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

Rosen Azad Chowdhury† and Bill Russell#

10 August 2017

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, bank lending channel

JEL code: C23, C22, G21

---

* †Corresponding author, School of Management, Swansea University, email: r.a.chowdhury@swansea.ac.uk. #University of Dundee School of Business, Dundee, DD1 4HN, Scotland, United Kingdome, 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 when applied to mean stationary data. 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 when the data are exogenous. 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 are endogenous. We also propose two new moment conditions for the Generalised Method of Moments (GMM) estimator that reduces the bias substantially when the dynamic panels contain structural breaks.

The effects of structural breaks in dynamic panels that incorporate endogenous variables are complicated. 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 biases outlined in Carrion-i-Silvestre *et al.* (2005) and Bai and Carrion-i-Silvestre (2009) when the variables are strictly exogenous. 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 demonstrating the moment conditions for the difference and system (which is the combination of difference and level) GMM estimators 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 when there are structural breaks in the data then for panels where the length of the data is short the double-D estimator out performs the difference and system GMM estimators for low levels of persistence (i.e. autoregressive coefficients less

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

than 0.6) and the difference and system GMM estimators perform marginally better when persistence is high. However, with longer length panels the double-D GMM estimator outperforms both the difference and system estimators even when the data are highly persistent. 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

2.1 *Structural Breaks and the difference GMM Estimator*

Consider the following AR(1) process before the break in period T_B;

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

where ‘i' represents the panel entity, η_i is an entity specific fixed effect and v_{it} is a random error term.

In keeping with Wachter and Tzavalis (2012) we assume that exogenous shocks have a heterogeneous impact on the unobservable fixed effects so that breaks may be correlated with past shocks and also with the fixed effects. Since the unobserved fixed effects are not identical across entities before the breaks the Wachter and Tzavalis assumption implies that following a shock the fixed effects may not all change by the same magnitude. In this case our AR(1) process after the break can be written;

$$y_{it} = \alpha y_{it-1} + (\eta_i + \delta_{iT_B}) + v_{it},\; t \geq T_B \tag{2}$$

where δ_{iT_B} is the mean shifts in the fixed effects caused by exogenous shocks to entity ‘i' and T_B is the break date where we assume that $T_B = 3$.

Assuming that $E\left(\delta_{iT_B} v_{it}\right) \neq 0$ for $t < T_B$, $E\left(\delta_{iT_B} v_{it}\right) = 0$ for $t \geq T_B$ and $E(\delta_{iT_B}\eta_i) \neq 0$ then for $t = 4$, the moment conditions for the Arellano Bond (1991) difference GMM estimator proposed when there are no structural breaks in the data is;

$$E[y_{it-s}(\Delta v_{it})] = 0 \;\; for \;\; t = 3,\; 4 \ldots . T \; and \; 2 \;\leq\; S \;\leq\; t-1 \tag{3}$$

However, if there are structural breaks the moment conditions in equation (3) will not be valid as additional terms will appear in the moment condition. In the presence of a structural break the differenced GMM moment conditions when $t = 4$ are:

$$E[y_{it-s}(\Delta v_{it})] \quad (4a)$$

$$= E[y_{i2}(\{y_{i4} - \alpha y_{i3} - \eta_i - \delta_{iT_B}\} - \{y_{i3} - \alpha y_{i2} - \eta_i\})] \quad (4b)$$

$$= E[y_{i2}\Delta y_{i4}] - \alpha E[y_{i2}\Delta y_{i3}] - E[y_{i2}\delta_{iT_B}] \quad (4c)$$

Revealing an additional term, $E[y_{i2}\delta_{iT_B}]$ in equation (4c) which is not equal to zero and hence the moment condition may be invalid and the estimator biased.

This bias has a number of underlying sources. 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 when the data are exogenous. They start with an AR(1) process, y_{it}, with a single level shift;

$$y_{it} = \alpha y_{it-1} + \theta_i DU_{i,t} + \eta_i + v_{it} \quad (5)$$

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. Note that equations (5) is equivalent to equations (1) and (2) above.

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

$$y_{it} = \alpha y_{it-1} + \eta_i + v_{it} \quad (6)$$

then the least square estimate of α is;

$$\hat{\alpha} = \alpha + \underbrace{\frac{\sum_{i=1}^{N} y' Q_T v_i}{\sum_{i=1}^{N} y' Q_T y_{i,-1}}}_{NE} + \underbrace{\frac{\sum_{I=1}^{N} y'_{i,-1} Q_T z \zeta_i}{\sum_{i=1}^{N} y'_{i,-1} Q_T\, y_{i,-1}}}_{PE} \quad (7)$$

where, $Q_{T=}I_T - X(X'X)^{-1}X'$, X' being the $T \times K$ matrix such that $X = (1, \ldots..1)' = e_T$, $Z = [e_{T,}DU]$ and $\zeta = (\alpha_i, \theta_i)'$ with θ is the magnitude of the break. Equation (7) indicates 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 (7). 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. The bias identified as *PE* in equation (7) 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.

