QUARTERLY MODELS OF WAGE DETERMINATION: SOME NEW EFFICIENT ESTIMATES

J.C.R. Rowley  D.A. Wilton
Queen’s University  Queen’s University

Department of Economics
Queen’s University
94 University Avenue
Kingston, Ontario, Canada
K7L 3N6

6-1971
QUARTERLY MODELS OF WAGE DETERMINATION:

SOME NEW EFFICIENT ESTIMATES

J.C.R. Rowley and D.A. Wilton
Queen's University

Discussion Paper No. 51

June 1971
The concept of stable Phillips' curves and their economic implications have been the subject of a series of disputes in recent years. Since Phillips introduced his simple disequilibrium model, empirical research in the field of wage determination has indicated a wide spectrum of explanatory variables which expands and contracts almost randomly in published studies. Apart from this difficulty in specification, the intertemporal instability of estimated coefficients for Phillips' curve variants is particularly disturbing.\(^1\) In addition, the theoretical bases for many of these variants are unclear and several economists have participated in a neo-classical counter-attack on the existence and stability of Phillips' curves from a theoretical viewpoint.\(^2\)

The objectives of this paper are to examine the consequences of certain invalid statistical procedures which are employed in many of these studies and to provide some empirical evidence of their effect. In particular, we focus attention on the implications of the aggregative procedure which provides the basis for the use of a "four quarter overlapping change" representation of the dependent variable in quarterly studies of wage determination and, also, for the use of moving averages of explanatory variables. A partial list of studies which make use of these specifications would include the following works: Anderson (1969); Bodkin et al. (1966), de Menil (1969), Dick-Mireaux and Dow (1959), Eckstein (1968), Evans and Klein (1968), Helliwell et al. (1969), Klein and Ball (1959), Kuh (1967), Levy (1967), Lipsey and Parkin (1970), Perry (1966), Pierson (1968), Reuber (1970), Schultze and Tryon (1965), Simler and Tella (1968).

Although these forms of variables are predominant in many studies,
the implications of the aggregative procedure for the error terms are usually either completely neglected or represented by simple statements of "possible" problems which are not discussed further. Kuh's assertion that "... artificial serial correlation is often introduced through the overlapping process which partially neutralizes the confidence to be placed in the larger t-coefficients" (Kuh, 1967, p. 346) is rare in its explicit reference to the problem of autocorrelation. In the first section of the paper we present a set of assumptions sufficient for the specification of the "four quarter overlapping change" model in wage analysis. Serial correlation in the error term is shown to be a necessary consequence of the model. However, if the collective explanatory variables can be calculated, then generalized least squares (GLS) estimators can likewise be calculated for the coefficients as well as ordinary least squares (OLS) estimators. These two sets of estimators provide the data for our assessment of the effect of the aggregative procedure.

Necessary and sufficient conditions for least-squares estimators to be fully efficient despite the presence of autocorrelation, so that both standard errors and Student's t-statistics are unaffected by it, have been established by Kruskal (1968), Mitra and Rao (1969), Rao (1967, 1968) and Zyslind (1962, 1967). These statistical results need to be supplemented in particular analyses by empirical results since they indicate that optimality of least-squares procedures may depend upon both the particular sample to be used and the weights adopted for the aggregative procedure. If the least-squares estimates of coefficients are not optimal, conventional estimates of standard errors and Student's t-statistics are based on inappropriate formulae. However, the parametric estimates remain unbiased
and the inaccuracies of these formulae may be small. Their extent must be established empirically, although Swindel (1968) and Watson (1955, 1967) provide theoretical accounts of bounds on these inaccuracies. Three tables, which contain OLS and GLS results for three different studies of wage determination, reveal that the use of these formulae leads to substantial errors in inference.

As a basis for discussion, the study by Perry provides an appropriate framework since it is one of the earliest and, perhaps, best elucidated studies on the use of overlapping annual changes in the context of quarterly data. We have presented a complete analysis of the general model of discontinuous wage changes elsewhere so only a brief statement of the set of sufficient conditions for the four quarter overlapping change model will be presented below. Most of these assumptions are either explicitly or implicitly stated by Perry.

