August 3, 2009

Alternative Estimators of Cointegrating Parameters in Models With Non-Stationary Data: An Application to US Export Demand (Working Paper)

James J Forest, *University of Massachusetts - Amherst*
Paul Turner

Alternative Estimators of Cointegrating Parameters in Models with Non-Stationary Data: An Application to US Export Demand

Jim Forest* and Paul Turner**

*Department of Finance, University of Connecticut, Storrs,CT, USA

** Department of Economics, Loughborough University, Loughborough, UK.

Abstract

This paper presents Monte Carlo simulations which compare the empirical performance of two alternative single equation estimators of the equilibrium parameters in a dynamic relationship. The estimators considered are Stock and Watson's dynamic ordinary least squares (DOLS) estimator and Bewley's transformation of the general autoregressive distributed lag model. The results indicate that the Bewley transformation produces a lower mean-square error as well as superior serial correlation properties even with lower truncation lags for the lagged variables included in the estimation equation. An application is then provided which examines the nature of the equilibrium relationship between aggregate US exports, world trade and the US real exchange rate. This confirms that estimation of the equilibrium parameters of this relationship by the Bewley transformation produces results which are superior to estimation by DOLS.

Keywords: DOLS estimator, Bewley transformation.

JEL Classification: C12, C15

Correspondence Address:

Dr Paul Turner
Department of Economics
Loughborough University
Loughborough
LE11 3TU

Tel: (44) 1509 222722
Fax: (44) 1509 223910
E-mail: p.m.turner@lboro.ac.uk

1. Introduction

In recent years Dynamic Ordinary Least Squares (DOLS) estimator introduced by Stock and Watson [9] has become a popular method for the estimation of equilibrium parameters in dynamic relationships between variables containing unit roots. The estimator has a number of features in common with an earlier estimator proposed by Bewley [1]. Despite the fact that Bewley's estimator (or the 'Bewley transformation' as it is popularly known) was introduced prior to the development of an extensive literature on the implications of unit roots in econometrics, we believe that it has a number of properties which make its performance superior to that of the DOLS estimator. As a result, we believe that this estimator has been unjustly neglected and that the comparison of its properties relative to that of the DOLS estimator presented in this paper are intended to demonstrate why it should be given more prominence in the applied econometrics literature.

In this paper we present Monte Carlo simulations which compare the empirical performance of the Bewley transformation and the DOLS estimator. The results indicate that the Bewley transformation produces a lower mean-square error as well as superior serial correlation properties even with lower truncation lags for the lagged variables included in the estimation equation. An application is then provided which examines the nature of the equilibrium relationship between aggregate US exports, world trade and the US real exchange rate. This confirms that estimation of the equilibrium parameters of this relationship by the Bewley transformation produces results which are superior to estimation by DOLS.

The plan of the paper is as follows: In section 2 we examine some statistical issues relating to the DOLS estimator and the Bewley transformation. We demonstrate that each of these estimators is superconsistent when the data series in the equilibrium relationship

1

contain a unit root. However, we also demonstrate that the Bewley transformation is necessarily biased in small samples because of the non-zero covariance of the current change in the dependent variable and the error term. In section 3 we set out our data generation process for a Monte Carlo investigation of the properties of these estimators and present simulations under a variety of scenarios. This is followed in section 4 by an application of the methodology to an equilibrium relationship between aggregate US exports, world trade and the US real exchange rate. Finally, in section 5 we present our conclusions.

2. Consistency of the DOLS estimator

The dynamic ordinary least squares (DOLS) estimator has become a popular method for the estimation of the parameters of cointegrating relationships between integrated variables. This estimator has been popularised by Stock and Watson [9] , building on earlier work by Saikkonen [8], and has since become established in the applied econometrics literature. Examples of its use can be found in Masih and Masih [5], Konno and Fukushige [4], Narayan [6] and Narayan and Narayan [7].

The DOLS estimator takes the form:

$$y_t = \gamma_1 x_t + \sum_{j=-k}^{k} \gamma_{j+k+2} \Delta x_{t+j} + u_t \tag{1}$$

where y and x are integrated series of the same order and u is a stationary random error. The parameter γ_1 measures the equilibrium or long-run response of y to x and the Δx terms capture the short-run relationship between the variables. The *truncation lag k* is chosen so that the errors are white-noise.

