Dundee Discussion Papers in Economics 268

Chowdhury, Rosen Azad; Russell, Bill

Publication date:
2012

Dundee Discussion Papers in Economics

The Difference, System and 'Double-D' GMM Panel Estimators in the Presence of Structural Breaks

Rosen Azad Chowdhury & Bill Russell

Department of
Economic Studies
University of Dundee
Dundee
DD1 4HN

Working Paper
No. 268
June 2012
ISSN: 1473-236X

The Difference, System and 'Double-D' GMM Panel Estimators in the Presence of Structural Breaks*

Rosen Azad Chowdhury# and Bill Russell†

21 June 2012

Abstract

The effects of structural breaks in dynamic panels are more complicated than in time series models as the bias can be either negative or positive. This paper focuses on the effects of mean shifts in otherwise stationary processes within an instrumental variable panel estimation framework. We show the sources of the bias and a Monte Carlo analysis calibrated on United States bank lending data demonstrates the size of the bias for a range of auto-regressive parameters. We also propose additional moment conditions that can be used to reduce the biases caused by shifts in the mean of the data.

***Keywords*:** Dynamic panel estimators, mean shifts/structural breaks, Monte Carlo Simulation

JEL code: C23, C22, C26

---

* #Corresponding author, Economic Studies, University of Dundee, Dundee, DD1 4HN, Scotland, United Kingdom, email: rchowdhury@dundee.ac.uk. †Economic Studies, University of Dundee, email: brussell@brolga.net. We thank Arnab Bhattacharjee, Hassan Molana and Dennis Petrie for their insightful comments. We also thank Tom Doan for making available his Bai-Perron programmes. Data are available at www.billrussell.info.

1. Introduction

Instrumental variable panel estimators are used in almost all fields of economics and are usually consistent and efficient. However, econometricians have noted that in some cases, like in the presence of heteroscedasticity or highly persistent data, instrumental variables estimators can perform poorly. Furthermore, Carrion-i-Silvestre *et al.* (2005) and Bai and Carrion-i-Silvestre (2009) demonstrate that unaccounted structural breaks bias the least squares estimates in standard auto-regressive panels. In this paper we add another dimension to this existing literature by showing how structural breaks in the mean of the variables can result in severely biased estimates in dynamic panels when the data is endogenous. We also propose two new moment conditions for the GMM estimator that reduces the bias substantially in dynamic time series panels.

Nickell (1981) shows the panel estimator of auto-regressive terms is subject to a positive bias due to unobserved fixed effects and this bias is present irrespective of structural breaks in the data. Compared to work done on structural breaks in time series the panel literature is still in its infancy. Notable work has been undertaken by Carrion-i-Silvestre *et al.* (2005) and Bai and Carrion-i-Silvestre (2009) who demonstrate that when the variables are strictly exogenous the power of panel unit root tests decreases in the presence of structural breaks making the data look more persistent. For example, unaccounted structural breaks in mean introduce a positive bias to the auto-regressive terms and the size of this bias depends on the magnitude and timing of the breaks and the sample length. This can be thought of as the Perron (1989) effect in panels. Carrion-i-Silvestre *et al.* (2005) and Bai and Carrion-i-Silvestre (2009) also show that the unaccounted breaks in mean introduces an additional bias to the Perron effect outlined above by changing the magnitude of the fixed effects bias of Nickell. While the Perron effect of unaccounted breaks in mean is always positive the effect of the unaccounted breaks on the fixed effects bias results in the bias being mostly negative but in some cases positive. Consequently the sign of the total bias is ambiguous in the presence of structural breaks.

However, in dynamic panels that incorporate endogenous variables the effect of structural breaks is more complicated. For example, Arellano and Bover (1995) and Blundell and Bond (1998) show that applying the difference GMM estimator to highly persistent data in dynamic models leads to invalid instruments which in turn causes a downward bias (in absolute terms) to the estimated coefficient on the lagged dependent variable. The usual way to overcome the problem of highly persistent data as suggested by these papers is to assume that the persistence has some economic rationale and estimate the model using the systems GMM estimator where the instruments are included as first differences. However, if the data looks persistent only because of structural breaks then this solution to 'imagined' persistence in the data leads to biased estimates and possibly incorrect inference. Consequently, unaccounted structural breaks in mean introduce an 'endogeneity' bias in difference and system GMM estimators which is over and above the Perron and Nickell biases outlined above. This paper seeks to identify the 'endogeneity' bias in the difference and system panel estimators before proposing two new moment conditions which can be used to reduce the 'endogeneity' bias.

In the next section we begin by briefly setting out the standard Carrion-i-Silvestre *et al.* (2002) analysis of the biases due to structural breaks in a dynamic panel assuming the variables are strictly exogenous. We extend this methodology to analyse the biases due to structural breaks assuming the data is endogenous. We identify three biases. The first two (the Perron and Nickell effects) are equivalent to those found when the data is exogenous. The third bias is due to the endogeneity of the data and is particularly important when the data is highly persistent. These biases indicate that the moment conditions are not zero in the presence of structural breaks. We therefore suggest two moment conditions that are zero in the presence of structural breaks and term the associated GMM estimator the 'double-D' GMM estimator.[1]

Section 3 uses a Monte Carlo analysis calibrated on United States bank lending data to examine the difference, system and double-D GMM estimators both without and with structural breaks in the data. We find that in the presence of structural breaks the double-D estimator out performs the difference and system GMM estimators for low levels of persistence (i.e. autoregressive coefficients less than 0.6) and the difference and system GMM estimators perform marginally better when persistence is high. A panel data model of the bank lending channel is then estimated in Section 4 to demonstrate the advantage of the double-D GMM estimator when estimating models in the presence of structural breaks.

2. Structural Breaks and their impact on the GMM panel estimators

Carrion-i-Silvestre *et al.* (2002) in a similar vein to Perron (1989) showed that the bias due to unaccounted mean shifts in panel data reduces the power of traditional unit root tests with exogenous data. They start with an AR(1) process, y_{it}, with a single level shift;

[1] The name of the estimator will become evident later in the paper.

$$y_{it} = \alpha y_{it-1} + \theta_i DU_t + \eta_i + v_{it} \quad (1a)$$

$$\mathrm{v_{it}} = \delta \mathrm{v_{it-1}} + \varepsilon_{\mathrm{it}},\ \delta = 0,\ \alpha < 1 \quad (1b)$$

where, i is the entity in the panel, η_i are the time invariant fixed effects, v_{it} is the error term and $DU = 1$ for $t \geq t_B$ and 0 elsewhere, with t_B indicating the date of the structural break.

