Endpoint Constraints and the St. Louis Equation: A Clarification

Authors	Dallas S. Batten, and Daniel L. Thornton
Working Paper Number	1983-001A
Creation Date	January 1983
Citable Link	https://doi.org/10.20955/wp.1983.001
Suggested Citation	Batten, D.S., Thornton, D.L., 1983; Endpoint Constraints and the St. Louis Equation: A Clarification, Federal Reserve Bank of St. Louis Working Paper 1983-001. URL https://doi.org/10.20955/wp.1983.001

ENDPOINT CONSTRAINTS AND THE ST. LOUIS EQUATION: A CLARIFICATION

Daniel L. Thornton & Dallas S. Batten*
Federal Reserve Bank of St. Louis
83-001

*Economists, Federal Reserve Bank of St. Louis. The views expressed here do not necessarily reflect those of the Federal Reserve Bank or of the Board of Governors of the Federal Reserve System.

We would like to thank R. Carter Hill and Rik Hafer for their helpful comments and Sarah Driver for her research assistance. Any remaining errors are our own.

ID 5632R
DLT:SD:lmw
Disk 62R
Draft 9
12/9/82

In a recent article in this journal, Seaks and Allen [9] argue that the imposition of endpoint constraints suggested by Almon [1] should not be rejected *a priori* as Schmidt and Waud [8] advise, but can be tested using a standard F-test. They argue that through testing researchers can avoid the bias resulting from imposing invalid restrictions, yet realize the efficiency gain from employing them. To illustrate this point, Seaks and Allen test the endpoint constraints imposed in the St. Louis equation and find that the data do not reject them. As a result, they conclude that the imposition of endpoint constraints does not appear to result in biased coefficient estimates. Unfortunately, Seaks and Allen misinterpret Schmidt and Waud's warning and incorporate restrictions that they might not have otherwise employed.

Schmidt and Waud point out that if the endpoint restrictions are false, they result in biased estimates of the distributed lag weights, however, they result in more efficient parameter estimates. Moreover, they question the use of endpoint constraints since "no convincing reason has ever been advanced as to why these 'endpoint constraints' should be true."[1/] They note that the imposition of endpoint constraints actually places restrictions on the polynomial outside of its relevant range, and that the behavior of the polynomial outside the range is simply irrelevant. Thus, there is no reason to expect these restrictions to be true *a priori*, and hence, they should not be applied routinely.

For Schmidt and Waud, there are three factors to consider about endpoint constraints: bias, efficiency, and lack of theoretical justification. Endpoint constraints which cannot be rejected by a F-test are data-based restrictions which may have no relevance in either econometric or economic theory. Since the only gain from imposing these restrictions is increased efficiency, one might prefer not to impose them even if they cannot be rejected, especially if the potential gains in efficiency appear to be small. At the other extreme, however, one may wish to impose these restrictions even if they are false, if the efficiency gains appear to be substantial, i.e., the biased estimator may have a smaller mean square error (MSE) than the unbiased one. In the latter instance, one might prefer to use a modified F-test based on a MSE-norm, such as those proposed by Toro-Vizcarrondo and Wallace [10] and Wallace [12].

In effect, Schmidt and Waud argue that endpoint constraints be evaluated in terms of their plausibility before blindly imposing them, even if they cannot be rejected by an F-test. We believe that had this procedure been followed, Seaks and Allen may have been more circumspect about advocating the use of endpoint constraints. In what follows, we outline a procedure for evaluating endpoint constraints, using the same specification of the St. Louis equation as did Seaks and Allen and using their data.[3/]

The Model

The St. Louis equation relates the growth rate of (or change in) nominal GNP ($\dot{Y}$) to the growth rate of narrowly defined money ($\dot{M}$) and the growth rate of high-employment federal spending ($\dot{E}$). Seaks and Allen used the specification:

$$\dot{Y}_t = a + \sum_{i=0}^{4} m_i \dot{M}_{t-i} + \sum_{i=0}^{4} e_i \dot{E}_{t-i} + u_t, \qquad t=1, 2, \ldots T$$

where the fourth-order distributed lags of $\dot{M}$ and $\dot{E}$ are each assumed to lie on different fourth-degree polynomials. This equation was estimated for the period 1953/II - 1977/IV.

