Modelling Employment, Output and Labour Costs: A General to Specific Methodology

Luis A. Valenzuela S.
MSc in Econ – University of London
Departamento de Economía, Recursos Naturales y Comercio Internacional, FAE
Universidad Tecnológica Metropolitana
luis.valenzuela@utem.cl

Abstract

The main purpose of this paper is to compare three log-linear models: a dynamic specification or unrestricted model, a partial adjustment model and a static one; using as raw material data about the UK level of manufacturing employment. We conclude that the dynamic model represents the best performance. This specification is a way to move from general to specific, which is an appropriate econometric methodology. A dynamic modelling process constitutes a more realistic approach to this sensitive demand side labour market.

Keywords: ágeneral to specificñádynamic specificationñácommon factor and labour marketñ

Resumen

El principal propósito de este artículo es comparar tres modelos logarítmico-lineales: una especificación dinámica o modelo no-restringido, un modelo de ajuste parcial y uno de carácter estático; utilizando como materia prima datos sobre el nivel de empleo manufacturero en el Reino Unido. Concluimos que el modelo dinámico presenta un mejor desempeño. Esta especificación es una forma de ir desde lo general a lo específico, lo cual es una metodología econométrica apropiada. Un proceso de modelamiento dinámico constituye una aproximación realista a este sensible mercado del trabajo por el lado de la demanda.

Palabras claves: ágeneral a específicoñ áespecificación dinámicañáfactor común y mercado laboralñ

INTRODUCTION

This paper discuss both the theoretical and empirical implications of a relationship between the level of employment “n” in the manufacturing sector of the United Kingdom and two explanatory variables: the level of output “y” in that sector -value added- and the real unit labour cost “w”.

In order to study this relationship we will compare three log-linear “l” models: (1) a dynamic specification or unrestricted model, in the same way as equations one and nine in Hendry and Mizon (1978); (2) a partial adjustment model, which is a restricted version of (1), because imposes the restriction that no lagged independent variables should be considered -only lagged dependent variables- in the modelling process; and (3) a static model with no lagged variables, which also corresponds to a restrictive formulation of (1), but valid when the latter has a common factor (COMFAC) - because it can be simplified to a static one with an autoregressive error.

The final goal of this paper is to discuss the advantages of the model which performs best the labour market by the demand side.

MODELLING THE EMPLOYMENT

The first question that arises here is where are the theoretical elements to support models as said above. The article of Hamermesh (1986) show us that if we start with a CES function -constant elasticity of substitution technology-, the linear homogeneous production function (see equation 16 of that paper) can be transformed into a demand for labour function (see equation 19 of that paper). If the marginal productivity condition (see equation 2a of that paper) or the Shephard condition holds, we can estimate an equation like:

1) $ln = k + b \cdot ly + d \cdot lw$,
where k (constant term), b and d are parameters, and with b=1 if the production function is characterized by constant returns to scale. It's necessary to remember that "d" corresponds to the elasticity of substitution, which is not the usual constant-output elasticity of the demand for labour. This elasticity, "N", is:

2) $N = -(1 - s) \cdot d$,
where "s" is the share of labour in total revenue. Since we have no possibility to calculate it from the set of given data, we shall consider "d" as a proxy for "N". It must be emphasized that this proxy won´t be a good one if "s" varies largely.

Thus, we can recognize this CES equation with model (3), being models (1) and (2) ways of representing, by using the same variables (n, y, w) -lagged or not lagged- a dynamic relationship and a partial adjustment specification, respectively.

Another way to study this relationship is to look at model (2):

3) $\Delta ln = k + a \cdot ln(-1) + b \cdot ly + d \cdot lw$
If we assume that a=1 -an extreme case of this model-, then the equation can be rewritten as:

4) $ln = ln - ln(-1) = k + b \cdot ly + d \cdot lw$,
which means that "y" and "w" essentially explain the rate of change in employment.

Additionally, if we believe that the rates of change in "y" and "w" are those which determine changes in employment, we are thinking more closely in a dynamic specification as model (1).

What seems to be clear is the importance of incorporating real wages to explain employment. As Nickell points out (see Hendry and Wallis 1984) in some of the key macroeconomic models of the British economy -those of the Treasury, the London Business School and the National Institute, among others-, employment is a simple function of the level of activity. He advocates to put the wage equation into the labour market, where it belongs.