2

The nature of the DOLS estimator can be illustrated using the example of the finite distributed lag model (cf. Hendry [3])This takes the form:

$$y_t = \beta_1 x_t + \beta_2 x_{t-1} + u_t \tag{2}$$

where x is a weakly exogenous random variable and u is a random error with the standard Gaussian properties. This can easily be reparameterised to the DOLS format of:

$$y_t = \gamma_1 x_t + \gamma_2 \Delta x_t + u_t \tag{3}$$

where $\gamma_1 = \beta_1 + \beta_2$ is the long-run elasticity of y with respect to x. The DOLS estimator of γ_1 can be derived as:

$$\hat{\gamma}_1 = \frac{\sum x_t y_t}{\sum x_t^2 - \frac{\left(\sum x_t \Delta x_t\right)^2}{\sum \left(\Delta x_t\right)^2}} - \frac{\sum x_t \Delta x_t \sum \Delta x_t y_t}{\sum x_t^2 \sum \left(\Delta x_t\right)^2 - \left(\sum x_t \Delta x_t\right)^2} \tag{4}$$

Now if x is an $I(1)$ process, it follows that $\sum x_t^2$ and $\sum x_t y_t$ are $O(T^2)$ while $\sum x_t \Delta x_t$, $\sum \Delta x_t y_t$ and $\sum \left(\Delta x_t\right)^2$ are $O(T)$. Thus, as the sample size gets large, the DOLS estimator converges on $\hat{\gamma}_1 = \frac{\sum x_t y_t}{\sum x_t^2}$ which is just the OLS estimator for a simple regression of y on x. By standard results, the simple OLS estimator is super-consistent under these assumptions and therefore the DOLS is also super-consistent. Superconsistency means that the estimator converges on the true value at a faster rate than would be the case for stationary series because of the non-stationary nature of the

data in the regression equation. More formally, we have $\hat{\gamma}_1 - \gamma_2 = O(T^{-1})$ for series with a single unit root while $\hat{\gamma}_1 - \gamma_2 = O\left(\sqrt{T}^{-1}\right)$ for stationary data, i.e. the estimation error tends to zero at a rate given by $1/T$ for $I(1)$ series but at rate $1/\sqrt{T}$ for $I(0)$ series.

The super-consistency property of the DOLS estimator arises because of the reparameterisation of the original regression equation which has the effect of writing the equation so that the y and x variables are of a higher order of integration than the other variables in the relationship. However, this is also be true of the simple regression equation which excludes Δx_t from the right hand side. The DOLS formulation is only superior to simple OLS regression (or other similar estimators) if either (a) it has lower small sample bias or (b) the convergence to the long-run value is faster. These questions can be addressed theoretically as in the work of Stock and Watson [9] but they can also be investigated empirically through the use of Monte Carlo simulations.

The standard DOLS estimator (1) includes lags of the Δx process but not lags of Δy. However, there are important cases in which lags of Δy arise naturally. Unfortunately this creates econometric problems because Δy_t will be correlated with u_t, meaning that a DOLS estimator which includes Δy_t will suffer from small-sample bias. More interesting however, is whether this bias disappears quickly as suggested by the super-consistency property. Consider the standard partial adjustment model given by equation (5):

$$y_t = \beta_1 x_t + \beta_2 y_{t-1} + u_t \tag{5}$$

This can also be reparameterised to yield:

$$y_t = \gamma_1 x_t + \gamma_2 \Delta y_t + u_t \tag{6}$$

where $\gamma_1 = \dfrac{\beta_1}{1-\beta_2}$ is the long-run elasticity of y with respect to x. However, in this case

$\text{cov}(\Delta y_t, u_t) = \sigma_u^2$ and therefore the OLS estimator of the parameters of (6) will be biased in small samples. In large samples however, the OLS estimator can again be shown to have the super-consistency property under the assumption that x is integrated of order 1. This follows because we can write the OLS estimator of the parameter of interest as:

$$\hat{\gamma}_1 = \frac{\sum x_t y_t}{\sum x_t^2 - \dfrac{\left(\sum x_t \Delta x_t\right)^2}{\sum \left(\Delta y_t\right)^2}} - \frac{\sum x_t \Delta y_t \sum \Delta x_t y_t}{\sum x_t^2 \sum \left(\Delta y_t\right)^2 - \left(\sum x_t \Delta y_t\right)^2} \tag{7}$$