2.1 *Structural Breaks and the difference GMM Estimator*

Carrion-i-Silvestre *et al.* (2002) demonstrates that if the shift term is unaccounted for and one estimates with least squares;

$$\mathrm{y_{it}} = \alpha \mathrm{y_{it-1}} + \eta_\mathrm{i} + \mathrm{v_{it}} \quad (2)$$

then α will be biased and the least square estimate of α is;

$$\widehat{\alpha} = \alpha + \underbrace{\frac{\sum_{i=1}^{N} y' Q_T v_i}{\sum_{i=1}^{N} y' Q_T y_{i,-1}}}_{\mathrm{NE}} + \underbrace{\frac{\sum_{i=1}^{N} y'_{i,-1} Q_T z \zeta_i}{\sum_{i=1}^{N} y'_{i,-1} Q_T\, y_{i,-1}}}_{\mathrm{PE}} \quad (3)$$

where, $Q_{T=}I_T - X(X'X)^{-1}X'$, X being the TXK matrix of non–stochastic regressors, $Z = [e_{T,}DU]$ where $e_T = (1, \ldots..1)' = X$ and $\zeta = (\alpha_i, \theta_i)'$ with θ being the magnitude of the break. Equation (3) shows the biases due to the unaccounted mean shifts is made up of two components. The bias identified by Nickell (1981) caused by fixed effects in OLS estimation is shown as NE in equation (3). Carrion-i-Silvestre *et al.* (2002) argue that this bias is negative although the sign is positive in Nickel's original paper which does not include structural breaks.[2] The bias identified as PE in equation (3) is positive and is similar to the Perron (1989) effect. Hence, the net bias when the data is exogenous depends on the relative magnitudes of the Nickell effect, NE, and Perron effect, PE, such that;

$$\widehat{\alpha} = \alpha + \mathrm{NE} + \mathrm{PE} \quad (4)$$

[2] Carrion-i-Silvestre *et al.* (2002) show that the sign of the denominator of NE in equation (3) depends on the magnitude of the auto regressive parameter and the break function involved. They conclude that in general the sign of the Nickell effect in the presence of structural breaks is negative.

We now extend this approach to consider the difference GMM estimator when the data is endogenous. Although the 'true' data generating process is as described by equation (1) we ignore the shift term and assume the process is as described by equation (2). In this case the standard difference GMM (Arellano and Bond 1991 type) orthogonal moment conditions can be written;

$$\mathrm{E}(\mathrm{y}_{\mathrm{it-s}}\,\Delta \mathrm{u}_{\mathrm{it}}) = 0 \;\text{ for }\; \mathrm{t} = 3,\; 4\ldots.\mathrm{T} \text{ and } 2 \;\leq\; \mathrm{S} \;\leq\; \mathrm{t}-1 \qquad (5)$$

Assuming $T = 3$ then the moment condition in equation (5) is exactly identified and the corresponding method of moments estimator reduces to a two stage least square estimator.[3] This implies the first stage of the instrumental variable regression is;

$$\Delta \mathrm{y}_{\mathrm{it}} = (\alpha - 1)\mathrm{y}_{\mathrm{it-1}} + \eta_{\mathrm{i}} + \mathrm{v}_{\mathrm{it}} \qquad (6a)$$

$$\Delta \mathrm{y}_{\mathrm{it}} = \beta \mathrm{y}_{\mathrm{it-1}} + \eta_{\mathrm{i}} + \mathrm{v}_{\mathrm{it}} \qquad (6b)$$

where $\beta = (\alpha - 1)$. The least squares estimator of equation (6b) is then;

$$\hat{\beta} = (\alpha - 1) + \underbrace{\frac{\sum_{i=1}^{N} y'_{i,,-1} Q_T v_i}{\sum_{i=1}^{N} y'_{i,-1} Q_T y_{i,-1}}}_{\mathrm{NE}} + \underbrace{\frac{\sum_{i=1}^{N} y'_{i,-1} Q_T Z \varsigma_i}{\sum_{i=1}^{N} y'_{i,-1} Q_T\, y_{i,-1}}}_{\mathrm{PE}} \qquad (7)$$

$$\hat{\beta} = (\alpha - 1) + \mathrm{NE} + \mathrm{PE} \qquad (8)$$

where the Nickell effect, NE, and Perron effect, PE, are the same as those in equation (4) when the data is exogenous. Arellano and Bond (1991) show that as the data becomes more persistent then without structural breaks $\alpha \rightarrow 1$ and $(\alpha - 1) \rightarrow 0$ in equation (6a) and y_{it-1} becomes an invalid instrument as the correlation between Δy_{it} and y_{it-1} declines. Therefore, the 'persistence bias' and the Nickell effect are negative while the Perron effect is positive in equation (8).

Consequently, the total bias is non-linear and depends on the relative magnitudes of the three biases. In equation (8) if α is small and the positive Perron effect is larger than the negative Nickell effect then $\hat{\beta}$ will be biased upwards. Alternatively, when persistence is high then $\alpha - 1$ tends to zero creating a negative bias to $\hat{\beta}$. If this negative bias along with the negative Nickell effect is

[3] Assuming $T = 3$ avoids the use of matrixes and greatly simplifies the exposition. If $\mathrm{T} > 3$ then the following results also apply to the other moment conditions.

greater than positive Perron effect then $\hat{\beta}$ will be biased downwards and the instruments will be less correlated with the Δy_{it} term.[4] Therefore, when estimating the model without accounting for the structural breaks the instruments may become invalid with the difference GMM estimator resulting in the estimates being biased. The standard response to finding the data are highly persistent is to estimate the model in equation (2) using the system GMM estimator and it is this estimator that we now turn to.

2.2 *Structural Breaks and the System GMM Estimator*

Arellano and Bover (1995) and Blundell and Bond (1998) demonstrated that when the data is persistent (i.e. when $\alpha \geq 0.8$) the difference GMM performs poorly for the reasons explained above. The solution proposed in both these papers is to use system GMM where lagged differenced terms are used as instruments instead of the lagged level terms as in difference GMM. They also demonstrate using Monte Carlo simulations that the system GMM performs better than the difference GMM when data is highly persistent.[5] Although it has been shown in the literature that system GMM adequately accounts for persistence in the data, we show that when persistence is caused by structural breaks in the mean of the data and these breaks are not accounted for then the moment conditions of system GMM may become invalid and the estimators biased.[6]

To show this, start with our simple AR (1) panel data model of equation (1) represented here as the period before the break;

$$y_{it} = \alpha y_{it-1} + \eta_i + v_{it}\,,\ t < T_B \tag{9a}$$

and the period after the break:

$$y_{it} = \alpha y_{it-1} + \eta_i + \delta_{iT_B} + v_{it},\ t \geq T_B \tag{9b}$$

4 However, Hayakawa (2009) argue that if the data is mean non stationary the moment condition of the first difference GMM may be valid even when the auto regressive parameter is high. This is due to the unaccounted fixed effects in the first stage of the regression.