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 (5) we ignore the shift term and assume the process is as described by equation (6). In this case the standard difference GMM (Arellano Bond 1991 type) orthogonal moment conditions can be written;

$$E(y_{it-s}\,\Delta v_{it}) = 0 \;\; for \;\; t = 3,\; 4 \ldots . T \; and \; 2 \;\leq\; S \;\leq\; t-1 \qquad (8)$$

The literature avoids the use of matrices and greatly simplifies the exposition by assuming $T = 3$ so that the moment condition in equation (8) is exactly identified and the corresponding method of moments estimator reduces to a two stage least square estimator. If $T > 3$ then the following results also apply to the other moment conditions with the difference GMM estimator. For $T = 3$ the first stage of the instrumental variable regression is;

$$\Delta y_{it} = (\alpha - 1) y_{it-1} + \eta_i + v_{it} \qquad (9a)$$

$$\Delta y_{it} = \beta y_{it-1} + \eta_i + v_{it} \qquad (9b)$$

where $\beta = (\alpha - 1)$. The least squares estimator of equation (9b) is then;

$$\hat{\beta} = (\alpha - 1) + NE + PE \qquad (10)$$

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 (9a) 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 (10).[2]

Consequently, the total bias is non-linear and depends on the relative magnitudes of the three biases. In equation (10) 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 greater than the positive Perron effect

2 Note however, some of the moment conditions prior to the break may remain valid in difference GMM. For more see Wachter and Tzavalis (2012).

then $\hat{\beta}$ will be biased downwards and the instruments will be less correlated with the Δy_{it} term.[3] Therefore, when estimating the model without accounting for the structural breaks the instruments may become invalid with the differenced 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 (6) using the system GMM estimator and it is this estimator that we now turn to.

2.2 *Structural Breaks in the Level and System GMM Estimators*

Arellano and Bover (1995) and Blundell and Bond (1998) demonstrate that when the data is persistent (i.e. when $\alpha \geq 0.8$) the difference GMM estimator performs poorly for the reasons explained above. Consider again equations (1) and (2) where δ_{i,T_B} is the mean shifts in the fixed effects caused by exogenous shocks and T_B is the break date where we assume that $T_B = 3$.[4] Assuming that $E(\delta_{i,T_B} v_{it}) \neq 0$ for $t < T_B$, $E(\delta_{i,T_B} v_{it}) = 0$ for $t \geq T_B$ and $E(\delta_{i,T_B} \eta_i) \neq 0$ then for $t = 4$, the level moment condition if there are no structural breaks in the mean of the data is;

$$E[\Delta y_{it-1}(v_{it} + \eta_i)] = 0 \qquad (11)$$

where Δy_{it-1} is the instrument and the equation is estimated in levels. However, if there are unaccounted breaks then the moment conditions in equation (11) 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:

$$E[\Delta y_{it-1}(v_{it} + \eta_i)] = [(y_{i3} - y_{i2})(v_{i4} + \eta_i)] \qquad (12a)$$

$$= E[(\alpha y_{it-2} + \delta_{i,T_B} + v_{i3} + \eta_i - y_{it-2})(v_{i4} + \eta_i)] \qquad (12b)$$

$$= E[((\alpha - 1)y_{it-2} + \eta_i)\eta_i] + E(\delta_{i,T_B} \eta_i) \qquad (12c)$$

Equation (12c) differs from the standard moment condition of no structural breaks level moment conditions of equation (11) by the term, $E(\delta_{i,T_B} \eta_i)$ 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. From equation (4c) and

---

3 Hayakawa (2009) argues that if the data is mean non-stationary the moment conditions of the difference GMM may be valid even when the auto regressive parameter is high. This is due to the fixed effects providing additional correlation between the lagged dependent variable and the instruments when persistence is high which may improve the finite sample behaviour and reduces the weak instrument problems of the difference GMM estimator.

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

now from equation (12c) it is evident that the difference and level moment conditions are respectively invalid when there are breaks in the data. And thus we can conclude that the system GMM which is a combination of the difference and level moment conditions is also invalid. This also implies that with structural breaks the system GMM initial moment conditions will not decay towards its long run mean set by the parameter α in equations (1) and (2).