If x is $I(1)$ it follows that $\sum x_t^2$ and $\sum x_t y_t$ are $O(T^2)$ while $\sum x_t \Delta y_t$, $\sum \Delta y_t y_t$ and $\sum \left(\Delta y_t\right)^2$ are $O(T)$ which means that $\hat{\gamma}_1 \to \dfrac{\sum x_t y_t}{\sum x_t^2}$ as $T \to \infty$. This is again, the simple OLS formula for a simple regression of y on x and, because simple OLS is super-consistent, it follows that this reparameterisation also yields super-consistent estimates. In this case however, finite sample estimates will be biased.

The reparameterisation (6) is familiar in the applied econometrics literature. It was originally suggested by Bewley [1]. However, Bewley [1] was writing before the literature on cointegration became established and therefore before econometricians became aware of the super-consistency property. The emphasis in this paper is on the need to allow for the endogeneity of the right-hand side variable Δy_t which is introduced by the transformation. Bewley [1] suggests the use of an instrumental variable estimator where Δy_t is instrumented by y_{t-1} and this can be shown to yield consistent estimates even in the case where x and y are $I(0)$ variables. As with any instrumental variable

estimator, there is a trade-off between bias and efficiency. In the case of integrated variables, where the bias disappears as the sample size gets larger, the use of an IV estimator may produce inferior results in terms of the mean square error. However, this is again an issue which can usefully be investigated using Monte Carlo methods. Our aim in this paper is to use such methods to compare the statistical performance of the DOLS estimator and the Bewley transformation .

A final issue concerning the DOLS estimator is the role of 'leads' in the DOLS regression equation. Stock and Watson [9] argue that the leads of Δx in equation (1) allow for feedback in cases where x is Granger caused by y. In a recent paper Hayakawa and Kurozumi [2] investigate the use of leads in the DOLS model and conclude that its performance can be improved significantly when the leads are excluded, provided that it is first established that x is not Granger caused by y.

3. Simulation results and comparison with alternative estimators

The data generation process (DGP) used for the simulations in this paper is the same as that used by Hayakawa and Kurozumi [2]. This DGP is used partly because it forms a standard benchmark for the analysis of cointegrating relationships between variables and partly to facilitate comparisons with the Hayakawa and Kurozumi paper. It is specified as follows:

$$
\begin{aligned}
y_t &= 1 + x_t + u_{1t} \\
x_t &= x_{t-1} + u_{2t}
\end{aligned}
\quad \text{where} \quad
\begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix} =
\begin{pmatrix} a_{11} & 0 \\ 0 & a_{22} \end{pmatrix}
\begin{pmatrix} u_{1t-1} \\ u_{2t-1} \end{pmatrix} +
\begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}
\quad \text{and} \quad
\begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix} \sim iidN \left(0, \begin{bmatrix} 1 & \sigma_{12} \\ \sigma_{12} & 1 \end{bmatrix} \right)
$$

This DGP is simulated using the EViews seeded random number generator for the normal distribution with $u_{10} = u_{20} = 0$ and the first 100 observations are discarded to avoid the

6

starting values contaminating the results. 10,000 replications are considered for each set of simulations. A non-zero value of a_{22} allows for Granger causality from y to x.

The simulations reported in this paper consider three possible strategies for the determination of the truncation lag. The first strategy involves choosing a value of k which ensures that there is no residual serial correlation in the estimated equation. We attempt to ensure this by beginning with $k=0$ and then progressively increasing its value if the Ljung-Box test statistic for serial correlation up to 10th order exceeds its critical value. In some cases this led to extremely long truncation lags and therefore a maximum value of k was fixed at 10. The two alternative criteria involve a choice of k which minimises either the Akaike Information Criterion (AIC) or the Schwartz Information Criterion (SIC). We allow for four different specifications of the estimated equation. The first two are variants of the standard DOLS model (1) in which we first estimate with lags only and then allow for the possibility of lead terms for Δx. The next two specifications are variants of the Bewley transformation equation (6). In the first, we estimate using OLS, implicitly relying on the super-consistency property to generate consistent estimates when $\text{cov}(\Delta y_t, u_t) \neq 0$. In the second we estimate by instrumental variables to allow for the endogeneity of the Δy_t variable explicitly. Each simulation is run for a variety of different values of $a_{22} = \{0.0, 0.5, 0.8\}$ to estimate the effects of Granger feedback from y to x.