The Model of Wage Determination

(A1) Wages are set annually for all workers, and, once established, remain fixed until the next annual negotiation and settlement.

(A2) The labour force is divided into four distinct groups on the basis of the quarter in which their annual wage negotiations and settlements take place.

(A3) The ratios of all seasonal groups in the labour force to the total labour force are constant.

(A4) The percentage change in wages for each of the four seasonal groups is a function of the same set of explanatory variables with the
same parameter values for each group. Explanatory variables \((X)\) and error term \((u)\) are dated in the quarter in which the wage negotiation and settlement took place \((j)\). That is,

\[
\frac{w^h_j - w^h_{j-4}}{w^h_{j-4}} = aX_j + u_j
\]

for \(h = 1, \ldots, 4\) and \(j = h + 4s\), where \(s\) is an integer and \(w^h_j\) is the wage-rate for the \(h\)-th group in the \(j\)-th quarter.

(A5) The relative change in the aggregate wage-rate is appropriately approximated by a moving average for the relative changes in the wage-rates of the four groups. The weights of this moving average are equal. Without loss of generality, we give them unit values.

\[
\frac{w_t - w_{t-4}}{w_{t-4}} = \frac{4}{\sum_{h=1}^{4}} \frac{w^h_t - w^h_{t-4}}{w^h_{t-4}}
\]

\[
\frac{w_t - w_{t-4}}{w_{t-4}} = a \left( \sum_{i=0}^{3} X_{t-i} \right) + \sum_{i=0}^{3} u_{t-i}
\]

where \(w_t\) is the aggregate wage-rate.

For statistical purposes, the model is completed by the specification of a list of explanatory variables and a distribution for the error term. Ignoring this latter ingredient, the final equation is the familiar one, with the overlapping annual wage change form of the dependent variable and fourth order moving averages for the explanatory variables. However,
the error term is likewise distributed as a moving average process with the same known weights. Let \( \{ u_t \} \) be a sequence of normally distributed errors which are mutually independent and have constant, equal variances. With this additional assumption and the knowledge of the known weights, the method of generalized least squares can be employed to calculate efficient estimates of the coefficients of the explanatory variables. Further, knowledge of the weights indicates a transformation which can be used to adjust the equation so that the problem of autocorrelation is eliminated.\(^6\) Then, the appropriate estimates of the standard errors of estimated coefficients (based on unbiased estimates of their variances) and Student's t-statistics can be calculated. In the absence of this adjustment, the OLS method yields biased estimates of standard errors and invalid Student's t-statistics unless the sample fulfills the conditions given in the references cited above. Calculation of GLS\(^7\) estimates permits us to observe how badly the OLS estimates are for the hypothetical model with a given sample. Note that even if this distribution of \( \{ u_t \} \) is inappropriate, as would be the case if the original errors for individual groups were autocorrelated, the results would remain of interest since they reveal the sensitivities of estimates to different specifications of the distribution for errors. Finally, notice that the aggregate error is not generated by a first order moving average process so the Durbin-Watson statistic is not an appropriate indicator of serial correlation and applications of the popular Hildreth-Lu procedure "to correct for serial correlation" will not do so in fact.
Empirical Results

Results for the model of Perry are supplemented below by additional estimates for the model of Bodkin et al. (1966) and the wage equation contained in the econometric model (RDXI) of the Bank of Canada. The role anticipated for these supplemental results stems from the danger that the inadequacies of empirical results for a single study might prove atypical. Although the use of three studies cannot eliminate this danger, we can present conclusions with greater confidence than would be the case with a single study.

