Sheffield Economic Research Paper Series

SERP Number: 2003003

Peter Hayes and Paul Turner

**Estimating Quarterly GDP for the Interwar UK Economy:
An Application to the Employment Function**

March 2003

Department of Economics
University of Sheffield
9 Mappin Street
Sheffield
S1 4DT
United Kingdom
www.shef.ac.uk/economics

Abstract

Chow and Lin (1971) set out a procedure for the generation of higher frequency estimates for series for which data is available at a low frequency using data on a related series at the higher frequency. In this paper we set out a simple algorithm for the generation of quarterly estimates for a series for which annual data is available and quarterly data is available for the related series. We apply this to data for interwar Gross Domestic Product using Industrial Production as the related series. Using this approach we generate quarterly GDP figures for the period 1920.1 to 1938.4. This series is valuable in that it can be used to estimate a cointegrating relationship between employment, real wages and aggregate output which is not possible when we use industrial production directly as our quarterly measure of aggregate demand.

Keywords: Chow-Lin procedure, interwar GDP, employment function.

JEL Numbers: C22, N14

Correspondence Address:

Dr P. M. Turner, Department of Economics, University of Sheffield, 9 Mappin Street, Sheffield, S1 4DT, UNITED KINGDOM, Tel: 0114 2223404, e-mail: p.turner@sheffield.ac.uk

I. INTRODUCTION

The problem discussed in this paper is one which frequently arises in applied econometrics. Suppose we have annual data on some variable of interest and quarterly data on some related variable or variables. We wish to use this data to generate quarterly estimates for the variable of interest. Chow and Lin (1971) consider this problem in some detail and show how a relatively simple procedure can be implemented to generate either interpolated or distributed estimates for the series of interest. In this paper we employ their method to show how we can generate efficient estimates of the parameters linking annual averaged data to a related quarterly series. We also demonstrate the utility of this method using as an example the relationship between UK employment and Gross Domestic Product for the interwar period.

We divide the Chow-Lin procedure into two stages. In the first we estimate the parameters linking the annual averages of the series of interest to the quarterly data on the related series. In this second we used these estimates to generate quarterly estimates of the series of interest. We then present an example of this approach and show how it can be used to generate improved estimates of a quarterly employment function for the interwar UK economy.

The plan of the paper is as follows. In section II we discuss the application of the Chow-Lin model to quarterly data and the algorithm used to estimate the parameter of interest. Section III presents an application of the procedure to the generation of quarterly estimates of interwar UK GDP. Finally, section IV presents our conclusions.

II. GENERATING QUARTERLY ESTIMATES USING ANNUAL AVERAGED DATA

We begin by assuming that there exists a quarterly relationship between a variable y of interest and some other variable x. This relationship takes the form:

$$y_j = \beta x_j + u_j \tag{1}$$

where $j = 1, 2, \ldots 4T$ is a time index and T is the number of years in the sample. u is a stochastic error term with known properties. In vector notation we have $y = x\beta + u$ where y, x and u are $4T \times 1$ column vectors.

The problem facing the investigator is that, although quarterly data is available for the related series x, it is not available for the series of interest y. Suppose the data available for the series of interest consists of annual averages of the underlying quarterly data. We can define the relationship between the annual average data and the underlying quarterly series by using the $T \times 4T$ distribution matrix C_D whose elements are defined as follows:

$$C_D(i, j) = \begin{cases} \frac{1}{4} \text{ for } j = 4(i-1) + k \;\; k = 1, 2, 3, 4 \\ 0 \text{ otherwise} \end{cases} \tag{2}$$

Now consider the annual model obtained by premultiplying (1) by the matrix C_D. We have:

$$y_i^A = \beta x_i^A + u_i^A \tag{3}$$

where $y^A = C_D y$, $x^A = C_D x$ and $u^A = C_D u$ are $T \times 1$ column vectors and the index $i = 1, 2 \ldots T$ picks out the individual annual observations. The nature of the efficient estimator of the parameter β will depend on the properties of the transformed error term u^A which in turn depend on the properties of the underlying quarterly error term u. Let $V = E(u u')$ be the variance-covariance matrix of the errors for the quarterly model. It follows that the variance-covariance for the annual model (3) can be derived as $V^A = C_D V C_D{}'$.