Table One presents the results for simulations in which the truncation lag is determined by the Ljung-Box test for serial correlation. The results reported for each simulation comprise the bias (average deviation of $\hat{\gamma}_1$ from its true value), the standard deviation, the mean square error, the average truncation lag and the percentage of the simulations in which $H_0: \gamma_1 = 1$ is rejected in favour of the alternative $H_1: \gamma_1 \neq 1$ using the 5% critical value from the standard normal distribution. Since all the processes considered contain a unit root, we would expect that the standard normal critical values would not be

appropriate. However, it is still interesting to examine the differences between the alternative models in terms of the rejection frequency.

Table One: Simulation Results – Estimates of γ with Truncation Lag determined by Ljung-Box Serial Correlation Test

Feedback Parameter a22		DOLS (Lags Only)	DOLS (Leads and Lags)	Bewley (OLS)	Bewley (IV)
0	Bias	0.002064357	0.001911935	0.008929348	0.001948995
	Standard Deviation	0.067992392	0.084420662	0.064458062	0.059454972
	MSE	0.004627227	0.007130504	0.004234575	0.003538692
	DW Stat	1.0903	1.1019	0.2515	1.9739
	Truncation Lag	9	8.9	10	0.4
	% H0 rejections*	24.89	25.11	58.76	7.11
0.5	Bias	0.000070174	-0.000144475	0.001446276	-0.007386494
	Standard Deviation	0.035287789	0.043951942	0.033124217	0.031429874
	MSE	0.001245233	0.001931794	0.001099305	0.001042397
	DW Stat	1.103	1.1159	0.2502	1.8003
	Truncation Lag	9	8.9	10	0.7
	% H0 rejections	24.81	25.06	59.31	12.33
0.8	Bias	-0.000494056	-0.000553651	-0.001148566	-0.007093028
	Standard Deviation	0.01600721	0.020011418	0.014891499	0.016510089
	MSE	0.000256475	0.000400763	0.000223076	0.000322894
	DW Stat	1.1146	1.1288	0.2578	1.7712
	Truncation Lag	8.9	8.7	10	0.8
	% H0 rejections	25.59	25.68	59.51	16.53

* Test of $H_0 : \gamma = 1$ against $H_1 : \gamma \neq 1$ using a 5% critical value. Bias, Standard Deviation, MSE, DW Stat and Truncation Lag are averages over 10,000 simulated values.

In terms of the bias and the mean square error, it is clear from Table One that all the estimators have reasonable properties in that both the bias and the mean square error are very low in absolute terms. This reflects the super-consistency property which holds when y and x are both unit root processes. In fact there is relatively little to choose between the estimators, though there is a tendency for the DOLS estimator with leads and lags to have a higher MSE. Interestingly the MSE tends to decline the higher is the a_{22} parameter and therefore the greater the extent of Granger feedback from y to x. However, there is no indication that the inclusion of lags improves the performance noticeably in these cases as suggested by Hayakawa and Kurozumi [2].

The Bewley transformation with IV estimation has two advantages over the other estimators. First, it is considerably more parsimonious as shown by the lower truncation lag in all cases. For the two DOLS estimators, the truncation lags are around 9 in all cases. However, even such long truncation lags do not completely remove serial correlation as can be seen by the low values of the Durbin-Watson test statistics which indicate that first order serial correlation remains a problem in all cases. The fact that the simulations pass the Ljung-Box test has more to do with its lack of power due to the portmanteau nature of the test, and the decision to test for serial correlation up to 10th order, rather than the ability of lags of Δx to capture residual serial correlation. The second advantage of the Bewley transformation is that it produces rejection frequencies for $H_0 : \gamma_1 = 1$ which are much closer to the 5% value which would be observed if the test had the correct size. The importance of using an IV estimator is indicated by the much larger truncation lags when OLS is used with the Bewley transformation as well as the higher MSE and rejection frequency. In fact the Bewley transformation, coupled with OLS estimation, produces the least satisfactory of all the simulation results as evidenced by the high MSE, the low values of the Durbin-Watson statistic and the excessive rejection frequencies for the null hypothesis $H_0 : \gamma_1 = 1$.