2.3 *The Double-D GMM Estimator*

One solution to overcome the bias due to the presence of a common break in the data is to split the panel at the break date into two sections: one prior and the other after the break date. The sections can then be 'stacked' into a larger unbalanced panel before estimating the model assuming the coefficients are the same before and after the break.[5] This approach is attractive when the time periods in both sections are large enough for the moment conditions to be valid in both of the sections in the stacked panel. This in turn allows the assumption that the expected value of the estimated coefficients to be the same in both sections to be valid.[6]

However, consider cases where (i) the break occurs either close to the beginning or the end of the panel, (ii) there are multiple breaks where the time span between the breaks is small, or (iii) the breaks are not common among the panel then the 'stacking' solution is not appropriate because the orthogonality conditions are not valid. Instead what is needed are further moment conditions to deal with these cases.

For example, the problem caused by unaccounted common structural breaks in the system GMM can be resolved by changing the moment condition in equation (11) to $E\big(\Delta y_{it-2}(\Delta v_{it} + \Delta\eta_i)\big)$. In this case the moment conditions will be valid and equal to zero, as demonstrated below for $t = 5$:

$$E\big(\Delta y_{it-2}(\Delta v_{it} + \Delta\eta_i)\big) = E(\Delta y_{i-2}(\Delta v_{it})) \tag{13a}$$

$$= E\left(\left(\alpha y_{it-3} + \delta_{i,T_B} + v_{i3} + \eta_i - y_{it-2}\right)(\Delta v_{it})\right) \tag{13b}$$

$$= E\big((\alpha - 1)y_{it-3} + \eta_i\big)(\Delta v_{i5})) + (\delta_{iT_B}\Delta v_{i5})) = 0 \tag{13c}$$

---

[5] See Russell (2011) who applies this 'stacking' approach to a time series of inflation data when estimating short-run Phillips curves for the United States.

[6] Consider the case when the true value of α is the same in both sections 1 and 2 of the data. The expected value of α in each section is therefore $\alpha + section\ specific\ bias$. The expected value of the section specific bias in each section is only the same if the time span in both sections are the same (i.e. the break occurs in the middle of the data) or if the time span in each section is sufficiently large that the expected value of the bias is insignificant.

The moment conditions in (13) can be generalized as $E(\Delta y_{i,t-s}\Delta v_{it})$ 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 $E(\delta_{i,T_B} v_{it}) \neq 0$ for $T_B < t$ then the moment condition $E(\Delta y_{i,s}(\Delta v_{it}))$ can also be used where $S \geq t+2$. In this case the instruments enter as forward differences of the data.[7] Thus for *t=5:*

$$E(\Delta y_{i7}(\Delta v_{i5})) = E\left[\left((\alpha-1)y_{i6} + \delta_{i,T_B} + \eta_i + v_{i6}\right)(\Delta v_{i5})\right] = 0 \qquad (14)$$

Importantly, because the moment conditions are valid for any value of t then this solution is valid for breaks that are either common, or not common, across the entities of the panel.[8]

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 differences and levels.[9] The moment conditions proposed in equations (13c) and (14) 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 (13c) the instruments are lagged whereas in equation (14) 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 (13c) and (14) 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.

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

---

7 Some may argue that the forward moment condition for the double-D estimator in equation (14) is invalid because Δv_{it} may be correlated with Δy_s when $S \geq t+2$. However, this issue is relevant in all moment conditions including those used in Arellano and Bond (1991) and Blundell and Bond (1998) if the variables are mean stationary because $E[y_{it}] = E[y_{it+1}] = E[y_{it-1}]$.

8 The stacked estimator mentioned in the introduction is consistent whenever the double-D estimator is consistent. However, the double-D estimator is likely to be more efficient due to using more moment conditions but suffer from a finite sample bias.

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

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 (DGP) is:

$$y_{it} = \alpha y_{it-1} + \eta_i + v_{it} \tag{15}$$

Where $\alpha < 1$ and η_i are randomly generated fixed effects assuming a mean zero normal distribution with unit variance, ε_{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}; and (v) the full system estimator uses all the above instruments.

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.

---

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 (2008), Windmeijer (2005) and Ziliakc (1997).

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

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 known 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. The second are breaks in mean as considered in the analysis above and may be due to changes in policy and shifts in the business cycle. 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 in the mean growth rates of the bank lending data.

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. Details of the Bai-Perron estimates are provided in the data appendix. The first is at $t = 5$ which corresponds to 1997 in our dataset and the second is at $t = 10$ which is the year 2002. The former coincides 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 structural breaks identified in the aggregate data using the Bai-Perron technique.[13] The DGP incorporating the structural breaks is;

$$y_{it} = \alpha y_{it-1} + \theta_1 DU_{1it} + \theta_2 DU_{2it} + \theta_3 DU_{3it} + \eta_i + v_{it} \text{ and } \alpha < 1 \qquad (16)$$

where, θ_1 is equal to 0.100 and DU_{1it}=1 for $t \leq 4$ and 0 in other periods, $\theta_2 = 0.085$ and $DU_{2it} = 1$ in $5 \leq t \leq 9$ and 0 in other periods, and $\theta_3 = 0.076$ and $DU_{3it} = 1$ in $t \geq 10$ and 0 in other periods.[14]

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

---

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

14 Another way to proceed is to include shift dummies in the estimated model to account for the common structural breaks if known. 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.