5 Roodman (2008) and Blundell and Bond (1998) argue that if the time period is small and the individual fixed effects are large then system GMM may perform poorly.

6 We consider breaks in the mean of the data but similar results can be obtained by changes in the auto-regressive term.

In equation (9b), δ_{iT_B} is the mean shift and T_B is the break date where we assume that $T_B = 3$.[7]

Assuming that $E(\delta_{T_B} v_{it}) \neq 0$ for $t < T_B$, $E(\delta_{T_B} v_{it}) = 0$ for $t \geq T_B$ and $E(\delta_{T_B} \eta_i) \neq 0$ then for $t = 4$, the moment conditions in the system GMM if there are no structural breaks in the mean of the data are;[8]

$$E[\Delta y_{it-1}(v_{it} + \eta_i)] = 0 \tag{10}$$

However, if there are unaccounted breaks then the moment conditions in equation (10) will not be valid and $E[\Delta y_{it-1}(v_{it} + \eta_i)] \neq 0$. With structural breaks, therefore, the moment conditions when $t = 4$ are:

$$E[\Delta y_{it-1}(v_{it} + \eta_i)] = [(y_{i3} - y_{i2})(v_{i4} + \eta_i)] \tag{11a}$$

$$= E[(\alpha y_{it-2} + \delta_{iT_B} + v_{i3} + \eta_i - y_{it-2})(v_{i4} + \eta_i)] \tag{11b}$$

$$= E[((\alpha - 1)y_{it-2} + \eta_i)\eta_i] + E[\delta_{T_B}\eta_i] \tag{11c}$$

Equation (11c) differs from the standard moment condition of no structural breaks system GMM of equation (10) by the term, $E(\delta_{T_B} \eta_{it})$ which is non-zero and therefore the moment condition, $E[\Delta y_{it-1}(v_{it} + \eta_i)]$, is not equal to zero and invalid along with the instruments. Moreover, in system GMM with structural breaks the initial moment condition will not decay towards its long run mean set by the parameter α in equations (9a) and (9b).

2.3 *The Double-D GMM Estimator*

The problem caused by unaccounted structural breaks in the system GMM can be resolved by changing the moment condition in equation (10) to $E(\Delta y_{it-2}(\Delta v_{it} + \Delta \eta_i))$. In this case the moment conditions will be valid and equal to zero, as demonstrated below for $t = 5$:

---

7 Note that the break date needs to be towards the start of the sample because if towards the end of the sample then the initial moment conditions may be valid even in the presence of a break.

8 This means that equations (9a) and (9b) follow from equation (2) instead of equation (1) as in the text. For more see Wachter and Tzavalis (2004).

$$\mathrm{E}\big(\Delta y_{it-2}(\Delta v_{it} + \Delta\eta_i)\big) = \mathrm{E}(\Delta y_{i-2}(\Delta v_{it})) \tag{12a}$$

$$= \Big(\big(\alpha y_{it-2} + \delta_{iT_B} + v_{i3} + \eta_i - y_{it-2}\big)(\Delta v_{it})\Big) \tag{12b}$$

$$= \big((\alpha - 1)y_{it-2} + \eta_i\big)(\Delta v_{it})) + (\delta_{iT_B}\Delta v_{it})) = 0 \tag{12c}$$

The moment condition in (12) can be generalized as $E(\Delta y_{t-s}\Delta v_{it})$ where $S \geq 2$ and the instruments enter as lagged differences of the data. Moreover, if we relax the restriction in Section 2.2 that the fixed effects are not correlated with the error term so that $E\big(\delta_{T_B} v_{it}\big) \neq 0$ for $T_B < t$ then the moment condition $E(\Delta y_s(\Delta v_{it}))$ can also be used where $s \geq t + 2$. In this case the instruments enter as forward differences of the data. Thus for $t = 5$:

$$\mathrm{E}\big(\Delta y_{it+2}(\Delta v_{it})\big) = \Big[\Big((\alpha - 1)y_{it+1} + \delta_{T_B} + v_{it+1}\Big)(\Delta v_{it})\Big] = 0 \tag{13}$$

Table 1 summarises the moment conditions set out above and allows us to compare the GMM estimators according to the practical implications of their moment conditions. For the difference GMM estimator the instruments are lagged and remain in levels while the equation is estimated in difference form. The moment conditions of the system GMM estimator implies that the instruments are also lagged and are both in levels and differences and the respective equation are also estimated in levels and differences.[9] Finally, the moment conditions proposed in equations (12) and (13) the instruments and the equation are both in differences and this gives rise to the name 'double-D' which is short for double-difference. Furthermore, with the moment conditions of equation (12) the instruments are lagged whereas in equation (13) the instruments are forward terms (or leads) and thus gives rise to two estimators; namely the backward and forward double-D GMM estimators respectively. Note that if the autocorrelation of y_{it} is low a GMM estimator based on moment conditions (12) and (13) may result in weak instruments leading to biased estimates of the auto regressive parameter. Finally, we can combine the moment conditions of all four estimators in a full system GMM estimator.

[9] Note that with the system GMM the instruments could instead only include the lagged differences. See Blundell and Bond (1998).

3. A Monte Carlo Analysis of the GMM estimators

In this section we undertake Monte Carlo simulations to examine the bias associated with mean shifts on the five GMM estimators outlined above. Two sets of simulations are undertaken for each of the GMM estimators. In the first set the data are generated without breaks and in the second set two mean shifts that are explained below are included in the generated data. The data generating process is calibrated on United States individual bank loan growth data for the period 1993 to 2007 in terms of the mean, variance and sample length of that data.[10]

3.1 *Simulations without structural breaks*

We create a panel of data where the number of entities $N = 100$ and time periods $T = 15$. The data generating process is:

$$y_{it} = \alpha y_{it-1} + \eta_i + v_{it} \qquad (14a)$$

$$v_{it} = \delta v_{it-1} + \varepsilon_{it} \quad \text{and} \quad \delta = 0 \qquad (14b)$$

where η_i are randomly generated fixed effects and y_{it} are randomly generated data with mean 0.087 and standard deviation of 0.163. Simulations are repeated 1000 times for a range of α parameter values between 0.1 to 0.99 to retrieve the mean values of the estimated auto-regressive coefficients and associated standard errors.