Each of the following three tables are divided vertically into two sections. The first section contains OLS estimates based on inappropriate formulae for standard error and Student's t-statistics (columns (1), (2), (3) for Table One, columns (1), (2), (3), (4) for Table Two and column (1) for Table Three), whereas the second section contains the appropriate GLS estimates on the assumption that the errors follow the distribution cited. Each cell in the tables contains three numbers; the estimated coefficient, the estimated standard error of this coefficient and a Student's t-statistic. To clarify exposition, the calculated numbers for the OLS estimates of standard errors and Student's t-statistics are prefixed by the term "pseudo" since we have reason to believe that the former are based on biased estimates of variances and the latter on a mis-specification of the distribution for the errors.

Two questions should be considered while the empirical results are read. First, would the set of OLS estimates lead the (typical) economist to draw invalid inferences with respect to the significance of variables in
Table 1. PERRY Wage Equations for U.S. Manufacturing Industry (1948 II - 1960 III)

<table>
<thead>
<tr>
<th></th>
<th>Ordinary Least Squares</th>
<th>Generalized Least Squares</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>( \hat{C}_{t-1} )</td>
<td>0.385</td>
<td>0.327</td>
</tr>
<tr>
<td></td>
<td>0.055</td>
<td>0.060</td>
</tr>
<tr>
<td></td>
<td>6.946</td>
<td>5.487</td>
</tr>
<tr>
<td></td>
<td>2.176</td>
<td>2.220</td>
</tr>
<tr>
<td></td>
<td>6.716</td>
<td>4.659</td>
</tr>
<tr>
<td>( R_{t-1} )</td>
<td>0.434</td>
<td>0.524</td>
</tr>
<tr>
<td></td>
<td>0.069</td>
<td>0.074</td>
</tr>
<tr>
<td></td>
<td>6.270</td>
<td>7.058</td>
</tr>
<tr>
<td>( \Delta R_t )</td>
<td>0.832</td>
<td>0.714</td>
</tr>
<tr>
<td></td>
<td>0.176</td>
<td>0.202</td>
</tr>
<tr>
<td></td>
<td>4.727</td>
<td>3.529</td>
</tr>
<tr>
<td>D.W.</td>
<td>1.188</td>
<td>0.811</td>
</tr>
<tr>
<td>F(4,44)</td>
<td>71.706</td>
<td>62.825</td>
</tr>
</tbody>
</table>
VARIABLE DEFINITION FOR THE PERRY MODEL

Dependent variable: annual percentage change in straight-time hourly earnings of production workers for total, durable and non-durable manufacturing, \( \left( \frac{W_t - W_{t-4}}{W_{t-4}} \right) \).

\( \dot{C}_{t-1} \) : four quarter moving average of one quarter percentage change in the consumer price index \( \left( \sum_{i=0}^{3} \frac{C_{t-i} - C_{t-1-i}}{C_{t-1-i}} \right) \), lagged one quarter.

\( 1/U_t \) : reciprocal of the four quarter moving average of the unemployment rate.

\( R_{t-1} \) : four quarter moving average of the annual profit rate (ratio of corporate earnings after taxes to stockholders equity), lagged one quarter, for total, durable and non-durable manufacturing.

\( \Delta R_t \) : first difference of the profit rate series.

\{ (1), (4) \} Total Manufacturing

\{ (2), (5) \} Non-Durables Manufacturing

\{ (3), (6) \} Durables Manufacturing
Table 2. BODKIN et al. Wage Equations for Canadian Manufacturing Industry (1953 I - 1965 II)