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.[15] 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 $\alpha = 1$.

Table 3 also shows the double-D estimators using either the leads or lags as instruments outperform 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.

Finally, for comparison purposes, the last two columns of Table 3 reports estimates from the stacked solution for the difference and system GMM estimators. As expected the stacked solutions to the breaks in the data outperform the Difference and System estimates over the full range of the true value of α in terms of smaller bias and standard errors. Table 3 also reveals the stacked solution marginally outperforms the double-D solutions when there are common breaks. However, the double-D solutions are also valid when the break dates are unknown and we turn to this issue in the next section.

3.3 *Simulations with structural breaks of unknown break dates*

In this section we report Monte Carlo analyses for two forms of breaks where the dates are unknown. The first are what we term 'fuzzy' breaks where there is a general break in the data but the break appears in the data for each entity with some variation. In our known break date case above, the first break is in T=5 and our corresponding 'fuzzy' case we distribute the breaks randomly for each entity in periods T=4, 5 or 6. Similarly, the second common break is in T=10 and in our 'fuzzy' case the break appears randomly in

---

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

periods T=9, 10 or 11 for each entity. The results are reported in Table 4 and we see little meaningful difference in terms of the bias and rankings of the estimators from the common break case reported in Table 3.

The second form of breaks is more unconstrained with 2 breaks occurring at random for each entity between periods T=2 and T=14. The results shown in Table 5 again indicate the double-D estimators outperform the difference and system GMM estimators for low levels of persistence in the same way as when there are common breaks across the entities.

3.4 *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. We can see from all three panels in Figure 1 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 $T = 60$ with common breaks in the same relative positions in the data of periods 20 and 40. The results are shown in Table 6 where we continue to see the double-D estimators dominate both the difference and system estimators at low levels of persistence in the data. However, the double-D GMM estimator now also dominates the difference and systems GMM estimators at high levels of persistence. In contrast, increasing the number of entities, N, from 100 to 500 keeping T=15 has little effect on the biases and the relative performance of the estimators at low and high levels of persistence.[16] It appears that increasing the length of the panel, T, has an important effect in reducing the bias of the double-D GMM estimator while increasing the number of entities N has little effect. We might conclude therefore that the general Monte Carlo results reported above are not due to the choice of lag structure for the instruments and the dimensions of the data set.

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

Changes in monetary policy affects both loan demand and loan supply simultaneously making it harder to identify whether a change in lending is due to supply or demand side influences. As a result the bank lending channel of the monetary policy transmission mechanism channel is difficult to identify in models using aggregate data and so researchers have turned 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 capture banks loan supply shifts due to changes in monetary policy by focusing on the heterogeneity among bank characteristics which can be identified by the use of the panel analysis. However, the data employed in these panels contain structural breaks and therefore the estimates are subject to the biases discussed above. We therefore estimate a Kashyap and Stein model of bank lending using the range of GMM estimators discussed above and disaggregated United States bank level data for the period 1992 to 2007 to demonstrate the advantages and disadvantages of the individual GMM estimators. The data appendix provides further details concerning the data.

4.1 The model

The Kashyap and Stein (2000) bank lending model is of the following form;

$$\Delta l_{it} = \alpha_i + \beta \Delta l_{it-1} + \sum_{j=1}^{3} \gamma_j \, \Delta R_{t-j} + \delta \, \Delta gdp_{t-1} + \sum_{j=1}^{2} \zeta_j \, inf_{t-j}$$
$$+\theta_0 LIQ_{it-1} + \theta_1 SIZE_{it-1} + \theta_2 CAP_{it-1} + \textstyle\sum_{j=1}^{2} \theta_{3j} \, SIZE_{it-1} \Delta R_{t-j} \qquad (17)$$
$$+ \sum_{j=1}^{2} \theta_{4j} \, LIQ_{it-1} \Delta R_{t-j} + \sum_{j=1}^{2} \theta_{5j} CAP_{it-1} \Delta R_{t-j} + v_{it}$$

where, the bank entity, i, with *n*= 5,820 and time, *t=1...........15.* 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

[16] The results are available from www.billrussell.info. We also considered T=30 with breaks occurring at the positions of periods 10 and 20. The general finding are part way between the T=15 results of Table 3 and the T=60 results of Table 6.