If we make the assumption that the quarterly error terms are zero mean, homoscedastic, serially independent random variables then the estimation problem is relatively easy. It is straightforward to show that if $V = \sigma^2 I_{4T}$ where I_{4T} is a $4T \times 4T$ identity matrix then $V^A = \frac{\sigma^2}{4} I_T$ where I_T is a $T \times T$ identity matrix. We can therefore estimate (3) efficiently using ordinary least squares.

A more complicated case arises when the errors in the quarterly model follow an autoregressive process. For example, consider the case where the errors follow an AR(1) process. That is we have:

$$u_j = a u_{j-1} + \varepsilon_j \qquad (4)$$

where ε is a zero mean, homoscedastic, serially independent stochastic process and we have $|a| < 1$. The subscript j is used to pick out the individual quarters. We can derive the following sequence of autocovariances for the u^A process:

$$
\begin{aligned}
\operatorname{var}\left(u_i^A\right) &= \frac{\sigma_\varepsilon^2}{16\left(1-a^2\right)}\left(4+6a+4a^2+2a^3\right) \\
\operatorname{cov}\left(u_i^A u_{i-1}^A\right) &= \frac{\sigma_\varepsilon^2}{16\left(1-a^2\right)}\left(a+2a^2+3a^3+4a^4+3a^5+2a^6+a^7\right) \\
\operatorname{cov}\left(u_i^A u_{i-k}^A\right) &= a^4 \operatorname{cov}\left(u_i^A u_{i-k+1}^A\right) \qquad\qquad k \geq 2
\end{aligned}
\tag{5}
$$

Therefore the error term for the annual model does not follow a simple AR(1) process even if that for the underlying quarterly model does have this property. However, if the autoregressive parameter a is known then it is possible to use a generalised least squares (GLS) estimator to obtain an efficient estimate of the unknown parameter β. Of course the parameter a is rarely known in practice and therefore we must make use of a feasible GLS estimator to obtain an estimate of the parameter of interest.

The algorithm we use to estimate the model is as follows:

1. For some estimate of the autoregressive parameter a^g construct the matrix V^A and estimate (3) by GLS.

2. Generate the residuals from step 1 and construct q = ratio of the first sample autocovariance to the sample variance. Then solve for the value of a which satisfies:

$$
q = \frac{a+2a^2+3a^3+4a^4+3a^5+2a^6+a^7}{4+6a+4a^2+2a^3}
$$

3. If $\left|a-a^g\right| < TOL$ where TOL is some stopping tolerance then stop and use the current value $\hat{\beta}$ as the estimate of the unknown parameter. Otherwise set $a^g = a$ and return to step 1.

The problem of distribution can be thought of as one of choosing a vector of estimates $\hat{y}$ such that the trace of the covariance matrix $\operatorname{cov}\left(\hat{y}-y\right)$ is minimised subject to the constraints $C_D\hat{y}=y^A$. Chow and Lin show that the best linear unbiased estimator of y in this sense is given by the following expression:

$$\hat{y}=x\hat{\beta}+VC_D'\left(C_DVC_D'\right)^{-1}\hat{u}^A \tag{6}$$

where $\hat{u}^A=y^A-x^A\hat{\beta}$ is the vector of residuals from the annual model. In this case where $V=\sigma^2 I_{4T}$ this reduces to the very simple formula in which the estimated values of the y series are the fitted values obtained by multiplying the related series x by the estimated parameter $\hat{\beta}$ plus one quarter of the residual from the annual regression equation. This becomes somewhat more complicated when there is serial correlation in the underlying relationship.