<table>
<thead>
<tr>
<th></th>
<th>Ordinary Least Squares</th>
<th>Generalized Least Squares</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>$P_t$</td>
<td>0.377</td>
<td>0.458</td>
</tr>
<tr>
<td></td>
<td>0.077</td>
<td>0.086</td>
</tr>
<tr>
<td></td>
<td>4.881</td>
<td>5.354</td>
</tr>
<tr>
<td></td>
<td>1.553</td>
<td>3.344</td>
</tr>
<tr>
<td>$(Z/Q)_{t-2}$</td>
<td>0.053</td>
<td>0.066</td>
</tr>
<tr>
<td></td>
<td>0.018</td>
<td>0.020</td>
</tr>
<tr>
<td></td>
<td>2.947</td>
<td>3.292</td>
</tr>
<tr>
<td>$w_{us_t}$</td>
<td>0.432</td>
<td>0.495</td>
</tr>
<tr>
<td></td>
<td>0.109</td>
<td>0.116</td>
</tr>
<tr>
<td></td>
<td>3.948</td>
<td>4.261</td>
</tr>
<tr>
<td>$w_{t-4}$</td>
<td>-0.092</td>
<td>-0.124</td>
</tr>
<tr>
<td></td>
<td>0.040</td>
<td>0.045</td>
</tr>
<tr>
<td></td>
<td>-2.279</td>
<td>-2.740</td>
</tr>
<tr>
<td>Const.</td>
<td>-4.122</td>
<td>-4.566</td>
</tr>
<tr>
<td>D.W.</td>
<td>1.45</td>
<td>1.11</td>
</tr>
<tr>
<td>$F(,)$</td>
<td>44.335</td>
<td>38.907</td>
</tr>
</tbody>
</table>
VARIABLE DEFINITION FOR THE BODKIN et al. STUDY

Dependent variable: annual percentage change in average hourly earnings of production workers in Canadian manufacturing industry,

\[ \frac{(W_t - W_{t-4})}{W_{t-4}} \].

\( \hat{P}_t \) : four quarter moving average of annual percentage change in consumer price index,

\[ \frac{1}{4} \sum_{i=0}^{3} \left( \frac{P_{t-i} - P_{t-4-i}}{P_{t-4-i}} \right) \].

\( 1/U_{t}^2 \) : squared reciprocal of a four quarter moving average of a two quarter average of the Canadian unemployment rate,

\[ \left[ \frac{1}{8} U_t + \frac{1}{4} \sum_{i=0}^{3} U_{t-i} + \frac{1}{8} U_{t-4} \right]^{-2} \]

\( (Z/Q)_{t-2} \) : four quarter moving average of the profit markup on output (index of corporate profits before tax divided by manufacturing production index), lagged two quarters.

\( \dot{W}_{us t} \) : four quarter moving average of the annual percentage change in average hourly earnings in U. S. manufacturing expressed in U. S. dollars,

\[ \left( \frac{1}{4} \sum_{i=0}^{3} \frac{W_{us t-i} - W_{us t-4-i}}{W_{us t-4-i}} \right) \]

\( \dot{W}_{t-4} \) : dependent variable lagged four quarters.
Table 3. RDXI Wage Equation for Canadian Economy (1955 I - 1965 IV)

<table>
<thead>
<tr>
<th></th>
<th>Ordinary Least Squares</th>
<th>Generalized Least Squares</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>( p^*_t )</td>
<td>0.940</td>
<td>0.516</td>
</tr>
<tr>
<td></td>
<td>0.240</td>
<td>0.633</td>
</tr>
<tr>
<td></td>
<td>3.922</td>
<td>0.815</td>
</tr>
<tr>
<td>( 1/u^2_t )</td>
<td>0.521</td>
<td>0.309</td>
</tr>
<tr>
<td></td>
<td>0.113</td>
<td>0.389</td>
</tr>
<tr>
<td></td>
<td>4.590</td>
<td>0.795</td>
</tr>
<tr>
<td>( R_t )</td>
<td>0.867</td>
<td>1.104</td>
</tr>
<tr>
<td></td>
<td>0.365</td>
<td>1.029</td>
</tr>
<tr>
<td></td>
<td>2.372</td>
<td>1.072</td>
</tr>
<tr>
<td>( w_{t-4} )</td>
<td>-0.224</td>
<td>0.154</td>
</tr>
<tr>
<td></td>
<td>0.138</td>
<td>0.140</td>
</tr>
<tr>
<td></td>
<td>-1.624</td>
<td>1.099</td>
</tr>
<tr>
<td>Const.</td>
<td>-4.712</td>
<td>-6.284</td>
</tr>
<tr>
<td>D.W.</td>
<td>1.75</td>
<td>2.61</td>
</tr>
<tr>
<td>F(5,39)</td>
<td>111.805</td>
<td>7.343</td>
</tr>
</tbody>
</table>
VARIABLE DEFINITION FOR RDXI WAGE EQUATION

Dependent variable: annual percentage change in the average hourly wage in the private sector, \( [(W_t - W_{t-4}) / W_{t-4}] \).