individual banks respectively, gdp_t 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 and Stein 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. A lagged dependent variable allows the model to capture 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 7 reports estimates of the lagged dependent variable and the long-run coefficients of the bank lending model.[17] 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 7 are relatively small although they have the signs predicated 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 7) we again obtain long-run estimates that are similar to the full system GMM estimates which we in turn believe to be poor because of the breaks in the data. Consequently, we might also question the validity of these estimates. The double-D estimates reported in columns 3 and 4 indicate the direct and indirect effects of monetary policy on banking 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,

---

17 The long run elasticity of lending with respect to monetary policy for the average bank is given by effect, $\frac{\sum_{j=1}^{3}\gamma_j}{(1-\beta)}$, while that with respect to the interaction term between capital and monetary policy is represented by $\frac{\sum_{j=2}^{2}\theta_{5j}}{(1-\beta)}$. Associated long-run standard errors are calculated using Taylor series progression.

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 serial correlation.

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 behaviour 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 range of GMM 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.

4.3 Applied Methodology

Based on the analysis above we suggest the following methodology when estimating panels.

(i) Are there 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 breaks among the entities, the next stage is to estimate the model using the double-D estimator. If estimated persistence using the double-D estimator is greater than 0.6 then the model should be re-estimated using the system estimator however we note that any improvement in the estimates over the double-D estimator may be minor. Note also that as the length of the panel data increases the performance of the double-D GMM estimators improve relative to the alternative GMM estimators.

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 estimators commonly used in the literature. However, when the data contains breaks in mean it is more complicated. If the researcher is confident that the entities experience breaks 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.

7. Reference

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.

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, 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 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, pp159-175.

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 shock and the unit root hypothesis. *Econometrica*, Vol. 57, pp 1361-1401.

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

Russell, B. (2011). Non-stationary Inflation and Panel Estimates of United States Short and Long-run Phillips Curves, *Journal of Macroeconomics*, vol. 33, pp. 406-19.

Wachter, S., Tzavalis, E. , 2012. Detection of structural breaks in linear dynamic panel data model. *Computational Statistics and Data Analysis,* vol. 56(1), pp.3020-3034.

Windmijer, 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 for 5,820 bank entities have been taken from Federal Reserve Bank of Chicago (www.chicagofed.org). The balance sheet items measured at the end of the December quarter each year. The data are 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 (1998) 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 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