To avoid the problem of over-fitting we do not use the full set of instruments/moment conditions when estimating the model.[11] Specifically, (i) the difference GMM estimator we use as instruments the third and fourth lags of y_{it}; (ii) the system GMM estimator the third and fourth lags of Δy_{it} and y_{it}; (iii) the backward double-D GMM estimator the third and fourth lags of Δy_{it-2}; (iv) the forward double-D GMM estimator the third and fourth leads of Δy_{it+2}; and (v) the full system estimator uses all the above instruments.

[10] See the data appendix for further details.

[11] The over-fitting problem is when a large set of instruments are individually valid but collectively invalid in finite samples because the number of instruments is greater than the number of entities. See Roodman (2006), Windmeijer (2005) and Ziliakc (1997).

Table 2 reports the mean estimates of the Monte Carlo simulations for the auto-regressive parameter, α, and the associated standard errors (in parentheses).[12] The bias measured as $\hat{\alpha} - \alpha$ is shown in square brackets. The table shows that when there are no mean shifts in the data the difference GMM and the system GMM estimators both perform well in an absolute sense and in the sense that the estimated values of α are less than two standard errors from their true values in the DGP. However, when the data is highly persistent and α is large and in the range of 0.8 to 0.99 the system GMM estimator outperforms the difference estimator as also reported in the simulations of Arellano and Bover (1995) and Blundell and Bond (1998). These results are consistent with the literature.

Table 2 also shows that without structural breaks the double-D estimators perform poorly relative to the other three estimators. This is because there is very low correlation between the instruments and the dependent variable as both enter as differences. Note however that the full system GMM which combines the moment conditions of all four GMM estimators (i.e. the difference, system and two double-D estimators) performs best and retrieves the data generating process to within 0.001 of the true value of α.

3.2 *Simulations with structural breaks*

Assuming the parameter values of the model are constant, there are two broad categories of breaks that are possible in the bank level data. The first are idiosyncratic breaks associated with each of the entities which in our case are banks. The second are common breaks across all entities. The most common cause of the first category of breaks is bank mergers which will introduce a one-off spike to the loan growth data and is not the type of structural break that we consider above. We therefore focus on the second category of breaks which we attribute to changes in policy and shifts in the business cycle.[13]

To calibrate the structural breaks in our generated data we apply the Bai-Perron multiple structural break test to the aggregate growth in loans data to obtain the number, weighted average size and dates of the breaks. Two significant break dates are found in the aggregated data.[14] The first is at $\boldsymbol{t = 5}$ which corresponds to 1997 in our dataset and the second is at $\boldsymbol{t = 10}$ which is the year

[12] Note that inference is unaffected by the use of the median rather than mean values of the estimates.

[13] The business cycle is generally thought to follow a stationary process. However, over finite samples the same cycle may look non-stationary and introduce a structural break to the bank lending data.

[14] Details of the Bai-Perron estimates are provided in the data appendix.

2002. The former is consistent with changes in United Sates bank regulations and the start of the 'boom' in the technology sector and the latter with the end of the technology bubble. The instruments, number of entities and time span remain the same as in our previous simulation.

We now generate a second panel of data which is identical to the first but incorporates the dates and magnitudes of the two Bai-Perron structural breaks identified in the aggregate data.[15] The DGP incorporating the structural breaks is;

$$y_{it} = \alpha y_{it-1} + \theta_1 DU_{t1} + \theta_2 DU_{t2} + \theta_3 DU_{t3} + \eta_i + v_{it} \qquad (15a)$$

$$v_{it} = \delta v_{it-1} + \varepsilon_{it} \text{ , } \delta = 0 \text{ and } \alpha < 1 \qquad (15b)$$

where, $\boldsymbol{\theta_1}$ is equal to 0.100 and $\mathbf{DU_{t1}}$=1 for $\boldsymbol{t \leq 4}$ and 0 in other periods, $\boldsymbol{\theta_2 = 0.085}$ and $\mathbf{DU_{t2} = 1}$ in $\boldsymbol{5 \leq t \leq 9}$ and 0 in other periods, and $\boldsymbol{\theta_3 = 0.076}$ and $\mathbf{DU_{t3} = 1}$ in $\boldsymbol{t \geq 10}$ and 0 in other periods.[16]

Table 3 presents the Monte Carlo results for the difference and system GMM without introducing shift variables to account for the structural breaks in the mean in the DGP. For both of these estimators we see that for values of α below 0.6 there is substantial and significant positive bias to the estimated values of α. For values of α between 0.6 and 0.99 however the bias is negative. These results demonstrate the non-linear nature of the bias introduced by the unaccounted breaks in mean as explained in Section 2. With low levels of persistence the total bias is positive because the Perron effect dominates. However, as α increases the negative bias due to the persistence itself increases along with the Nickell effect until the total bias becomes negative. With our generated data the total effect of the three biases 'cross-over' somewhere between the true values for α between 0.5 and 0.6.[17] Note however there is also a non-linearity in the negative range of the bias when the value of α approaches one. In this range the concept of a structural break in very highly persistent data becomes less relevant and in some sense is undefined in the limit when $\alpha = 1$.

15 The magnitude of the parameters θ_2 and θ_3 are Bai Perron estimates of the breaks.

16 Another way to proceed is to include shift dummies in the estimated model to account for the known structural breaks. However, if the magnitude of the break is different for each entity then one needs to include shift dummies for each individual entity which is not practical when the number of observations is small.

17 If the DGP incorporates larger shifts in mean then the 'cross-over' point is higher.

Table 3 also shows the double-D estimators, either using the leads or lags as instruments, outperforms the difference and system GMM estimators by a wide margin when $\alpha < 0.6$. However, for values of $\alpha \geq 0.6$ the system GMM estimator performs better in terms of the absolute size of the bias although the improvement is small and most likely to be insignificant.

These Monte Carlo results conform to our theoretical analysis in Section 2. When there are structural breaks the use of levels in the estimation is problematic because of the Perron effect. This explains why the double-D estimators which incorporate only differences dominate the other estimators which include levels in the estimation procedure. However, as the level of persistence increases the bias due to the breaks is reduced and so the advantage of using only differences is also reduced to the point where the system estimator outperforms the double-D estimator.

3.3 *A robustness check of the results*

Because of the finite nature of our generated data we are required to specify the lag structure of the instruments to avoid over-fitting the model. Furthermore, our Monte Carlo analysis has been constrained in other dimensions so as to conform to our annual bank lending data. Some observers may feel uncomfortable about our Monte Carlo analysis and wonder if the results are simply dependent on our modelling choices or are more 'global' in nature. To this end we undertake the following analysis of the robustness of our results.