\( \hat{p}_t^* \) : four quarter moving average of the annual change in the implicit price index of consumer non-durable expenditure,
\[
(3 \sum_{i=0}^{4} \frac{p_{t-i} - p_{t-4-i}}{p_{t-4-i}}).
\]

\( 1/U_t^2 \) : squared reciprocal of a four quarter moving average of the unemployment rate.

\( R_t \) : four quarter moving average of the ratio of corporate profits before taxes (less income tax accruals) to real gross national expenditure.

\( W_{t-4} \) : dependent variable lagged four quarters.
the models. Second, if these invalid inferences are drawn, are they sufficiently severe to suggest that the additional burden of computation for the GLS estimates is worthwhile.

Looking only at the OLS estimates in these three tables, one would conclude that these Phillips curve variants perform reasonably well. The three common explanatory variables, unemployment, price increases and profitability, are pseudo-significant in all cases save one (the unemployment variable in equation (1) of Table 2). With the possible exception of the Durbin-Watson statistic, all of the conventional statistical checks are satisfied, i.e. sign patterns, Student t-statistics greater than two, etc.

On the other hand, statistical inferences drawn from the GLS equations are largely inconclusive. While F-tests indicate that in all cases the joint hypothesis that all coefficients for explanatory variables are zero must be rejected (since critical .05 values for the F-statistic fall within the 2.4-3.5 range), the results for individual explanatory variables are not encouraging. Only the profit variable retains significance in even half of the GLS equations. Unemployment and price increases, perhaps the two most common explanatory variables in all wage research, are significant in only two of eight individual cases. One can only speculate whether the various authors would have advanced the Phillips curve model had they been faced with the GLS estimates rather than the OLS estimates.

Turning from individual explanatory variables to a comparison of the three studies, the Perry model holds up best under GLS, but again the results are not that encouraging. While all variables are pseudo-significant under OLS, there is no consistent pattern for variables under GLS. No
variable's coefficient retains significance in each of the three sectors. The basic Phillips curve variable is significant in only one sector while price and profit variables are significant in two of three sectors. In sectoral terms, no one equation exhibits significant coefficients for all four explanatory variables.

The GLS results for the two Canadian studies are perhaps more conclusive. None of the usual explanatory variables would appear to be reliable from usual statistical tests. In place of fifteen pseudo-significant coefficients in the Bodkin et al. study, only three parameter estimates are significant under GLS. Their basic model, as represented by equations (1) and (5) in Table 2, fails to retain one of the five pseudo-significant parameters. As shown in Table 3, all explanatory variables in the RDX1 wage equation likewise prove to be insignificant under GLS, again a revealing contrast from the OLS results. The only consolation rests with the F-statistic, even though it is greatly reduced under GLS. As a group the explanatory variables are significant at the .05 level.

In short, the potential bias in estimates of standard errors and Student t-statistics using conventional OLS techniques appears to be substantial. Employing the GLS technique, which produces unbiased estimates of the variances of estimated coefficients, all standard-error estimates exceed those obtained using OLS with more than two-thirds of the individual coefficient standard errors increasing by more than 100%. In terms of standard significance tests, the results are even more pronounced. All Student t-statistics fall under GLS as OLS pseudo-t-statistics are at least twice as large for over three quarters of all parameter estimates and at least three times as large for more than half of the parameter coefficients.
Two final comments can be made concerning the estimates. Since both OLS and GLS give unbiased parameter estimates, one might expect our coefficient estimates to be reasonably similar for both estimation techniques. Almost two-thirds (20 of 32) of the OLS parameter estimates are within one standard error of the GLS estimates. Finally, the appropriateness of the Durbin-Watson statistic is somewhat dubious. As pointed out above, it fails to test for a fourth order moving average process in the error term, as exists in the OLS equations. While the Durbin-Watson statistics are close to two under GLS, a fourth order autoregressive pattern (i.e. \( u_t = p u_{t-4} + e_t \), where \( e_t \) satisfies the well-known classical properties) would appear more plausible given the annual bargaining patterns in the four distinct labour groups.