Next consider Table Two, in which the truncation lag is chosen by the process of minimising the Akaike Information Criterion. For the DOLS estimators, the truncation lags are much lower than we when use the Ljung-Box criterion. In these simulations the average k is typically between 2 and 3, as is also the case when the Bewley transformation with IV is used. The Bewley transformation with OLS estimation is again the worst performing of the four estimators with truncation lags at the maximum value of 10, low values of the Durbin-Watson statistics and high rejection frequencies for the null that the cointegrating parameter for x equals its true value. The best performing estimator is again the Bewley transformation with IV estimation as evidenced by the fact that the Durbin-Watson statistics are close to 2 coupled with rejection frequencies which, while still too high, are closer to the theoretical size of the test. Table Three gives the results when the truncation lag is chosen by the Schwartz Information Criterion. The results are qualitatively very similar to the AIC results, although the truncation lags are generally lower reflecting the more parsimonious nature of the SIC.

Table Two: Simulation Results – Estimates of γ with Truncation Lag determined by Akaike Information Criterion

Feedback Parameter a22		DOLS (Lags Only)	DOLS (Leads and Lags)	Bewley (OLS)	Bewley (IV)
0	Bias	0.00384808	0.005710912	0.008935709	0.002177571
	Standard Deviation	0.060481625	0.063711014	0.0644623	0.059990239
	MSE	0.003672835	0.004091708	0.004235235	0.003603571
	DW Stat	1.0835	1.0858	0.2514	1.9824
	Truncation Lag	2.5	2.2	10	1
	% H0 rejections	25.65	26.38	58.77	7.39
0.5	Bias	0.000119103	0.000137337	0.001449597	0.000477436
	Standard Deviation	0.029992372	0.030879353	0.033132893	0.030874223
	MSE	0.000899557	0.000953553	0.00109989	0.000953446
	DW Stat	1.0934	1.0946	0.2499	1.9752
	Truncation Lag	1.3	1.2	10	1.0
	% H0 rejections	25.06	25.16	59.33	7.85
0.8	Bias	-0.000792338	-0.001125289	-0.001150984	-0.000288335
	Standard Deviation	0.014098662	0.014767737	0.014890968	0.013830719
	MSE	0.0001994	0.000219352	0.000223066	0.000191372
	DW Stat	1.1058	1.1082	0.2575	1.9818
	Truncation Lag	2.4	2.2	10	1.0
	% H0 rejections	26.01	26.34	59.51	8.26

Table Three: Simulation Results – Estimates of γ with Truncation Lag determined by Schwartz Information Criterion

Feedback Parameter a22		DOLS (Lags Only)	DOLS (Leads and Lags)	Bewley (OLS)	Bewley (IV)
0	Bias	0.007264562	0.010274228	0.009875948	0.002169948
	Standard Deviation	0.05937801	0.060451518	0.063979632	0.060012366
	MSE	0.003578522	0.003759946	0.004190928	0.003606193
	DW Stat	1.0826	1.0828	0.2483	1.9824
	Truncation Lag	1.8	1.4	9.1	1.0
	% H0 rejections	25.86	26	58.6	7.41
0.5	Bias	0.000200279	0.000194395	0.001624559	0.000473153
	Standard Deviation	0.029656667	0.030227026	0.032741218	0.030893838
	MSE	0.000879558	0.000913711	0.001074627	0.000954653
	DW Stat	1.0929	1.0941	0.2464	1.9751
	Truncation Lag	1	1	9	1.0
	% H0 rejections	24.82	24.76	58.67	7.88
0.8	Bias	-0.001497444	-0.002070976	-0.001223095	-0.000287102
	Standard Deviation	0.01389029	0.014198143	0.014747346	0.013838305
	MSE	0.000195183	0.000205876	0.00021898	0.000191581
	DW Stat	1.1043	1.1051	0.2543	1.9815
	Truncation Lag	1.8	1.4	9.1	1.0
	% H0 rejections	25.98	26.11	58.91	8.33