III. AN APPLICATION TO INTERWAR GDP

Using this procedure set out in the previous section we now construct quarterly estimates of Gross Domestic Product (GDP) for the UK economy over the period 1920 to 1938. While annual GDP data are readily available c.f. Feinstein (1972), quarterly data are not generally available. However, there are numerous related series such as industrial production which are available at a quarterly or even higher frequency c.f. Capie and Collins (1983).

Equation (7) gives estimates for the relationship between Gross Domestic Product (GDP) and Industrial Production (IP) for the UK economy based on annual data from 1920 to

1938. GDP figures are taken from Feinstein and the industrial production figures are annual averages of the monthly data given in Capie and Collins. The parameter estimates are sensible with an elasticity of GDP with respect to industrial production of about 0.27. The positive time trend captures the fact that the service sector tended to grow faster than the industrial sector during this period. The implicit value of the autoregressive parameter for the quarterly relationship can be calculated as $a = 0.8114$ while the ratio of the first autocovariance to the variance of the annual model is $q = 0.5825$. The variance-covariance matrix used to calculate the t-ratios given in parentheses below the coefficients is given by the GLS formula $\hat{\sigma}^2 \left(X^{A\prime} \left(\hat{V}^A \right)^{-1} X^A \right)^{-1}$. The results, given in equation (7) indicate a significant effect from both the industrial production series and the time trend:

$$\ln\left(GDP\right)_t = \underset{(6.91)}{3.2342} + \underset{(2.64)}{0.2701} \ln\left(IP\right)_t + \underset{(3.69)}{0.0173} t + \hat{u}_t^A \tag{7}$$

Table 1 gives the quarterly estimates of GDP obtained from the Chow-Lin procedure i.e using equation (6) and Figure 1 shows the relationship between the quarterly estimates and the corresponding annual data. Note that the annual figures are placed at the second quarter of each year while strictly they should be placed between the second and third quarters. Nevertheless Figure 1 confirms that the quarterly estimates track the trend given by the annual data closely while allowing for some quarterly variation estimated by movements in industrial production.

[Insert Table 1 and Figure 1 here]

The value of our quarterly GDP measure can be seen in terms of its role in generating an economically meaningful equilibrium employment relationship. Suppose we wish to estimate an employment function linking employment to a measure of aggregate output and the real wage rate. As a first approximation we will use the industrial production index as our measure of aggregate output. Estimation using quarterly data from 1924.I to 1938.IV yields the following results:

$$\ln \hat{E} = \underset{(39.6)}{6.9513} + \underset{(4.3)}{0.0978}\, D26 + \underset{(12.8)}{0.4815}\, \ln IP + \underset{(4.9)}{0.4203}\, \ln\left(\frac{W}{P}\right) \tag{8}$$
$$R^2 = 0.88 \qquad SEE = 0.0256 \qquad DW = 0.77$$

where E is aggregate employment, $D26$ is a dummy variable designed to capture the effects of the General Strike which equals 1 for the quarters 1926.II to 1926.IV and 0 for other time periods, IP is the industrial production index, W is the wage rate and P is the retail price index. The t-ratios given in parentheses below the coefficients are calculated using the Newey-West heteroscedasticity and autocorrelation consistent coefficient covariance matrix.

The equilibrium relationship given in equation (8) is problematic both in terms of its statistical and economic properties. In statistical terms, we find that we cannot reject the null hypothesis that the residuals from this equation are integrated of order 1. Applying the augmented Dickey-Fuller test to the residuals, we obtain a test statistic of –3.25 (we set the number of lagged differenced terms at 0 as indicated by the Schwarz criterion). This compares with a 1% critical value of –4.54 and a 5% critical value of –3.89 as determined by the MacKinnon (1991) response surfaces. In addition it is hard to interpret (8) sensibly in terms of economic theory. The coefficient on the real wage rate is positive when we would expect a negative effect and even the sign on the General Strike dummy variable is opposite to what we would expect.