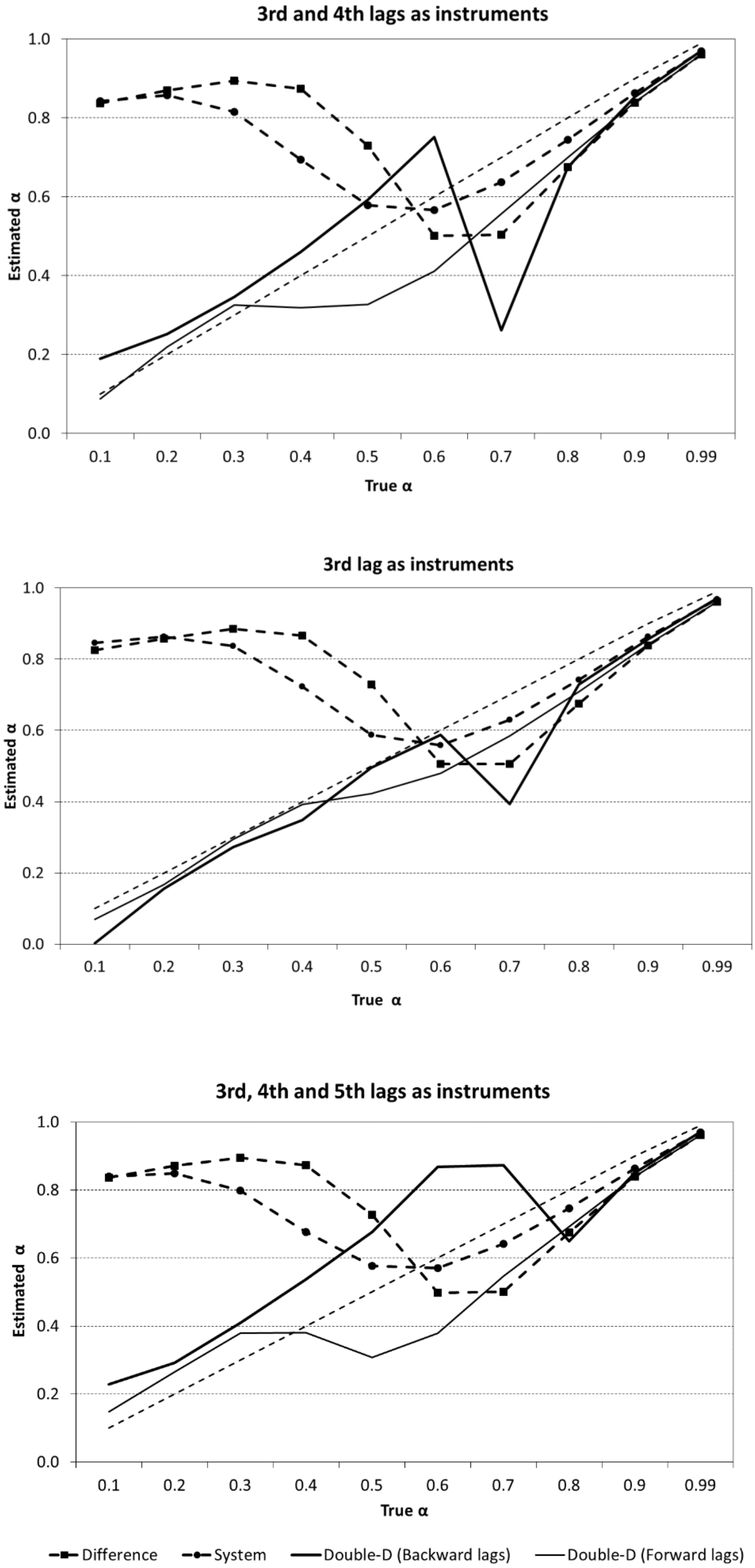

Table 1: Moment conditions used in each of the GMM estimators

	Moment conditions				
1	$E(y_{it-s}\,\Delta v_{it}) = 0; \quad (2 \leq S \leq t-1)$				
2	$E[\Delta y_{it-1}(v_{it} + \eta_i)] = 0$				
3	$E(\Delta y_{it-s}\Delta v_{it}) = 0; \quad (S \geq 2)$				
4	$E\big(\Delta y_{is}(\Delta v_{it})\big) = 0; \quad (S \geq t+2 \; and \; t \geq 2)$				
	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) [-0.190]	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) [-0.0096]	0.991 (0.008) [0.001]

Notes: Reported are the mean values from the Monte Carlo simulations of $\hat{\alpha}$ from an AR(1) process with no structural breaks. Shown in ( ) are mean standard errors and in [ ] are the mean estimated bias. The simulation is designed with T=15, N=100, a 'seed' value of 1010 and undertaken in Stata 11.1. The mean of the data is 0.087 and standard deviation is 0.163. All simulations are run with 1,000 replications. See Section 3.1 for details concerning the generation of the data.

Table 3: Monte Carlo simulation results assuming 2 common structural breaks

	Mean $\hat{\alpha}$						
True α	**Difference**	**System**	**Double-D (backward lags)**	**Double-D (forward lags)**	**Full System**	**Stacked (Difference)**	**Stacked (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]	-0.181 (0.976) [-0.281]	0.168 (0.039) [0.068]
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]	0.122 (0.035) [-0.078]	0.300 (0.034) [0.100]
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]	0.285 (0.115) [-0.015]	0.406 (0.028) [0.106]
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]	0.397 (0.082) [-0.003]	0.502 (0.023) [0.102]
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]	0.498 (0.052) [-0.002]	0.581 (0.018) [0.081]
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]	0.598 (0.036) [-0.002]	0.619 (0.015) [0.019]
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]	0.699 (0.027) [-0.001]	0.648 (0.011) [-0.052]
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]	0.799 (0.010) [-0.001]	0.740 (0.005) [-0.060]
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]	0.899 (0.008) [-0.001]	0.875 (0.001) [-0.025]
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]	0.989 (0.012) [-0.011]	0.987 (0.001) [-0.003]

Note: Reported are the mean values from the Monte Carlo simulations of $\hat{\alpha}$ from an AR(1) process with two common breaks. See Section 3.2 for details concerning the generation of the data. The common breaks are assumed to occur in periods 5 and 10. The means of the three periods are 0.100, 0.085 and 0.076 with standard deviations of 0.012, 0.010, and 0.011. The break dates and means of the data are consistent with the bank level data used in the application example in the paper. See also notes to Table 2 and Section 3.2.