ABOUT THE DATA

The econometric estimates use seasonally adjusted quarterly data in an index number form for the period 1978 - 1990.

Observing the data for the period 1978.1-1990.4, it's possible to conclude that:

- The level of employment in the manufacturing sector decreased quarter by quarter from 137 -index- in 1978 to 100 in 1987, mantaining this latter level until 1990.

- The real unit labour cost, on the other hand, increased dramatically in the period, probably reflecting a growth in productivity in the UK.

- The output -value-added- seems to follow a "U" shaped curve; first fell in general terms (1978-1983.2) and then increased up to the last period.

By looking at the data we suspect a negative relationship between employment and wages and a positive correlation between the former and output in the same period.

ECONOMETRIC ESTIMATES

First, we check the signs of the estimated coefficients, finding that they are well behaved in terms of a labour demand funcion. Output is positively related with employment, while salary is to the contrary. Even though in model (1) the variable "lw" is not significantly different from zero, the lagged variable "lw(-1)" is important. This problem disappears in models (2) and (3), where only "lw" is relevant, suggesting a high degree of collinearity between the variable labour cost and its lagged values.

To verify this problem we run a regression between "lw" and "lw(-1)". As we know, a high coefficient of determination and a little standard error -or a high t-calculated- in the coefficient of the variable "lw(-1)" indicate us a strong degree of correlation between them. This is also clear by looking at the data of this variable: "lw" is always increasing, so the current value of this variable is strongly related with its lagged value. From an empirical perspective, one of them seems to be enough to assure that the impact of wages on employment will be considered. By doing a similar estimate, we obtained a weak correlation between "ly" and "lw".

In order to test the statistical significance of any independent variable, including the constant term, we examine the values of the t-tests under the null-hypothesis that individual coefficients are not signicantly different from zero. Another important point is that we cannot reject, by using the appropriate t-test, the hypothesis that the coefficients of variable "ln(-1)" in models (1) and (2) are, also in an individual way, significantly equal to 1, which indicate us that we are essentially explaining the rate of change in the level of employment -more than its level- through a set of variables (y, w), lagged or not.

There are not great differences between the estimated standard errors of the regression coefficients and the heteroscedastic consistent standard errors -HCSE-, which is only a first approach to the absence of heteroscedasticity. The significance of the regression equation as a whole is checked by the F-test. The coefficient of determination -R squared- is a measure of the fit of the model. All the estimated equations exhibit a high level of goodness of fit -adjusted or unadjusted-, being in models (1) and (2) close to one.

The Lagrange-Multiplier test for residual autocorrelation -LM- show us if errors are white noise under the null hypothesis that there is no autocorrelation. This problem only seems to be clear in the case of model (3). While we cannot be so conclusive with models (1) -here we ought to test first COMFAC- and (2), the situation of (3) is clear.

If there is residual autocorrelation, then it is taken as a symptom of poor model design and the whole specification process should be reviewed. It is possible to verify this conclusion for model (3), by using the Durbin-Watson test - DW. As we know, we cannot use DW to test for serial correlation in autoregressive models, that is, models containing lagged values of the dependent variable as explanatory variables. In addition, this test is limited to disturbances that are generated by a first-order autoregressive scheme. In the case of this static model we think at first that the serial correlation might be due to the noninclusion of lagged values of both the explained and explanatory variables, which is a misspecification of the dynamic process - omitted variables. We check the behaviour of the estimated residuals for each equation. It is possible to appreciate that the residuals have a more systematic pattern of behaviour in model (3), because of the smoothly trend followed by the calculated errors. This is not so clear in the other cases, where the trend of residuals is changing more abruptly.

The test for autoregressive conditional heteroscedasticity checks whether the residuals have an ARCH structure, under a negative null-hypothesis - homoscedasticity. In models (1) and (2) the null-hypothesis is accepted, but in model (3) it is rejected, meaning that the variance of the error term depends on the past history of such errors.

The normality test detects the presence of observations whose behaviour or values are out of the pattern exhibited by the variables - outliers. The test for heteroscedastic errors checks if the disturbances have constant variances against the alternative that the squared disturbances depend on the original and squared regressors. This test was not conclusive about the presence of heteroscedasticity on the original and squared regressors. Finally, we check the reset test, under the null-hypothesis that there is no functional form misspecification. In model (3) we reject it, which means that the log-linear functional form is here a wrong way of specification.