First, to examine if our results depend on the choice of lags (and leads) of our instruments we re-run the Monte Carlo analysis for a range of lag structures for the instruments. The three panels of Figure 1 report the mean estimates of α for a range of lag structures for the instruments. The dotted line in each panel indicates the 'true' value of α from the DGP. Shown with square markers, circular markers, thick and thin lines are the mean estimates from the difference, system, double-D with backward lags and double-D with forward lags respectively.

The top panel shows graphically the results from Table 3. On a visual basis we can see that the double-D estimators dominate the difference and system estimators for values of $\alpha < 0.6$. The middle panel of the figure re-estimates the model only with the third lag (or lead where appropriate) as instruments. And finally the lower panel estimates the models using the third to fifth lags as instruments. We can see from all three panels that the double-D estimators perform better than the difference and system estimators at low levels of persistence but outperformed marginally by the system estimator at high levels of persistence.

Second we consider whether our results are dependent on the dimensions of the data set, in particular whether the number of periods (i.e. the size of T) and the number of entities (i.e. N) influence our results. We re-run the Monte Carlo analysis assuming $\mathrm{T} = 30$ and $\mathrm{T} = 60$. Given the position of the breaks may affect the results we also undertake the analysis assuming the breaks are in their initial positions (i.e. periods 5 and 10) and in the same relative positions in the data set (i.e. 10 and 20 when $\mathrm{T} = 30$ and at 20 and 40 when $\mathrm{T} = 60$). We find that in both cases the bias is reduced for all five estimators as might be expected but that the ranking of the estimators in terms of bias is unaffected by the longer data sets. We also find that increasing the number of entities by a factor of 5 has little impact on the bias or the ranking of the estimators.

We might conclude therefore that the Monte Carlo results reported above are largely unaffected by the choice of lags structures for the instruments and the dimensions of the data set.

4. An application of the double-D GMM estimator to the bank lending channel

Traditionally the transmission of monetary policy has been thought of in terms of the demand and supply side channels. The former includes the transmission through interest rates, exchange rates, the effect on the balance sheets of non-financial firms and the effect on the valuation of the firm's assets. In contrast, the supply side transmission of monetary policy focuses on the willingness of banks to lend which includes the bank lending channel.

The bank lending channel is difficult to identify in models using aggregate data and so researchers have turned more recently to the use of time series panel techniques to model this channel. The standard panel bank lending model is that of Kashyap and Stein (2000). This model attempts to identify how banks respond to changes in monetary policy by focusing on the heterogeneity among bank characteristics which can be incorporated in the panel analysis. However, the data employed in these panels contain structural breaks and therefore the estimates are subject to the biases discussed above to demonstrate the advantages and disadvantages of the five GMM estimators in the above analysis. We therefore estimate a Kashyap model of bank lending using the range of GMM estimators. The model is estimated with disaggregated United States bank level data for the period 1993 to 2007. The data appendix provides further details concerning the data.

4.1 The model

The Kashyap bank lending model is of the following form;

$$\Delta l_{it} = a_i + b\Delta l_{it-1} + \sum_{j=1}^{3} c_j\, \Delta R_{t-j} + d\, \Delta gdp_{t-1} + \sum_{j=1}^{2} e_j\, inf_{t-j}$$

$$+ g_0 \mathrm{LIQ}_{it-1} + g_1 \mathrm{SIZE}_{it-1} + g_2 \mathrm{CAP}_{it-1} + \textstyle\sum_{j=1}^{2} g_{3j}\, \mathrm{SIZE}_{it-1} \Delta \mathrm{R}_{t-j} \qquad (16)$$

$$+ \sum_{j=1}^{2} g_{4j}\, LIQ_{it-1} \Delta r_{t-j} + \sum_{j=1}^{2} g_{5j} CAP_{it-1} \Delta R_{t-j}$$

where the bank entity, i, with *N*= 5,820 and time, *t=1* to *16*. In the above equation l_{it} is total loans, R_t is the federal funds rate, $size_{it}$, cap_{it} and liq_{it} are the size of the balance sheet, capitalization and liquidity of individual banks respectively, gdp is gross domestic product measured at constant prices and inf_t is inflation. Lower case variables are in natural logarithms and Δ represents the change in the variable.

In the Kashyap model the growth in loans depends on two aggregate variables (i.e. the growth in GDP and prices) that represent the demand side of the economy and a range of characteristics of the individual banks. The lagged dependent variable models the dynamics in the data. The direct effects of monetary policy are represented in the model by the interest rate, R. The indirect effects of monetary policy are due to the interaction of changes in interest rates with the heterogeneous bank characteristics and these effects are incorporated in the model as multiplicative terms. We estimate the model using the five GMM estimators discussed above and our primary interest is the estimates of the lagged dynamic term and the indirect monetary policy effects captured by the multiplicative terms.

4.2 Results

Table 4 reports the long-run estimates of the bank lending model. Columns 1 to 5 report the models estimated with the difference, system, double-D backwards, double-D forwards and full system GMM estimators respectively. While there are some similarities in the estimates across the five estimators there are also some important differences. For example, if there are no breaks in the data then we know from the simulation results in Table 2 that the full system GMM estimates are the least biased by a considerable margin. In this case the estimated direct and indirect effects of monetary policy reported in column 5 of Table 4 are relatively small although they have the signs predicted in the monetary policy transmission literature.

However, there is every indication that the growth in bank lending data contains structural breaks. If we apply the difference and system GMM estimators to the model (see columns 1 and 2 in Table 4) we again obtain long-run estimates that are similar to the full system GMM estimates which in turn we believe to be poor because of the breaks in the data. Consequently, we might also question the validity of the difference and system GMM estimates. The double-D estimates reported in columns 3 and 4 indicate the direct and indirect effects of monetary policy on bank lending are substantially larger. For example, the effect of the size of the balance sheet on bank lending is around 10 times larger when the model is estimated with the double-D estimators than the estimates from the difference, system and full system estimators. Similarly, the direct effect of monetary policy is around 4 times larger when estimated with the double-D estimator. Note that as expected the residuals from the difference, system and full system models display second order serial correlation while the residuals from the models estimated with the double-D estimators are largely free of second order serial correlation.[18]

Finally, the dynamics of the models estimated with the double-D estimator appear more relevant than the estimated dynamics using the other estimators. The estimated coefficient on the lagged dependent variable in the difference, system and full system models are around -0.4. This suggests that the bank lending data is relatively slow to revert to its mean and that during convergence the data oscillates strongly about its mean. Given the models are estimated with annual data this description of the bank lending data appears difficult to sustain. In contrast, the double-D estimates suggest that the data are also mean reverting but the reversion is substantially quicker and the data does not routinely overshoot the mean on its path back to its mean. These differences in the estimates between the estimators are exactly as would be expected if the data contained structural breaks and the breaks are not adequately accounted for by the difference, system and full system estimators.