13

4. Example: US Export Demand 1978q4 to 2005q4

As an example of the relative performance of the DOLS methodology and the Bewley procedure, this section presents the results from estimation of an equation for exports in the US economy over the period 1978q4 to 2005q4. The equation relates the US$ value of US exports at constant 2000 prices to the total exports of the five large industrial economies (France, Germany, Italy, UK and Japan) also measured in US$ at 2000 prices as well as the real effective exchange rate (measured on a relative unit labour cost basis). The three variables used as the basis of the equation are shown in Figure 1 with each variable being expressed in logarithmic form. All data are taken from the IMF International Financial Statistics database.
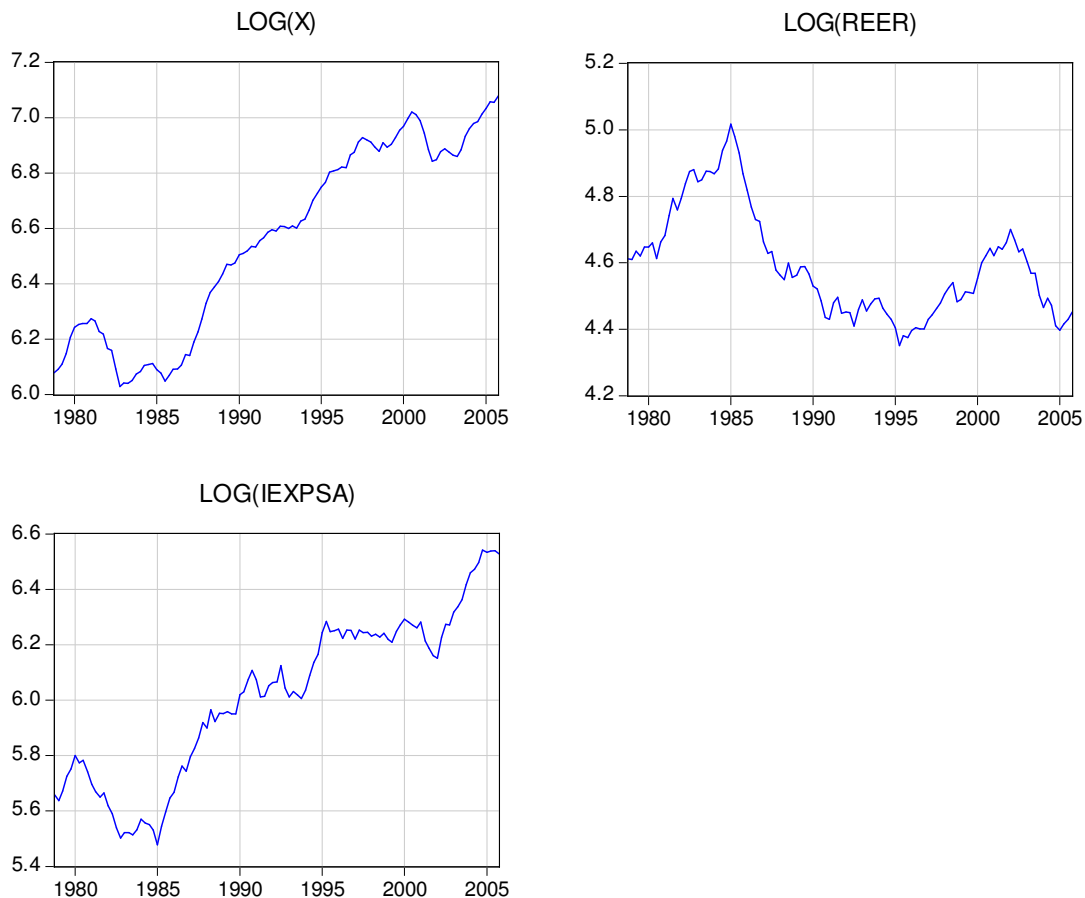
Figure 1: US export model variables – X = US exports, IEXPSA = Total Exports of France, Germany, Italy, UK and Japan, REER = real effective exchange rate (relative unit labour cost basis.