The use of the polynomial distributed lag (PDL) technique for this specification is redundant. Since the order of the distributed lags and the degree of the polynomials were assumed to be the same, there are no PDL constraints on the parameters.[4/] To see this, consider the matrix form of a general PDL model given below:

$$\underline{Y} = a + X\underline{\beta} + \underline{u},$$

where $\underline{Y}$ is a Tx1 vector of observations on the dependent variable, X is a Tx(ℓ +1) vector of observations on a ℓth-order distributed lag on the independent variable X, and $\underline{\beta}$ is a (ℓ +1)x1 vector of parameters, i.e., $\underline{\beta} = (\beta_0, \beta_1, \ldots, \beta_\ell)'$. The Tx1 vector of random disturbances is given by $\underline{u}$. A PDL model assumes that the distributed lag weights fall on a polynomial of degree p, i.e.,

$$\beta_i = \alpha_0 + \alpha_1 i + \alpha_2 i^2 + \ldots + \alpha_p i^p.$$

In more general matrix terms, the vector $\underline{\beta}$ is related to the vector $\underline{\alpha} = (\alpha_0, \alpha_1, \ldots, \alpha_p)'$ as:

$$\underline{\beta} = H\underline{\alpha},$$

where H is a $(\ell+1)$ by $(p+1)$ matrix of known coefficients.[5] In the general case, p is usually less than ℓ, so that $(\ell-p)$ homogeneous restrictions are being employed on $\underline{\beta}$; hence, the gain in efficiency. In Seaks and Allen's case, however, $\ell = p$, and H is nonsingular. Thus, OLS estimates and PDL estimates of $\underline{\beta}$ are identical. The only gains in efficiency come through imposing the endpoint constraints.

Endpoint Constraints

The endpoint constraints suggested by Almon are

$$\beta_{-1} = \beta_{\ell+1} = 0.$$

These homogeneous restrictions on the distributed lag weights outside of the relevant range of the polynomial imply homogeneous restrictions on the distributed lag weights inside the relevant range, via restrictions on $\underline{\alpha}$. To see this, note that the above restrictions imply

$$\alpha_0 + \alpha_1 (-1) + \alpha_2 (-1)^2 + \ldots + \alpha_p (-1)^p = 0$$

$$\alpha_0 + \alpha_1 (\ell+1) + \alpha_2 (\ell+1)^2 + \ldots + \alpha_p (\ell+1)^p = 0,$$

or

$$R\underline{\alpha} = 0.$$

Since $\alpha = H^{+}\underline{\beta}$, where H^{+} is $(H'H)^{-1} H'$,

$$R\underline{\alpha} = RH^{+}\underline{\beta} = R^{*}\underline{\beta} = 0.$$

Thus, the endpoint constraints imply a set of homogeneous restrictions, R*, on $\underline{\beta}$. It is easy to show that

$$R^{*} = \begin{bmatrix} 5 & -10 & 10 & -5 & 1 \\ 1 & -5 & 10 & -10 & 5 \end{bmatrix}$$

for each distributed lag variable of the above model.[6/] A priori it is difficult to imagine why one would wish to employ such restrictions. Nevertheless, there will be a gain in efficiency from imposing them.

At this point, it is desirable to consider two aspects of imposing these restrictions, other than their plausibility: the potential gain in efficiency and the likelihood that these restrictions will not be rejected by the data. Given that the only reason for imposing endpoint constraints is the potential efficiency gain, one would like to know if the endpoint constraints are optimal in some sense. Fomby, Hill and Johnson [4] have shown that the Principal Components restrictions result in the maximum percentage reduction in the trace of the covariance matrix of the restricted least squares estimator of $\underline{\beta}$, for a given number of linearly independent restrictions that are true. Furthermore, Fomby and Hill [5] have derived a formula for the maximum percentage gain in efficiency under this norm.[7/] This suggests that if one were to impose the endpoint constraints in a PDL model in order to increase the precision of the estimates of $\underline{\beta}$, one would impose them only if they spanned the space given by an equal number of true Principal Components restrictions, given the trace norm.

Of course, there may not be an equivalent number of Principal Components restrictions that cannot be rejected by an F-test, or it may be that one would prefer a different norm than the trace norm suggested by Fomby et. al.[8/] Nevertheless, this illustrates the questionableness of imposing endpoint constraints in a PDL model without first, specifying a loss function and second, investigating the potential gains in efficiency given this criterion. The latter point is particularly important given that Fomby and Hill's formula shows that the gain in efficiency from one Principal Components restriction that is true can be substantial in colinear data sets such as distributed lag models like the St. Louis equation.[9/]

Given two or more PDL variables in an equation, it is possible that the endpoint constraints may hold for one but not for other PDL variables. Thus, it is desirable to investigate the plausibility of these constraints for the individual PDL variables. This can be done in a rough fashion by determining what the restrictions imply for the unconstrained estimates of $\underline{\beta}$ or by comparing the restricted with the unrestricted estimates.[10/] In the latter instance, we would not expect the constrained estimates of $\underline{\beta}$ to differ significantly from the unconstrained ones if the constraint is not binding. When this investigation is undertaken for this version of the St. Louis equation, substantial changes in the parameter estimates for the constrained and unconstrained models appear (see table 1). These differences are especially notable for the estimated coefficients of the distributed lag of $\dot{M}$. The estimates of all five coefficients change dramatically when the endpoint constraints are imposed with the estimated coefficients of the first and second lags of $\dot{M}$ becoming statistically significant while those of the third and fourth lags become statistically insignificant. Thus, one might be suspicious that the endpoint constraints are not valid in the case of $\dot{M}$.