Table 4: Monte Carlo simulation results assuming 2 'fuzzy' structural breaks

	Mean $\hat{\alpha}$				
True α	**Difference**	**System**	**Double-D (backward lags)**	**Double-D (forward lags)**	**Full System**
0.1	0.838 (0.052) [0.738]	0.839 (0.050) [0.739]	0.200 (0.160) [0.100]	0.157 (0.063) [0.057]	0.557 (0.033) [0.457]
0.2	0.869 (0.047) [0.669]	0.849 (0.044) [0.649]	0.313 (0.131) [0.113]	0.279 (0.074) [0.079]	0.659 (0.033) [0.459]
0.3	0.900 (0.044) [0.600]	0.801 (0.037) [0.501]	0.440 (0.109) [0.130]	0.301 (0.080) [0.001]	0.742 (0.028) [0.442]
0.4	0.883 (0.041) [0.443]	0.659 (0.026) [0.259]	0.550 (0.090) [0.150]	0.410 (0.051) [0.010]	0.777 (0.033) [0.377]
0.5	0.687 (0.033) [0.187]	0.542 (0.016) [0.042]	0.512 (0.115) [0.012]	0.552 (0.032) [0.052]	0.655 (0.029) [0.155]
0.6	0.426 (0.019) [-0.174]	0.551 (0.009) [-0.049]	0.587 (0.192) [-0.013]	0.400 (0.195) [-0.200]	0.434 (0.018) [-0.066]
0.7	0.504 (0.010) [-0.196]	0.637 (0.005) [-0.063]	0.558 (0.086) [-0.142]	0.541 (0.013) [-0.159]	0.505 (0.009) [-0.195]
0.8	0.691 (-0.004) [-0.109]	0.750 (0.003) [-0.050]	0.694 (0.019) [-0.106]	0.710 (0.006) [-0.090]	0.691 (0.002) [-0.109]
0.9	0.849 (0.002) [-0.051]	0.868 (0.001) [-0.032]	0.855 (0.004) [-0.045]	0.848 (0.002) [-0.052]	0.849 (0.002) [-0.051]
0.99	0.966 (0.001) [-0.034]	0.972 (0.000) [-0.028]	0.972 (0.001) [-0.028]	0.965 (0.001) [-0.035]	0.966 (0.001) [-0.034]

Note: Reported are the mean values from the Monte Carlo simulations of $\hat{\alpha}$ from an AR(1) process with two 'fuzzy' breaks. See Section 3.2 and 3.3 for details concerning the generation of the data and the 'fuzzy' breaks. The two 'fuzzy' break occurs with equal probability in periods 4, 5 and 6 and then again in periods 9, 10 and 11. The means and standard deviations in the three sections are the same as those used in the common break example in Table 3. See also notes to Table 2.

Table 5: Monte Carlo simulation results assuming 2 unknown structural breaks

	Mean $\hat{\alpha}$				
True α	**Difference**	**System**	**Double-D (backward lags)**	**Double-D (forward lags)**	**Full System**
0.1	0.865 (0.074) [0.765]	0.835 (0.063) [0.735]	0.210 (0.160) [0.110]	0.120 (0.069) [0.020]	0.532 (0.040) [0.432]
0.2	0.903 (0.066) [0.703]	0.824 (0.052) [0.624]	0.310 (0.181) [0.110]	0.217 (0.080) [0.017]	0.653 (0.041) [0.453]
0.3	0.934 (0.062) [0.634]	0.762 (0.062) [0.462]	0.334 (0.210) [0.034]	0.273 (0.093) [0.027]	0.755 (0.043) [0.455]
0.4	0.870 (0.055) [0.470]	0.667 (0.026) [0.267]	0.544 (0.200) [0.144]	0.305 (0.070) [-0.095]	0.779 (0.043) [0.379]
0.5	0.630 (0.037) [0.130]	0.614 (0.076) [0.114]	0.722 (0.192) [0.222]	0.400 (0.071) [-0.100]	0.634 (0.033) [0.134]
0.6	0.487 (0.021) [-0.113]	0.628 (0.070) [0.028]	0.712 (0.176) [0.112]	0.521 (0.020) [-0.079]	0.499 (0.021) [-0.101]
0.7	0.566 (0.011) [-0.136]	0.689 (0.007) [-0.011]	0.500 (0.092) [-0.200]	0.615 (0.075) [-0.085]	0.572 (0.071) [-0.228]
0.8	0.715 (0.026) [-0.085]	0.776 (0.004) [-0.024]	0.650 (0.029) [-0.150]	0.740 (0.028) [-0.260]	0.716 (0.006) [-0.084]
0.9	0.884 (0.003) [-0.016]	0.877 (0.002) [-0.023]	0.838 (0.009) [-0.062]	0.853 (0.004) [-0.047]	0.855 (0.003) [-0.045]
0.99	0.966 (0.002) [-0.024]	0.974 (0.001) [-0.016]	0.965 (0.004) [-0.035]	0.963 (0.002) [-0.027]	0.966 (0.002) [-0.034]

Note: Reported are the mean values from the Monte Carlo simulations of $\hat{\alpha}$ from an AR(1) process with two unknown breaks. See Section 3.2 for details concerning the generation of the data. The two breaks occur with equal probability between periods 2 and 14. The means and the standard deviations of the sections are the same as those used in the common break example in Table 2. See also notes to Table 2.

