS regression, allowing also for conditional heteroscedasticity.

The resulting GMM estimates are reported in Table 3. The estimated output elasticity of labor demand rises from .5 under the assumption of serially uncorrelated measurement errors to almost .9 under the independent MA(2) assumption. Note that the independent MA(1) and MA(2) estimates of the output coefficient are not statistically different from unity. The absolute value of the real wage elasticity shows a parallel decline, although its changes do not seem highly significant. Allow-

⁹Other specifications with "asymmetric" error processes (e.g., MA(2) for output, MA(1) for the real wage) were also attempted. They were always rejected on the basis of Hausman tests in favor of the ones reported in the text.

ing for higher-order serial correlation of the measurement errors results in a substantial increase of the estimated standard errors (which allow for conditional heteroscedasticity), as the number of available instruments becomes gradually smaller.

The last two columns of Table 3 provide a formal check of the validity of the assumptions underlying the alternative models. Column 4 reports the statistic for Hansen's (1982) test of the overidentifying restrictions, which under the null hypothesis of a correct specifications is distributed as a chi-square with degrees of freedom equal to the number of available instruments minus the number of estimated parameters. For the case of uncorrelated measurement errors the test statistic is 59.7 with 28 degrees of freedom, which provides strong evidence of misspecification. In contrast, the test cannot reject the validity of the overidentifying restrictions for the MA(1) (joint or independent) and MA(2) specifications.

The fifth column of Table 3 reports the statistics for Hausman's (1978) specification test. Each one of them is distributed as a chi-square with 3 degrees of freedom under the null hypothesis that the model specification is correct. On the basis of this test, the uncorrelated errors assumption is rejected in favor of the joint MA(1), which in turn cannot be rejected when compared to the individual MA(1), joint MA(2) (not reported in the Table) or individual MA(2). Thus, mutually uncorrelated MA(1) or MA(2) processes for the measurement errors appear to provide an adequate representation of their stochastic properties.

To conclude this section, it may be worth to briefly summarize our empirical results. We found evidence that output and real wages may be correlated with firm-specific effects, and that conventional cross-section and panel data estimates of the demand for labor in Spanish manufacturing may suffer from errors in variables biases. In particular, the output elasticity obtained from conventional first-difference estimates appears to be strongly biased toward zero, as predicted by Griliches and Hausman (1986). By employing instrumental variables techniques we were able to obtain consistent estimates, that place the output elasticity of labor demand between .8 and .9 (and not significantly apart from 1), and the real wage elasticity around .6. These results are in fact fairly similar to the ones obtained using time-series data for Spanish manufacturing¹⁰.

To reach these conclusions, we allowed for errors in the measurement of the output and real wage variables relevant for firms' employment decisions. The finding that these errors may follow a MA(1) or MA(2) process could suggest that only persistent changes in the level of demand or real wages (that last for one or two years) are likely to be reflected in firms' employment policy.

¹⁰For example, the time series estimates reported by Raymond et al. (1986) yield the following long-run employment equation (standard errors in parentheses): $l = y - .456 w/p - .035 t$
(.157) (.010)

CONCLUSIONS

The responsiveness of labor demand to manufacturing output and labor cost conditions is a key issue in the current policy debate about the remedies to the unemployment problem in Spain. In this paper we have estimated a simple model of labor demand using panel data for 820 Spanish manufacturing firms. We found that conventional panel data estimates are very likely to understate the response of employment to changes in real output. By means of a straightforward extension of the instrumental variable techniques proposed by Griliches and Hausman (1986) to the case of several explanatory variables, we were able to obtain consistent estimates of our labor demand equation. The empirical results agree fairly well with those obtained in other studies using time series data, and they seem to imply that only persistent changes in real output and real labor costs are likely to be fully reflected in firms' employment decisions.

TABLE 1

ESTIMATES OF THE EMPLOYMENT EQUATION (N=4100)

<u>Method</u>	<u>Coefficient</u>		
	<u>Output</u>	<u>Real wage</u>	<u>Time</u>
GLS	.587 (.003)	-.645 (.014)	-.020 (.002)
Within	.461 (.011)	-.607 (.014)	-.017 (.002)

Note: standard errors in parentheses.

TABLE 2

ESTIMATES OF THE EMPLOYMENT EQUATION FOR
ALTERNATIVE DIFFERENCING PERIODS

<u>Dependent variable</u>	<u>Coefficient</u>			<u>N</u>
	<u>Output</u>	<u>Real wage</u>	<u>Time</u>	
$l_{it} - l_{it-1}$	.378 (.011)	-.615 (.011)	-.014 (.002)	3280
$l_{it} - l_{it-2}$	.420 (.013)	-.653 (.015)	-.017 (.004)	2460
$l_{it} - l_{it-3}$	.510 (.016)	-.624 (.022)	-.019 (.006)	1640
$l_{it} - l_{it-4}$	.532 (.023)	-.424 (.035)	-.019 (.010)	820

TABLE 3

GMM ESTIMATES OF THE EMPLOYMENT EQUATION

Assumed error process	Coefficient			Test of overidentifying restrictions (D.F.)	Hausman test
	Output	Real wage	Time		
MA(0)	.505 (.030)	-.736 (.063)	-.018 (.002)	59.7* (28)	MA(0) vs joint MA(1) 27.3*
Joint MA(1)	.590 (.163)	-.560 (.097)	-.022 (.004)	12.5 (12)	Joint MA(1) vs joint MA(2) 5.4
Individual MA(1)'s	.326 (.137)	-.636 (.107)	-.027 (.004)	17.2 (20)	Indiv MA(1) vs joint MA(1) 5.0
Individual MA(2)'s	.891 (.168)	-.589 (.140)	-.028 (.005)	14.2 (14)	Indiv MA(1) vs indiv MA(2) 1.0

Note: * significant at the 5% level.

REFERENCES

Dolado, J. et al (1986): "Spanish Industrial Unemployment: some explanatory factors", Economica, 53.

Griliches, Z. and J. Hausman (1986): "Errors in Variables in Panel Data", Journal of Econometrics, 31, pp 93-118.

Hansen, L. (1982): "Large sample properties of Generalized Method of Moments Estimators", Econometrica, 50, pp 1029-1054.

Hausman, J. (1978): "Specification Tests in Econometrics", Econometrica, 46, pp 1251-1271.

Hsiao, C. (1986): Analysis of Panel Data, Cambridge University Press.

Mairesse, J. and B. Dormont (1985): "Labor and Investment Demand at the Firm Level", NBER Working Paper num. 1554.

Raymond, J.L. et al (1986): "Factores Explicativos de la Demanda de Empleo", Papeles de Economía Española 26.

Sebastián, C. (1986): "Excedente Empresarial, Inversión y Demanda de Empleo en la Empresa Privada", Papeles de Economía Española 26.

Servén, L. (1986): "The Decline of Manufacturing Employment in Spain: a Disequilibrium Approach" mimeo, MIT.

White, H. (1982): "Instrumental Variables regression with independent observations", Econometrica, 50, pp483-499.