At a more formal level, each of the separate endpoint constraints could be tested, along with combinations of them. Care must be taken, however, in that this procedure gives rise to further problems of preliminary test estimation.[11] The results of these tests, which are reported in table 2, confirm our suspicions concerning the validity of the joint head and tail constraints on $\dot{M}$. The imposition of both head and tail constraints on $\dot{M}$ is rejected at the .05 percent level. Nevertheless, the test of all four endpoint constraints cannot be rejected.[12] In effect, the F-test on all four constraints confuses the effect of significant endpoint constraints with that of the insignificant ones. Thus, contrary to Seaks and Allen's conclusion, there appears to be evidence that the endpoint constraints on money growth do bias the estimated parameters of the distributed lag coefficients on $\dot{M}$.

Furthermore, it is important to note that the presence or absence of endpoint constraints has no influence on th- qualitative conclusions that a change in the rate of growth of the money stock has a significant long-run effect while a change in the rate of growth of government expenditures has no long-run effect on the rate of growth of nominal income. Both money and government expenditures, however, have significant short-run effects.

Conclusions

We have argued that the primary concern over imposing the endpoint constraints of a PDL model is that such constraints are not justified by either economic or econometric theory. Given this and the fact that such constraints will only improve the efficiency of the estimator, one might not wish to impose them. Also, it is likely that another set of restrictions will exist that, if true, will produce a larger reduction in the variance of the estimator than do the endpoint constraints. Thus, if the gain in efficiency is important, one would probably want to choose restrictions other than those imposed by the endpoint constraints. At a minimum, care should be taken to specify the loss function associated with the imposition of these constraints. In this instance, the constraints might be imposed even if they are false. It may be preferable, however, to impose homogeneous restrictions that have more of a basis in economic or econometric theory.

In any event, we have argued that it is desirable to write out the restrictions implied by the endpoint constraints to judge whether they are desirable and to determine whether they are likely to be obeyed by the data. Indeed, it may be desirable to perform separate F-tests on the endpoint constraints. We have showed that, had this procedure been followed by Seaks and Allen, the joint endpoint constraints on money growth would have been rejected. Seaks and Allen should not be criticized too severely, however, since most of the empirical work employing PDLs have followed Almon's suggestion and imposed endpoint constraints without testing them.

FOOTNOTES

1/ Schmidt and Waud [8, p. 12].

2/ In this regard, one might just as well impose any arbitrary set of homogeneous linear restrictions on the parameters. Presumably, by Seaks and Allen's criterion, the "best" set would be that which cannot be rejected by an F-test and which produces the greatest increase in efficiency. A similar problem exists for another data-based restricted least squares estimator which employs Principal Components. In this case, however, there is at least some statistical criterion for employing such restrictions. See, Dhrymes [3, pp. 53-65].

3/ We would like to thank Terry G. Seaks and Stuart D. Allen for supplying us with their data so that our estimates are directly comparable to theirs.

4/ In order to estimate a fourth-degree polynomial with no endpoint constraints, one must at least consider a fourth-order distributed lag.

5/ The H-matrix is of the general form

$$H = \begin{bmatrix} 1 & 0 & 0 & \cdots\cdots & 0 \\ 1 & 1 & 1 & \cdots\cdots & 1 \\ 1 & 2 & 2^2 & \cdots\cdots & 2^p \\ \cdot & & & & \\ \cdot & & & & \\ 1 & \ell & \ell^2 & \cdots\cdots & \ell^p \end{bmatrix}$$

6/ In the case where $\ell = p$, $H^+ = H^{-1}$, therefore $R^* = RH^{-1}$.

7/ The maximum gain in efficiency given by Fomby and Hill [5] is

$$V_p = \frac{\sum_{i=k-p+1}^{k} (1/\lambda_i)}{\sum_{i=1}^{k} (1/\lambda)_i} \cdot 100;$$

where V_p denotes the maximum efficiency gain for a set of p valid restrictions, and λ_i denotes the ith ordered eigenvalue of (X'X) going from the largest, λ_1, to the smallest, λ_k.

8/ For examples see Toro-Vizcarrondo and Wallace [10] and Wallace [12] or Judge, et. al., [6, pp. 644-47].

9/ The estimate of V_1 (see footnote 7 above) for one principle components restriction is 44.7 percent. An F-test of this restriction cannot reject the null hypothesis at the .01 level. The F-value was 1.07.