As an alternative we replaced industrial production with our quarterly GDP measure and obtained the following results:

$$\ln \hat{E} = \underset{(50.0)}{5.3272} - \underset{(-6.1)}{0.0484}\, D26 + \underset{(36.2)}{0.8384} \ln GDP - \underset{(-11.8)}{0.4151} \ln\left(\frac{W}{P}\right) \tag{9}$$

$$R^2 = 0.97 \qquad SEE = 0.0125 \qquad DW = 1.25$$

This specification of the model proves superior both in terms of its statistical and economic properties. The t-ratios are again calculated using the Newey-West standard errors. Applying the ADF test to the residuals of this model produces a test statistic equal to –4.92. Therefore in this case we reject the null of a unit root in the residuals at both the 5% and 1% significance levels. In terms of economic theory we now observe that the real wage elasticity now has the correct negative sign and the coefficient on the General Strike dummy variable also has its expected negative sign. Therefore the use of the quarterly GDP has produced a potentially useful cointegrating vector linking employment to GDP and the real wage rate.

Since the Engle-Granger cointegration test has established that an equilibrium relationship exists, it is reasonable to make use of this to estimate a dynamic model for aggregate employment. We began with a general autoregressive distributed lag model with four lags on employment, GDP and the real wage and conducted a specification search to obtain a parsimonious model. This model was then reparameterised into error correction form to obtain the equation given in (10). The equation was estimated by non-linear least squares so as to identify the equilibrium elasticities directly as well as the adjustment coefficient which measures the speed at which disequilibrium is eliminated.

$$\begin{aligned} \Delta \ln \hat{E}_t = {} & \underset{(4.52)}{3.02} - \underset{(2.55)}{0.025}\, D26 + \underset{(10.66)}{0.81} \Delta \ln GDP_t \\ & - \underset{(4.63)}{0.57} \left(\ln E_{t-1} - \underset{(22.18)}{0.85} \ln GDP_{t-1} + \underset{(6.78)}{0.39} \ln\left(\frac{W}{P}\right)_{t-1} \right) \end{aligned} \tag{10}$$

$$R^2 = 0.78 \qquad \hat{\sigma} = 0.0117 \qquad DW = 1.78 \qquad LM_4 = 0.75(0.56)$$

$$ARCH = 0.46(0.50) \quad NORM = 1.16(0.56)$$

R^2 is the coefficient of determination, $\hat{\sigma}$ is the standard error of the regression, *DW* is the Durbin-Watson statistic, LM_4 is the F-form of the Lagrange Multiplier test for 4th order serial correlation in the residuals, *ARCH* is the F-form of the Lagrange Multiplier test for 1st order autoregressive conditional heteroscedasticity in the residuals and *NORM* is the Jarque-Bera test statistic for non-normally distributed residuals. Numbers in parentheses are the p-values for the various test statistics.

The error correction model reported in equation (10) has good statistical properties. None of the standard diagnostic test statistics indicates the presence of misspecification and the parameter estimates are consistent with economic theory. Note also that the adjustment coefficient takes a value of 0.57 indicating that over half the disequilibrium in the level of employment is eliminated in each quarter. The fact that this coefficient is statistically significant also reinforces the conclusion that there exists an equilibrium relationship between employment, aggregate output and the real wage for the period considered.

It is interesting to compare our results with existing work based on annual data. A useful summary of this work can be found in Broadberry (1986). Hatton (1983) estimates separate equations linking employment to output and the real wage rate using annual data for the period 1921-1938. His estimates of the elasticities of employment with respect to output and the real wage rate are 0.59 and –0.79 respectively which compare with our estimates of 0.85 and –0.39. We can obtain an estimate of returns to scale by inverting the output elasticity to obtain the percentage response of output to a 1% increase in employment. Hatton's results yield an output elasticity with respect to employment of 1.69 which indicates strongly increasing returns. Our estimate of this elasticity is 1.18 which still indicates increasing returns but at a rather more modest rate than Hatton's estimate. Dimsdale (1984) estimates an equation which links employment to the real wage rate and variables intended to proxy the level of aggregate demand. His estimate of the real wage elasticity is –0.71. Thus our estimate of the real wage elasticity is rather lower than that obtained in these papers. The most likely reason for this is that our equation includes output as an explanatory variable. Thus the real wage variable captures only the effects of changes in the real wage on the substitution between labour and capital for a given level of output and does not include any induced effects of changes in real wages on the optimum level of output itself.