Table 6: Large 'T' Monte Carlo simulation assuming 2 common structural breaks

	Mean $\hat{\alpha}$						
True α	**Difference**	**System**	**Double-D (backward lags)**	**Double-D (forward lags)**	**Full System**	**Stacked (Difference)**	**Stacked (System)**
0.1	0.848 (0.029) [0.748]	0.934 (0.020) [0.834]	0.047 (0.047) [-0.053]	0.075 (0.026) [-0.025]	0.614 (0.013) [0.514]	-0.123 (0.144) [-0.223]	0.140 (0.016) [0.040]
0.2	0.856 (0.022) [0.656]	0.940 (0.170) [0.740]	0.152 (0.047) [-0.048]	0.167 (0.029) [-0.033]	0.707 (0.012) [0.507]	0.148 (0.117) [-0.052]	0.258 (0.016) [0.058]
0.3	0.872 (0.017) [0.572]	0.942 (0.014) [0.642]	0.245 (0.046) [-0.055]	0.258 (0.032) [-0.042]	0.787 (0.001) [0.487]	0.225 (0.031) [-0.075]	0.368 (0.015) [0.068]
0.4	0.887 (0.013) [0.487]	0.935 (0.011) [0.535]	0.339 (0.044) [-0.061]	0.352 (0.033) [-0.048]	0.850 (0.010) [0.450]	0.373 (0.048) [-0.017]	0.447 (0.012) [0.047]
0.5	0.893 (0.010) [0.393]	0.920 (0.009) [0.420]	0.430 (0.040) [-0.070]	0.466 (0.028) [-0.034]	0.888 (0.008) [0.388]	0.491 (0.026) [-0.008]	0.588 (0.013) [-0.001]
0.6	0.881 (0.007) [0.281]	0.914 (0.009) [0.314]	0.557 (0.031) [-0.043]	0.584 (0.017) [-0.016]	0.889 (0.007) [0.289]	0.590 (0.014) [-0.010]	0.631 (0.007) [0.031]
0.7	0.848 (0.004) [0.148]	0.840 (0.003) [0.140]	0.691 (0.016) [-0.009]	0.694 (0.009) [-0.006]	0.851 (0.004) [0.151]	0.699 (0.007) [-0.001]	0.706 (0.004) [0.006]
0.8	0.822 (0.002) [0.022]	0.841 (0.002) [0.41]	0.780 (0.005) [-0.020]	0.798 (0.004) [-0.002]	0.816 (0.002) [0.016]	0.797 (0.004) [-0.003]	0.791 0.002 [-0.009]
0.9	0.866 (0.001) [-0.034]	0.891 (0.000) [-0.009]	0.867 (0.001) [-0.033]	0.897 (0.001) [-0.001]	0.864 (0.000) [-0.034]	0.899 (0.001) [0.001]	0.884 (0.000) [-0.006]
0.99	0.981 (0.000) [-0.009]	0.983 (0.000) [-0.007]	0.986 (0.000) [-0.004]	0.981 (0.002) [-0.009]	0.981 (0.001) [-0.009]	0.988 (0.001) [-0.002]	0.988 (0.002) [-0.002]

Notes: The sample was increased from T=15 to T=60. The breaks occur in the same relative position as in Table 3 in periods 20 and 40. See also notes to Table 3.

Table 7: United States Estimates of the Bank Lending Chanel

	1	2	3	4	5
Variables	**Difference**	**System**	**Double-D (backward lags)**	**Double-D (forward lags)**	**Full System**
Δl_{it-1}	-0.469**	-0.409**	0.126**	0.097**	-0.397**
	(0.000)	(0.000)	(0.057)	(0.097)	(0.611)
Long-Run Coefficients					
Δgdp_{t-1}	1.144**	1.004**	0.175*	0.180*	0.992**
	(0.000)	(0.738)	(0.213)	(0.179)	(0.000)
inf_{t-j}	0.003*	0.002*	0.002**	0.006**	0.709**
	(0.001)	(0.000)	(0.001)	(0.001)	(0.000)
R_{t-j}	-0.007**	-0.004**	-0.018**	-0.0176**	-0.005*
	(0.000)	(0.063)	(0.000)	(-0.017)	(0.000)
$\mathrm{SIZE}_{it-1}\Delta \mathrm{R}_{t-j}$	0.007*	0.020**	0.051**	0.044**	0.020**
	(0.004)	(0.003)	(0.008)	(0.000)	(0.004)
$LIQ_{it-1}\Delta R_{t-j}$	0.004	0.016**	0.001	0.001	0.037
	(0.004)	(0.004)	(0.005)	(0.005)	(0.004)
$CAP_{it-1}\Delta R_{t-j}$	0.006	0.043	0.013*	0.172**	0.015**
	(0.074)	(0.063)	(0.075)	(0.086)	(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 $\boldsymbol{\Delta l_{it}}$. 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. The models are estimated using Stata 11.