10/ Of course, if one considers the plausibility of these restrictions by looking at the unrestricted estimates of $\underline{\beta}$, one must be careful to consider the covariance of $\underline{\beta}$. Furthermore, the head and tail constraints together imply restrictions which may not be apparent from looking at the individual constraints alone. That is, if R_1 and R_2 denote the first and second rows of R* respectively, then $R'_3 = \delta_1 R'_1 + \delta_2 R'_2$ (for all δ_1 and δ_2 such that both are not zero) and must lie in the space spanned by R_1 and R_2.

11/ For a general discussion of these problems see Batten and Thornton [2].

12/ Given the problems with preliminary test estimates, results by Toyoda and Wallace [11] and Sawa and Hiromatsu [7] indicate that one might want to choose a much larger α-level than is usually chosen in order to minimize the risk function. Toyoda and Wallace recommend the unrestricted estimator if fewer than five restrictiuons are imposed. Nevertheless, if one chooses a critical value for F of 2.0 as they suggest if the number of restrictions is five or more, the test of all four endpoint constraints would be rejected.

REFERENCES

[1] Almon, S., "The Distributed Lag Between Capital Appropriations and Expenditures," Econometrica, January 1965, pp. 178-96.

[2] Batten, Dallas S. and Daniel L. Thornton, "Polynomial Distributed Lags and the St. Louis Equation," Review Federal Reserve Bank of St. Louis, forthcoming.

[3] Dhrymes, Phoebus J., Econometrics, Harper and Row, NY, 1970.

[4] Fomby, Thomas B., R. Carter Hill and Stan R. Johnson, "An Optimal Property of Principal Components in the Context of Restricted Least Squares," Journal of the American Statistical Association, March 1978, pp. 191-93.

[5] Fomby, Thomas B. and R. Carter Hill, "Multi-Colinearity and the Value of A Priori Information," Communications in Statistics - Theory and Methods, 1974, pp. 477-86.

[6] Judge, George G., William E. Griffiths, R. Carter Hill, and Tsoung-Chao Lee, The Theory and Practice of Econometrics, John Wiley and Sons, NY, 1980.

[7] Sawa, Takamitsu and Takeshi Hiromatsu, "Minimax Regret Significance Points for a Preliminary Test in Regression Analysis," Econometrica, November 1973, pp. 1093-1101.

[8] Schmidt, Peter and Roger N. Waud, "The Almon Distributed Lag Technique and the Monetary Versus Fiscal Policy Debate," Journal of the American Statistical Association, March 1973, pp. 11-19.

[9] Seaks, Terry G. and Stuart D. Allen, "The St. Louis Equation: A Decade Later," Southern Economic Journal, January 1980, pp. 817-29.

[10] Toro-Vizcarrondo, C. and T.D. Wallace, "A Test of the Mean Square Error Criterion for Restrictions in Linear Regression," Journal of the American Statistical Association, June 1968, pp. 558-72.

[11] Toyoda, T. and T. D. Wallace, "Optimal Critical Values for Pre-Testing in Regression," Econometrica, March 1976, pp. 365-75.

[12] Wallace, T.D., "Weaker Criteria and Tests for Linear Restrictions in Regression," Econometrica, July 1972, pp. 689-98.

Table 1

St. Louis Equation With and Without Endpoint Constraints

	Without Endpoint Constraints		With Both Endpoint Constraints[1/]		With Endpoint Constraints on $\dot{E}$ only[1/]	
Lag	$\dot{M}$	$\dot{E}$	$\dot{M}$	$\dot{E}$	$\dot{M}$	$\dot{E}$
0	.59 (3.06)	.05 (1.17)	.38 (2.77)	.08 (2.10)	.61 (3.22)	.06 (1.72)
1	.22 (.91)	.10 (2.40)	.42 (5.50)	.06 (2.52)	.18 (.75)	.07 (2.75)
2	.08 (.32)	-.02 (-.41)	.27 (2.23)	.01 (.18)	.10 (0.42)	.02 (0.66)
3	.57 (2.36)	-.02 (-.57)	.06 (.71)	-.05 (-2.16)	.56 (2.32)	-.05 (-2.10)
4	-.41 (-2.17)	-.10 (-2.60)	-.08 (-.58)	-.07 (-2.01)	-.41 (-2.16)	-.09 (-2.50)
Sum	1.05 (5.82)	.01 (.09)	1.05 (5.71)	.03 (.30)	1.04 (5.84)	.01 (.16)
α	2.93		2.81		2.89	
$\bar{R}^2$	.40		.38		.41	
S.E.	3.63		3.72		3.63	

1/ Estimated with restricted least squares.

T-ratios in parentheses.

Table 2

F-tests on Combinations of Endpoint Constraints

	Head	Tail	Head and Tail
$\dot{M}$	.112	2.059	3.145*
$\dot{E}$	1.780	1.738	.947
$\dot{M}$ and $\dot{E}$	--	--	2.110

*Significant at the .05 level.