IV. CONCLUSIONS

In this article we have shown how the Chow-Lin procedure can be implemented to obtain quarterly estimates of a variable for which only annual average data are available using data on a related variable for which quarterly data is available. By using an iterative generalised least squares approach we obtain efficient estimates of the parameters linking the two series. This approach has then been applied to generate quarterly estimates of UK GDP for the period 1920.I to 1938.IV using industrial production as the related variable. We then demonstrate the value of this procedure by showing that the estimated series enables us to estimate an economically meaningful cointegrating vector linking employment to aggregate output and the real wage rate. Finally, we show that the estimated GDP series can be used to estimate a quarterly error correction model for employment which yields plausible econometric results.

References

Broadberry, S. N. (1986) *The British Economy Between the Wars: A Macroeconomic Survey*, Oxford: Basil Blackwell.

Chow, G.C. and Lin, A.L. (1971) 'Best Linear Unbiased Interpolation, Distribution, and Extrapolation of Time Series by Related Series', *Review of Economics and Statistics*, Vol 53, No 4, pp. 372-375.

Capie, F. and Collins, M. (1983) *The Inter-War British Economy: A Statistical Abstract*, Manchester: Manchester University Press.

Dimsdale, N. H. (1984) 'Employment and Real Wages in the Interwar Period', *National Institute Economic Review*, No 110, pp. 94-103.

Feinstein, C.H. (1972) *National Income, Expenditure and Output of the UK 1988-1965*, Cambridge: Cambridge University Press.

Hatton, T. J. (1983) 'Unemployment Benefits and the Macroeconomics of the Interwar Labour Market', Oxford Economic Papers, Vol. 35, pp. 486-505.

Table 1: Quarterly GDP Estimates 1920.I to 1938.IV

1920	I	94.4	1927	I	105.4	1934	I	110.8
	II	94.9		II	105.5		II	113.1
	III	95.2		III	105.0		III	112.6
	IV	90.3		IV	104.1		IV	113.9
1921	I	84.9	1928	I	104.2	1935	I	113.8
	II	75.0		II	105.9		II	117.9
	III	85.3		III	104.3		III	118.2
	IV	84.4		IV	105.6		IV	120.6
1922	I	85.4	1929	I	105.6	1936	I	121.1
	II	88.3		II	108.3		II	124.5
	III	91.1		III	108.3		III	124.8
	IV	89.3		IV	108.5		IV	125.6
1923	I	90.1	1930	I	107.6	1937	I	125.2
	II	91.5		II	107.3		II	129.8
	III	91.0		III	105.1		III	129.7
	IV	93.0		IV	104.1		IV	130.2
1924	I	93.2	1931	I	101.6	1938	I	128.9
	II	96.2		II	101.7		II	127.4
	III	97.1		III	102.3		III	125.4
	IV	97.6		IV	103.5		IV	126.7
1925	I	97.6	1932	I	103.3			
	II	98.8		II	102.3			
	III	98.4		III	101.3			
	IV	100.9		IV	103.2			
1926	I	103.3	1933	I	104.0			
	II	96.9		II	106.6			
	III	91.9		III	107.6			
	IV	94.6		IV	109.1			

Original GDP Figures are Output Data at constant factor cost 1913=100 taken from Feinstein (1972) Table 6.