---

[18] Arellano-Bond type GMM estimators require that the error terms are serially correlated. If v_{it} in equation (1a) is serially uncorrelated then $\Delta\mathrm{v}_{\mathrm{it}}$ are correlated because $\mathrm{Cov}(\Delta\mathrm{v}_{\mathrm{it}}\Delta\mathrm{v}_{\mathrm{it-1}}) \neq 0$. However, $\Delta\mathrm{v}_{\mathrm{it}}$, will not be correlated with $\Delta\mathrm{v}_{\mathrm{it-k}}$ for $k \geq 2$.

5. Conclusion

The Monte Carlo analysis above suggests that if the researcher is confident that there are no structural breaks in the means of the data of the individual entities then the full system GMM estimator dominates all of the alternative estimators considered above in terms of lowest bias. This includes the standard difference and system GMM estimators commonly used in the literature.

However, when the data contains breaks in mean it is more complicated. The first difficulty is that when the panel data is of the 'large *N* relative to *T*' variety the individual graphing of the data is laborious and not practical. If the researcher is confident that there is a common break in the data of the entities then the double-D GMM estimator (estimated either with leads or lags for instruments) is the preferred option when estimated persistence is less than 0.6 and the system estimator when persistence is high.

We therefore, suggest the following methodology when estimating panels.

(i) Are there breaks in the data? This issue has two dimensions. First, has there been a change in the regulations or market structures that may lead to a break in the means of the data of the individual entities? Second, does the data span a short period of time so that the data appears non-stationary? For example, the business cycle may make some data look non-stationary over a few years when the same data are stationary over a longer span of time. It may be helpful to graph the aggregate data at this stage as an aid to understanding common breaks in the data.

If the researcher concludes that it is highly unlikely that there are breaks in the mean of the data then the full systems estimator that combines the moment conditions of the difference, system and both double-D GMM estimators should be applied to the data.

(ii) Breaks and Persistence. It is fortunate that none of the estimators considered above estimate the data to have low persistence when the true level of persistence is high. This implies that when choosing the 'correct' estimator the researcher does not need to know the 'true' level of persistence in the data and the estimated level of persistence can guide our choice when we believe there are structural breaks. Therefore, having decided that there are common breaks among the entities, the next stage is to estimate the model using the double-D estimator. However, if persistence is greater than 0.6 then the model should be re-estimated using the system estimator. We note that any improvement in the estimates over the double-D estimator may be minor.

Finally, the analysis above further demonstrates that ‘good’ empirical work is a sophisticated ‘art’. The researcher needs to understand the data they are working with and, importantly, the nature of any breaks that may be in the data. The analysis also suggests that there may be considerable benefit in undertaking Monte Carlo simulations similar to that above so as to understand the properties of the available estimators given the particular dimensions of the data set and prior beliefs concerning the nature of the breaks.

7. References

Arellano, M., and Bond, S., 1991.Some tests of specification for panel data: Monte Carlo evidence and application to employment equations, *Review of Economic Studies*, vol. 58, pp. 277–297.

Arellano, M.,and Bover O., 1995.Another look at the instrumental variable estimation of error components models, *Journal of Econometrics*, vol. 68, pp. 29-52.

Bai, J., and Carrion-i- Silvestre, J. Ll., 2009. Structural Changes, Common Stochastic Trends, and Unit Roots in Panel Data, *Review of Economic Studies*, vol. 76(2), pp. 471-501.

Bai, J., and Perron, P., 2003. Computational and analysis of multiple structural-change models. *Journal of Applied Econometrics*, vol. 18(1), pp. 1-22.

Blundell, R., and Bond, S., 1998. Initial conditions and moment restrictions in dynamic panel models. *Journal of Econometrics*, Vol. 87, pp. 115-143.

Carrion-i-Silvestre, J. Ll., del Barrio-Castro T. and López-Bazo, E., 2002. Level shifts in a panel data based unit root test: an application to the rate of unemployment, Proceedings of the 2002 North American Summer Meetings of the Econometric Society: Econometric Theory, edited by D.K. Levine and W.Zame. http://www.dklevine.com/proceedings/theoretical-econometrics.htm.

Carrion i-Silvestre, J. Ll., del Barrio –Castro T., Lopez-Bazo, E., 2005. Breaking the panels: An application to the gdp per capita, *Econometrics Journal*, vol. 8, pp. 159-175.

Hayakawa, K., 2009. On the effects of mean non–stationarity in dynamic panel data models. *Journal of Econometrics*, vol. 8, pp. 159-175.

Kashyap, A., and Stein J. C., 2000. What do a million observations on banks say about the transmission of monetary policy, *American Economic Review*, vol. 90(3), pp. 407-428.

Nickell, S., 1981. Biases in dynamic models with fixed effects, *Econometrica,* vol. 49(6), pp. 1417-1426.

Perron, P., 1989. The great crash, the oil price shockand the unit root hypothesis. *Econometrica*, vol. 57, pp. 1361-1401.

Roodman, D., 2006. How to do Xtabond2: An introduction to difference and system GMM in Stata. Centre for Global Development Working paper series, No. 103.

Roodman, D., 2008. A note on the theme on too many instruments. Centre for Global Development Working paper series, No. 125.

Wachter, S. ,Tzavalis, E. , 2004. Detection of structural breaks in linear dynamic panel data model. Queen Mary, University of London Working paper series, No. 505.

Windmeijer, F., 2005. A finite sample correction for variance of linear efficient two-step GMM estimators, *Journal of Econometrics*, vol. 26(1), pp. 25-51.

Ziliak, J. P., 1997. Efficient estimation with panel data when instruments are predetermined: An empirical comparison of moment–condition estimators. *Journal of Business and Economic Statistics*, vol. 15(4), pp. 419-431.

APPENDIX 1: DATA APPENDIX.

Balance sheet items are measured at the end of the December quarter each year and from the Federal Reserve Bank of Chicago (www.chicagofed.org). The data were downloaded between 25th October 2009 and 10th November 2009. Total loans (mnemonic Rcfd1400) are defined as the aggregate gross book value of total loans (before deduction of valuation reserves) including (i) acceptances of other banks and commercial paper purchased in open market, (ii) acceptances executed by or account of reporting bank and subsequently acquired by it through purchases or discount, (iii) customer's liability to reporting bank on drafts paid under letter of credit for which bank has not been reimbursed, and (iv) all advances. The data are in natural logarithms. All data and the Stata 'do files' are available at www.billrussell.info.