Preliminary data analysis shows that each of the variables contains a unit root. Application of the Johansen cointegration test yields a trace statistic of 29.69 and a maximum eigenvalue statistic of 24.72 which compare with 5% critical values of 31.88 and 21.49 respectively. Thus the evidence for a cointegrating vector is mixed with the trace test (only just) failing to reject the null of no cointegrating vectors while the maximum eigenvalue test rejects. On this basis we thought it reasonable to assume that a cointegrating relationship exists and that the DOLS and Bewley estimates could be interpreted as methods for the estimation of the parameters of this vector.

Next, we estimated the parameters of the equilibrium relationship using both the DOLS procedure and the Bewley transformation. The results of the DOLS procedure are given in equation (8). A truncation lag of 4 was chosen on the basis that this minimises the Schwartz criterion. The coefficient estimates for the lags and leads of the differenced explanatory variables are not reported because these are of no direct interest. Standard errors are given in parentheses below coefficients.

Equation (8) yields estimates of the long-run elasticities of US exports with respect to world trade and the real exchange rate of 1.23 and 0.15 respectively. The estimate of the world trade elasticity is considerably above one and, on the basis of a conventional *t* test would be significantly greater than one. The estimate of the real exchange rate elasticity has the 'wrong' sign in that economic theory would suggest a negative effect and is moreover statistically significant when we use a *t*-test. However, the *t*-test is not appropriate here because the equation is badly dynamically misspecified. A truncation lag of 4 is clearly not sufficient to produce residuals which are serially uncorrelated as

evidenced by the Durbin-Watson and Breusch-Godfrey test statistics. Moreover, extending the truncation lag does little to eliminate serial correlation. Truncation lags of up to 10 were tried with little change in either of the test statistics for serial correlation. The equation also shows clear evidence of misspecification in the form of a significant first-order ARCH process in the residuals as well as a significant Jarque-Bera test for non-normality. In summary, the DOLS procedure may give unbiased estimates of the long-run elasticities but there is little else to recommend it since it produces a badly misspecified equation when estimated.

$$
\ln X_t = -\ \underset{(0.3578)}{1.5288} +\ \underset{(0.0247)}{1.2306}\ \ln W_t +\ \underset{(0.0492)}{0.1508}\ \ln E_t
$$

$$
+ \sum_{i=0}^{4} \hat{\gamma}_i \Delta W_{t-i} + \sum_{i=1}^{4} \hat{\gamma}_{i+4} \Delta W_{t+i} + \sum_{i=0}^{4} \hat{\phi}_i \Delta E_{t-i} + \sum_{i=1}^{4} \hat{\phi}_{i+4} \Delta W_{t+i} + \hat{u}_t
$$

(8)

$$
DW = 0.3718 \qquad JB = 7.43\,(0.024) \qquad \hat{\sigma} = 0.041908
$$
$$
BG_4 = 38.44\,(0.00) \qquad ARCH = 157.44\,(0.00)
$$

DW is the Durbin-Watson statistic, JB is the Jarque Bera statistic for non-normality of the residuals which is distributed as $\chi^2(2)$ under the null of normal residuals, $\hat{\sigma}$ is the standard error of the regression, BG_4 is the χ^2 form of the Breusch-Godfrey test for 4th order serial correlation and ARCH is the χ^2 form of the Lagrange Multiplier test for first order autoregressive conditional heteroscedasticity in the residuals. P-values are given in parentheses next to test statistics.

Next, we estimated an equilibrium relationship using the Bewley transformation. Again, the Schwartz criterion was used to determine the truncation lag for the number of differenced variables and this produced a truncation lag of zero in this case, i.e. the current value of each of the differences. The results are reported in equation (9) which was estimated by the method of instrumental variables using $\ln X_{t-1}$ as an instrument for the endogenous right-hand side variable $\Delta \ln X_t$.

$$\ln X_t = 1.8315 + 1.0540 \ln W_t - 0.3398 \ln E_t$$
$$ (1.3251) \quad (0.0873) \qquad (0.1903)$$

$$+ \hat{\gamma}_1 \Delta \ln X_t + \hat{\phi}_1 \Delta \ln W_t + \hat{\eta}_1 \Delta \ln E_t + \hat{u}_t$$

(9)

$$DW = 1.7821 \qquad JB = 0.32(0.85) \qquad \hat{\sigma} = 0.01719$$
$$BG_4 = 3.70(0.44) \qquad ARCH = 0.0432(0.84)$$