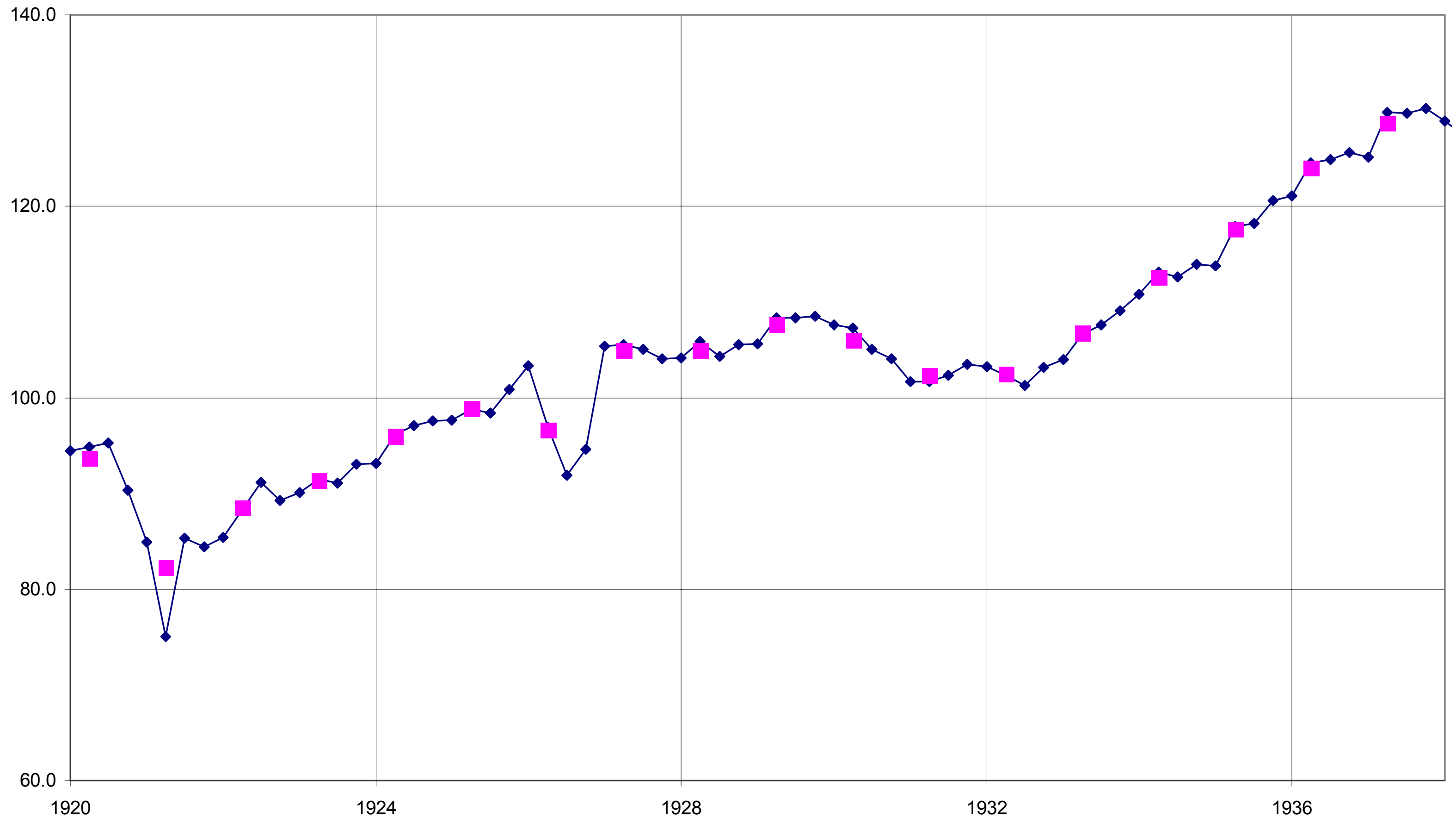

Figure 1: Quarterly GDP Estimates – (Blocks show annual figures)