The Bai and Perron (2003) approach minimises the sum of the squared residuals to identify the number and dates of k breaks in the model: $\Delta l_t = \gamma_{k+1} + \tau_t$ where Δl_t is the annual change in the natural logarithm of total loans, γ_{k+1} is a series of $k+1$ constants that estimate the mean growth of loans in each of $k+1$ 'regimes' where the mean is constant in a statistical sense and τ_t is a random error. The model is estimated with a minimum regime size (or 'trimming') of 5 years out of a total sample of 15 years. The final model is chosen using the Bayesian Information Criterion. The model is estimated for the period 1993 to 2007. The results of the estimated model are reported in the table below. The Bai-Perron technique was estimated using Rats 7.2 using baiperron.src and multiplebreaks.src written by Tom Doan and kindly made available on the Estima internet site.

Table A1: Estimated breaks in the Growth in Total Loans

Regime	*Dates of the 'Regimes'*	*Mean Growth Rate of Loans*
1	1993 - 1997	0.0996
2	1998 - 2002	0.0858
3	2003 - 2007	0.0761

Figure 1: GMM estimators assuming different lag structures

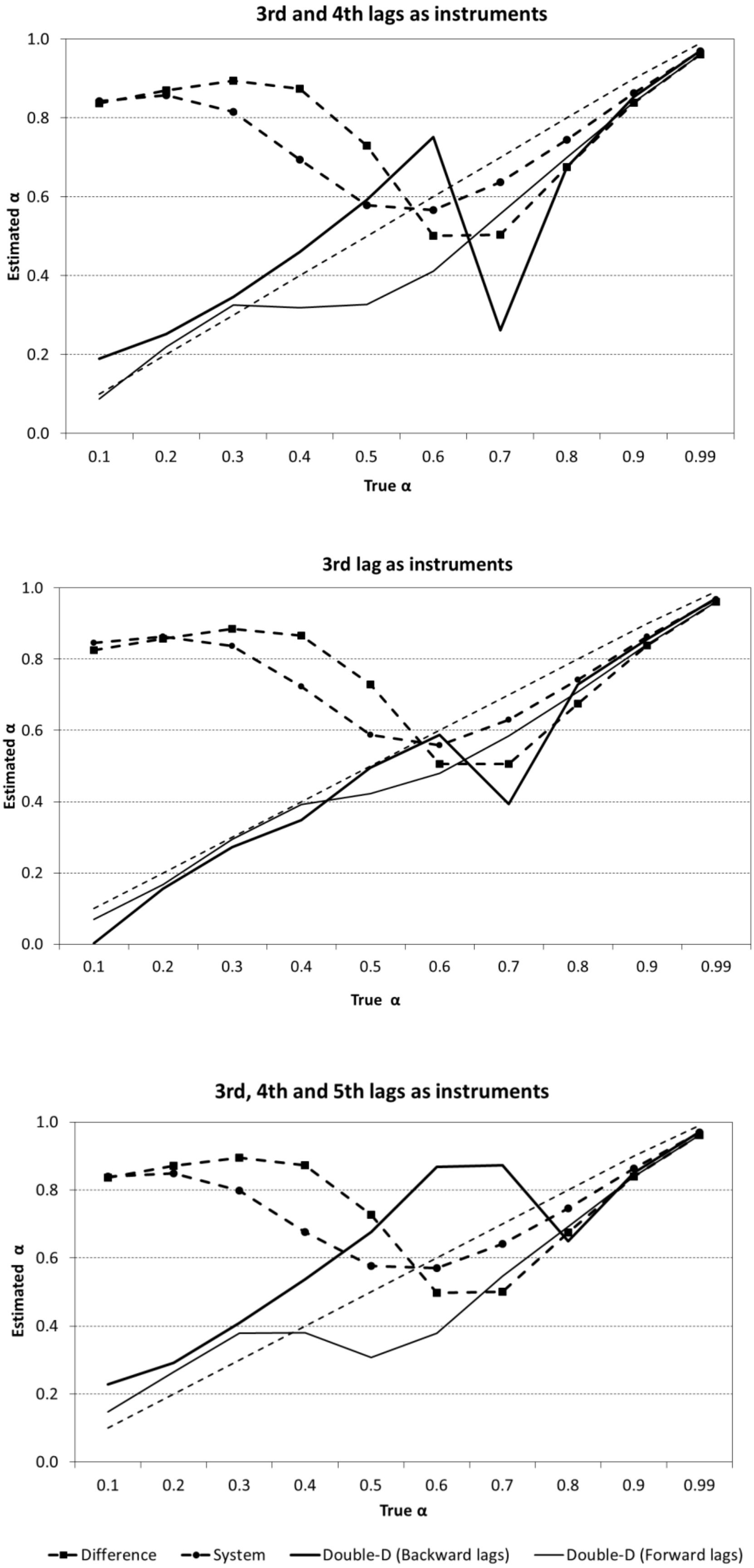

Table 1: Moment conditions used in each of the GMM estimators

	Moment conditions				
1	$E(y_{it-s}\,\Delta u_{it}) = 0; \qquad (2 \leq \mathrm{S} \leq \mathrm{t} - 1)$				
2	$E[\Delta y_{it-1}(v_{it} + \eta_i)] = 0$				
3	$E(\Delta y_{t-s} \Delta v_{it}) = 0; \qquad (S \geq 2)$				
4	$E\big(\Delta y_s (\Delta v_{it})\big) = 0; \qquad (S \geq t + 2\ )$				
	Moment conditions used in each estimator				
	Difference	*System*	*Double-D (backward lag)*	*Double-D (forward lag)*	*Full system*
	1	1, 2	3	4	1, 2, 3, 4