The estimates of the long-run elasticities given in equation (9) are closer to what we would expect from economic theory. The estimate of the world trade elasticity is 1.05 which is well within two standard errors of one (even given the likely downward bias of the standard errors). The estimate of the real exchange rate elasticity has the expected sign, and is reasonable in magnitude, but is not significantly different from zero when we use a t-test. The most important difference between the two equations we have estimated however, can be seen in the misspecification tests. For example, the Durbin-Watson and Breusch-Godfrey tests for equation (9) both indicate than we cannot reject the null hypothesis that the residuals are serially uncorrelated. In addition there is no evidence of an ARCH process or of non-normality in the residuals. Thus equation (9) appears to be a statistically well specified equation whereas equation (8) is clearly badly misspecified.

A comparison of the two methods therefore confirms that it is much easier to obtain a well-specified, parsimonious equation using the Bewley transformation than using the DOLS methodology. Our Monte Carlo analysis indicates that even with the Bewley transformation, it remains the case that the distribution of the long-run coefficient estimates is non-standard. However, the distribution is much closer the t or normal distribution than would be the case for the dynamically misspecified DOLS equation. This can be useful when testing hypotheses of interest. For example, in equation (9) when we wish to test $H_0 : \beta_2 = 1$, where β_2 is the long-run elasticity of US exports with respect to world trade, the Monte Carlo analysis shows us that the standard errors are probably biased downwards slightly. Therefore the fact that we cannot reject the null at the usual 5% significance level adds to our confidence that this is a reasonable restriction. In

contrast that fact that we can reject the same null for the DOLS model tells us nothing because the misspecification leads to such biased estimates of the standard errors that it is almost inevitable that the null will be rejected in this case.

5. Conclusions

This paper has investigated alternative single equation estimation methods for cointegrating parameters. All the methods investigated produce generally low bias and MSE which reflects the fact that all these methods will produce super-consistent estimates when applied to non-stationary data. Although the DOLS method is widely used in the current applied econometric literature, our results indicate that the Bewley transformation estimated using instrumental variables produces better results in terms of slightly lower MSE and noticeably better serial correlation properties. We have investigated an applied example which relates the demand for US exports to the level of industrial country exports and the US real effective exchange rates. This indicates that the Bewley transformation produces more plausible estimates of the parameters of interest (the long-run elasticities) as well as producing a statistically well specified equation which is considerably more parsimonious. Therefore, both our Monte Carlo analysis and our example are consistent with the argument that the Bewley transformation provides a more satisfactory method for the estimation of long-run parameters in a dynamic model.

References

[1] R. A. Bewley, The Direct Estimation of the Equilibrium Response in a Linear Model, Economics Letters 3 (1979) 375-381.

[2] K. Hayakawa, E. Kurozumi The role of "leads" in the dynamic OLS estimation of cointegrating regression models Mathematics and Computers in Simulation (2008) (in press) doi:10.10106/j.matcom.2008.02.027.

[3] D.F. Hendry, Dynamic Econometrics, Oxford University Press,Oxford, 1995.

[4] T. Konno, M. Fukushige, Did Nafta cause the structural changes in bilateral import functions between the US and Mexico?, Journal of Policy Modeling 25 (2003) 53-59.

[5] R. Mashih, A.M.M. Masih, Stock-Watson dynamic OLS (DOLS) and error correction modelling approaches to estimating long- and short-run elasticities in a demand function: new evidence and methodological implications from an application to the demand for coal in mainland China, Energy Economics 18 (1996) 315-334.

[6] P. K. Narayan, Do public investments crowd out private investments? Fresh evidence from Fiji, Journal of Policy Modeling 26 (2004) 747-753.

[7] P.K. Narayan, S. Narayan, (2005) 'Estimating income and price elasticities of imports for Fiji in a cointegration framework', Economic Modelling 22 (2005) 423-438.

[8] P. Saikkonen, Asymptotically efficient estimation of cointegration regressions, Econometric Theory 7 (1991) 1-21.

[9] J.H. Stock, M.W. Watson, A simple estimator of cointegrating vectors in higher order integrated systems, Econometrica 61 (1993) 783-820.