SET OF SEASONAL DUMMIES

We test the statistical significance of a set of seasonal dummies, in order to review if the data have or not a seasonal pattern. First, we check if each dummy, in an individual way, is or not statistically different from zero. It is clear by looking al the t-tests that none dummy-variable is significantly different from zero.

Now, we test the statistical significance of a set of seasonal dummies by imposing linear restrictions ($R \cdot q = r$, see Greene 2003). The results of the Wald criterion, with a table F-value of 2,84 are: Model (1) = $F(3,42) = 0{,}3539$; Model (2) = $F(3,44) = 0{,}3118$; and Model (3) = $F(3,45) = 0{,}0810$. Therefore, we accepted the null hypothesis that the coefficients of seasonal dummies are statistically equal to zero, i.e., r = transpose of [0 0 0].

These results seem to be very consistent with three facts: First, we are using variables in a "log" form, which smooth oscillating movements in the data set. Second, we are working with lagged variables, at least in equations (1) and (2), thus reducing the problem as well - even though the lag is only one quarter. And third, as we said, series were previously seasonally adjusted. We shall consider that inappropriate procedures of seasonal adjustment can create serial correlation and dynamic specification problems, which lead to inconsistent and inefficient estimates, and that adequate methods for this matter should not alter the lag structure.

The last test imposes to model (1) the restriction that the sum of coefficients of the seasonal dummy-variables is equal to zero, i.e., r = [0], accepting this null-hypothesis.

COMMON FACTOR RESTRICTIONS

To test the validity of COMFAC -common factor restrictions, used to evaluate error autocorrelation- for the unrestricted dynamic specification, we first need to transform model (1):

5) $ln = k + a \cdot ln(-1) + b \cdot ly + c \cdot ly(-1) + d \cdot lw + e \cdot lw(-1) + u$

6) $Lx = x(-1)$,

the lag operator (first order polynomial), and "x" is any variable.

6) in 5):

7) $(1 - aL) \cdot ln = k + (b + cL) \cdot ly + (d + eL) \cdot lw + u$

If:
8) $c = -ab$ and $e = -ad$,

then equation 7) becomes:
9) $(1 - aL) \cdot ln = k + b\,(1 - aL) \cdot ly + d\,(1 - aL) \cdot lw + (u / 1 - aL)$,

so that the terms involving "n", "y" and "w" have a common factor of (1 – aL) and the polynomials in "L" have a common root of "a", the serial correlation coefficient. We reject the null-hypothesis, meaning that the model cannot be re-expressed with an autoregressive error, i.e., it can not be simplified to a static one -as model (3)- with an autoregressive error. Thus, we cannot improve efficiency by reducing the number of parameters to be estimated and also indicates the inadequacy of moving from a dynamic specification (1) to a static model (3).

ELASTICITIES AND LINEAR RESTRICTIONS

i) We also test the following demand-elasticities, which are always present in the economic literature of the labour market: Ey= 1 (elasticity of "n" with respect to "y") and Ew = -1 (elasticity of "n" with respect to "w").

In the long-run:

10) $x = x(-1)$,

so model (1) can be transformed in:

11) $(1 - a) \cdot ln = k + (b + c) \cdot ly + (d + e) \cdot lw + u$

Thus:

12) $ln = k^* + [(b + c) / (1 - a)] \cdot ly + [(d + e) / (1 - a)] \cdot lw + u$,

where: $k^* = [k / (1 - a)]$

13) $Ey = [(b + c) / (1 - a)] = 1$,
then: (*) $a + b + c = 1$

14) $Ew = [(d + e) / (1 - a)] = -1$,
then: (**) $d + e - a = -1$

We test (*) and (**) by using linear restrictions, in which r = transpose of [11]. We reject the null-hypothesis that these restrictions are valid. This means that there is not a one-to-one relationship in the long-term between "y", "w" and "n".

ii) In order to get model (2) from (1), we test c = 0 and e = 0 by using linear restrictions, where r= transpose of [0 0]. The F-test was considered not conclusive in this case. Unfortunately, a test of c + e = 0 is not useful in this situation to clear the problem, because the coefficients have opposite signs.

iii) In order to get model (3) from (1), we test a = 0, c = 0 and e = 0 by using linear restrictions, where r = transpose of [0 0 0]. The F-test strongly reject the null-hypothesis and also confirms the accumulated evidence against the static model (3). An optional way was to test for linear restrictions that a + c + e = 0, thus verifying the above conclusion.