Table 2: Monte Carlo results assuming no structural breaks

	Mean $\hat{\alpha}$				
True α	**Difference**	**System**	**Double-D (backward lags)**	**Double-D (forward lags)**	**Full System**
0.1	-0.166 (0.172) [-0.266]	0.086 (0.133) [-0.014]	-0.254 (0.225) [-0.354]	0.077 (0.054) [-0.023]	0.099 (0.029) [-0.001]
0.2	0.030 (0.139) [-0.17]	0.187 (0.109) [-0.013]	-0.197 (0.228) [-0.397]	0.171 (0.060) [-0.029]	0.199 (0.029) [-0.001]
0.3	0.319 (0.095) [0.019]	0.289 (0.089) [-0.011]	-0.138 (0.222) [-0.438]	0.262 (0.067) [-0.038]	0.298 (0.029) [-0.002]
0.4	0.434 (0.085) [0.034]	0.391 (0.073) [-0.009]	-0.072 (0.223) [-0.472]	0.349 (0.075) [-0.051]	0.398 (0.029) [-0.002]
0.5	0.538 (0.080) [0.038]	0.492 (0.060) [-0.008]	0.010 (0.216) [-0.490]	0.429 (0.084) [-0.071]	0.498 (0.028) [-0.002]
0.6	0.625 (0.085) [0.025]	0.593 (0.049) [-0.007]	0.107 (0.211) [-0.493]	0.492 (0.102) [-0.108]	0.598 (0.027) [-0.002]
0.7	0.617 (0.122) [-0.083]	0.695 (0.038) [-0.005]	0.200 (0.216) [-0.500]	0.525 (0.102) [-0.175]	0.699 (0.024) [-0.001]
0.8	0.747 (0.180) [-0.053]	0.798 (0.027) [-0.002]	0.230 (0.211) [-0.570]	0.448 (0.124) [-0.352]	0.800 (0.019) [0.000]
0.9	0.710 (0.106) [-0.19]	0.900 (0.016) [0.000]	0.290 (0.218) [-0.610]	0.339 (0.175) [-0.561]	0.900 (0.013) [0.000]
0.99	0.981 (0.026) [-0.009]	0.990 (0.009) [0.000]	0.907 (0.096) [-0.083]	0.894 (0.080) [-.0096]	0.991 (0.008) [0.001]

Notes: The simulations where undertaken in Stata 11.1 with a 'seed' value of 1010. See Section 3.1 for details concerning the generation of the data. Shown in ( ) are the mean standard errors of $\hat{\alpha}$. Shown in [ ] is the estimated bias.

Table 3: Monte Carlo simulation results assuming structural breaks

	Mean $\hat{\alpha}$				
True $\hat{\alpha}$	**Difference**	**System**	**Double-D (backward lags)**	**Double-D (forward lags)**	**Full System**
0.1	0.837 (0.046) [0.737]	0.842 (0.042) [0.742]	0.190 (0.119) [0.090]	0.088 (0.059) [0.012]	0.606 (0.030) [0.506]
0.2	0.874 (0.041) [0.674]	0.857 (0.037) [0.657]	0.252 (0.087) [0.052]	0.224 (0.070) [0.024]	0.702 (0.030) [0.502]
0.3	0.895 (0.037) [0.595]	0.816 (0.031) [0.516]	0.346 (0.074) [0.046]	0.326 (0.072) [0.026]	0.777 (0.029) [0.477]
0.4	0.874 (0.033) [0.474]	0.695 (0.023) [0.295]	0.460 (0.066) [0.060]	0.319 (0.053) [-0.081]	0.801 (0.028) [0.401]
0.5	0.730 (0.027) [0.230]	0.579 (0.014) [0.079]	0.592 (0.062) [0.092]	0.327 (0.033) [-0.173]	0.703 (0.024) [0.203]
0.6	0.501 (0.017) [-0.099]	0.566 (0.008) [-0.034]	0.753 (0.082) [0.153]	0.411 (0.019) [-0.189]	0.500 (0.016) [-0.100]
0.7	0.504 (0.009) [-0.196]	0.637 (0.005) [-0.063]	0.261 (0.230) [-0.439]	0.557 (0.010) [-0.143]	0.505 (0.009) [-0.195]
0.8	0.675 (0.004) [-0.125]	0.744 (0.003) [-0.056]	0.675 (0.020) [-0.125]	0.700 (0.006) [-0.100]	0.675 (0.004) [-0.125]
0.9	0.839 (0.002) [-0.061]	0.863 (0.001) [-0.037]	0.852 (0.004) [-0.048]	0.840 (0.002) [-0.060]	0.839 (0.002) [-0.061]
0.99	0.961 (0.001) [-0.029]	0.969 (0.000) [-0.021]	0.971 (0.001) [-0.029]	0.961 (0.001) [-0.039]	0.961 (0.001) [-0.039]

Notes: Reported are the mean values of $\hat{\alpha}$ from the Monte Carlo simulations. See Section 3.2 for details concerning the generation of the data. See also notes to Table 2.

Table 4: United States Estimates of the Bank Lending Channel

	1	2	3	4	5
Variables	**Difference**	**System**	**Double-D (backward lags)**	**Double-D (forward lags)**	**Full System**
Δl_{it-1}	-0.469** (0.000)	-0.409** (0.000)	0.126** (0.057)	0.097** (0.097)	-0.397** (0.611)
Long-Run Coefficients					
Δgdp_{t-1}	0.779** (0.000)	0.738** (0.738)	0.020 (0.213)	0.200 (0.179)	0.709** (0.000)
inf_{t-j}	0.003* (0.001)	0.002** (0.000)	0.002** (0.001)	0.006** (0.001)	0.002** (0.000)
R_{t-j}	-0.007** (0.000)	-0.004** (0.063)	-0.018** (0.000)	-0.0176** (-0.017)	-0.005* (0.000)
$\mathrm{SIZE}_{it-1}\Delta \mathrm{R}_{t-j}$	0.007* (0.004)	0.020** (0.003)	0.051** (0.008)	0.044** (0.000)	0.020** (0.004)
$LIQ_{it-1}\Delta r_{t-j}$	0.004 (0.004)	0.016** (0.004)	0.001 (0.005)	0.001 (0.005)	0.015** (0.004)
$CAP_{it-1}\Delta R_{t-j}$	0.006 (0.074)	0.043 (0.063)	0.126* (0.075)	0.172** (0.086)	0.037 (0.059)
Diagnostics – probability values					
J-Stat	0.180	0.390	0.630	0.055	0.060
AR(1)	0.781	0.005	0.000	0.000	0.009
AR(2)	0.001	0.000	0.188	0.191	0.000

Notes:** significant at 5% level,* significant at 10% level. Standard errors reported as ( ). Dependent variable is $\mathbf{\Delta l_{it}}$. Long-run values calculated as the sum of the estimated coefficients divided by 1 minus the coefficient on the lagged dependent term. Associated long-run standard errors are calculated using Taylor series progression. J-Stat, AR(1) and AR(2) are the Hansen J statistic of moment condition over-identification and Arellano-Bond tests of auto-correlated residuals of order 1 and 2 respectively (see also footnote 18). A linear trend is included in the models which are estimated using Stata 11.1.