**Conclusion**

A necessary first step in empirical research is the specification of the form of the dependent variable, that is, the sequence of values which are to be explained. Quarterly analyses of wage determination have adopted overlapping annual changes as the appropriate form of their dependent variable. Explicit recognition of the statistical implications of this choice is extremely rare. We have presented a set of assumptions which are sufficient to explain this particular specification for the dependent variable. Serial correlation in the error term is a direct consequence of the aggregative procedure indicated by these assumptions. In particular, if researchers specify the same moving average for all explanatory variables, then the aggregate error must exhibit the same moving average features. Hence, all of the conventional formulae for appropriate standard errors and test statistics are potentially biased,
with the extent of biases to be determined empirically.

For the three models which we have considered these biases are shown to be substantial. A comparison of Student's t-statistics based on correct formulae with the pseudo-results suggest that the latter must be reduced by at least fifty per cent in most cases. The inferential consequences of these reductions are severe and cannot be neglected. The roles of certain variables, which are usually considered to be of major importance in the determination of wage levels, are in doubt. For example, in the eight equations for which GLS estimates are calculated, unemployment is significant at the .05 level twice and not seven times as indicated by the results for the conventional OLS procedure. Similarly, the change in prices is significant twice at this critical level although the OLS results indicate eight times.

While there are dangers in attempts to generalize these results which are based on only three studies, every test indicates the potential severity of the problem of unacknowledged serial correlation. Fortunately, as we have demonstrated, there are simple ways to correct for this bias in estimation. Correctional techniques require few additional assumptions other than those which already form the bases for specifications of the dependent variable and moving averages of the explanatory variables. Only after these corrections are made will economists be able to assess the true significance of the alternative theories which have been advanced to explain the determination of aggregate wage changes.
FOOTNOTES

1. For example, in our attempt to reproduce the estimates of the study by Bodkin et al. (1966), we mistakenly added two additional observations to our initial sample. The estimated coefficient of the unemployment term increased from 10.4 to 17.3. For further evidence of the inherent instability of Phillips' curve variants see Levy (1967) and Perry (1966, pp. 72-84).

2. See the studies by E.S. Phelps et al. (1970).

3. There are exceptions. For example, see Dicks-Mireaux and Dow (1959), Sparks and Wilton (1969) and Sargan (1964). In addition, Kuh (1967) and Eckstein (1968) present models with both quarterly and overlapping annual changes in wage-rates which indirectly illustrate the nature of the problem. The use of quarterly changes avoids the problem of autocorrelation which we shall illustrate but introduces other difficulties. Students' t-statistics are much lower for models of quarterly changes than for the models of overlapping annual changes.


6. The basic details of the GLS method are presented in Rowley and Wilton (1971).

7. GLS estimates are seldom calculated even if we have precise knowledge of the process generating the errors because of their computational burden. Wilkinson (1965) demonstrates how this burden can be sub-
stantially reduced for the efficient Cholesky technique in our particular case of a moving average.

8. Our choice of studies was primarily dictated by the accessibility of data. In this regard, we wish to thank the Bank of Canada and Steve Kaliski for making data available to us. We would point out that both of these latter studies employ different wage data and time horizons from the study by Perry.
BIBLIOGRAPHY


KRUSKAL, W., "When are Gauss-Markov and Least Squares Estimators Identical? A Coordinate-Free Approach", Annals of Mathematical Statistics,