COMPARISON OF THE THREE MODELS

The first thing we do is to summarize the main results obtained above for the three models:

	Model (1)	**Model (2)**	**Model** (3)
High R squared?	Yes	Yes	Yes
Are signs well?	Yes	Yes	Yes
Are t-tests well?	No	Yes	Yes
Collinearity [w, w (-1)]?	Yes	-	-
Coefficient of ln(-1) = 1?	Yes	Yes	-
F of the regressions?	Yes	Yes	Yes
DW – f.o. autocorrelation?	-	-	Yes (+)
LM – autocorrelation?	N C	N C	Yes
ARCH - heteroscedasticity?	No	No	Yes
Outliers?	No	No	No
Heteroscedastic errors?	N C	N C	N C
Functional form misspecification?	No	No	Yes
Importance of seasonal dummies?	No	No	No
COMFAC exists?	No	-	-
N C= non-conclusive			

The main statistical problem detected in model (1) is the strong collinearity between "w" and "w(-1)". In model (3) the problems are autocorrelation, ARCH, functional form misspecification and omitted variables.

From an exclusively statistical point of view, comparing the three models and their tests, we think that model (2) represents a more reasonable approach to the explained variable - employment. It is clear that model (3) is the worst way to do it - inefficient and inconsistent estimators. Model (2) also have better relevant tests than (1), probably due to higher degrees of freedom. As we know, collinearity among explanatory variables won't necessarily cause any problems in inference - this seems to be the case, see Maddala (1992) - and there are several solutions to this problem. In this sense, model (2) cannot be considered as superior to (1).

A model should never be selected only for statistical reasons. If model (2) is enough to assure a good statistical relationship, we'll see that is a poor specification of the demand side labour market. First of all, it can be said that only model (1) is a dynamic specification of this market. Models like the static regression usually present high autocorrelation, so that inferences about its parameters are invalid. This model corresponds to the most restrictive one, which makes difficult -and not so correct as strategy- to move from simple to general. Moreover, this specification stresses that the level of employment in a given period is well explained by the level of real output and real wages in the same period, which means that the market labour is cleared -in the sense of velocity of adjustment- in only one period, in this case one quarter, which represents a poor economic understanding of this market. On the other hand, a model of partial adjustment like (2) could be very useful if we assume that there is a long-run desired target for the level of employment and also when we think that the rate on change in employment is well explained by the levels of "y" and "w". With quarterly data it's difficult to think that is possible to reach a "natural rate" in the labour market or something like this - for instance, it's very useful to assume that individuals adjust period by period a part of their monetary balances to a desired level, because of the adjustment costs, but it has more problems in the labour market. Again, this specification only contains current variables "y" and "w", without considering that the impact of a change in these variables takes more time than one quarter to be completely felt in the labour market.

CONCLUSION

It's clear that since an economic point of view model (1) represents the best performance. This specification is a way to move from general to specific, which is an appropriate econometric methodology. Also, it considers the possibility that past values of explanatory variables can influence the behaviour of current employment, which is now a reasonable approximation to reality, because there are lags in market adjustments and economic agents do not respond inmediatly to changes in these variables – it's very costly for employees to change decisions rapidly. An obvious recommendation for future estimates is to incorporate a greater number of lagged variables in order to test then the validity of more restricted formulations. To conclude, a modelling process which begins by general model (1) constitutes a more realistic approach to this change-reactive market.

References

1. Greene, W. H. 2003. Econometric Analysis. Fifth Edition; Prentice Hall.

2. Hamermesh, D. S. 1986. The Demand for Labor in the Long Run. In Ashenfelter, O. C. and R. Layard, Editors, Handbook of Labor Economics, North-Holland, chapter 8.

3. Hendry, D. F. and Wallis K. F. 1984. Econometrics and Quantitative Economics. Oxford: Basil Backwell.

4. Hendry, D. F. and Mizon G. E. 1978. Serial Correlation as Convenient Simplification, Not a Nuisance: A Comment on a Study of the Demand for Money by the Bank of England. The Economic Journal, vol. 88, Nº 351: 549-563.

5. Maddala, G. S. 1992. Introduction to Econometrics, Second Edition, Englewood Cliffs, New Jersey: Prentice-Hall.