



# Multiple Structural Breaks in Interactive Effects Panel Data Models

Jan Ditzen<sup>1</sup> | Yiannis Karavias<sup>2</sup> | Joakim Westerlund<sup>3,4</sup>

<sup>1</sup>Faculty of Economics and Management, Free University of Bozen-Bolzano, Bozen, Italy | <sup>2</sup>Department of Economics and Finance, Brunel University London, London, UK | <sup>3</sup>Department of Economics, Lund University, Lund, Sweden | <sup>4</sup>Department of Economics, Deakin University, Melbourne, Victoria, Australia

Correspondence: Yiannis Karavias (yiannis.karavias@brunel.ac.uk)

Received: 18 May 2023 | Revised: 13 June 2024 | Accepted: 3 September 2024

Keywords: bank lending | cross-section dependence | interactive fixed effects | panel data | quantitative easing | structural breaks

## ABSTRACT

This paper develops new econometric methods for multiple structural break detection in panel data models with interactive fixed effects. The new methods include tests for the presence of structural breaks, estimators for the number of breaks and their location, and a method for constructing asymptotically valid break date confidence intervals. The new methodology is applied to a large panel of US banks for a period characterized by massive quantitative easing programs aimed at lessening the impact of the global financial crisis and the COVID-19 pandemic. The question we ask is as follows: Have these programs been successful in spurring bank lending in the US economy? The short answer turns out to be: "No".

# 1 Introduction

Accounting for structural change has always been an important issue in economics and elsewhere. However, because of major events such as the 2007–2008 global financial crisis, the 2007–2010 subprime mortgage crisis, the 2016 Brexit referendum, the 2020 COVID-19 outbreak, and the 2022 war in Ukraine, interest has recently intensified.

The time series literature concerned with the estimation and testing for breaks is huge, and there is by now considerable accumulated empirical evidence of breaks in all kinds of economic relationships, especially in macroeconomics and finance (see, e.g., Stock and Watson 1996, 2003). Most of this evidence is based on econometric techniques requiring that there is (at most) one break (see, e.g., Andrews 1993). However, there are also techniques that allow for an unknown number of breaks, which is the most relevant scenario in practice. An important contribution in this area is Bai and Perron (1998), "BP98" henceforth, who develop a methods for testing and dating multiple

breaks in linear time series regression models. The methodology includes (i) a number of tests for the presence of breaks, including a sequential test procedure to estimate the number of breaks, (ii) a breakpoint estimator, and (iii) a breakpoint confidence interval. The BP98 methodology is widely applicable, it is computationally attractive, and it is readily available in many software programs, such as GAUSS, EViews, MATLAB, R and most recently Stata (see Ditzen, Karavias, and Westerlund 2022). BP98 has been instrumental not only for empirical economists but also for econometricians who have used it as a launching pad for numerous extensions (see the 2021 Journal of Econometrics special issue in honor of Pierre Perron for overviews of some of these works).

Panel data relationships are also susceptible to breaks, a fact that is by now well-understood in the literature, and it is not difficult to find empirical evidence in its support (see Antoch et al. 2019; Boldea, Drepper, and Gan 2020; Kaddoura and Westerlund 2022; Karavias, Narayan, and Westerlund 2023; Zhu, Sarafidis, and Silvapulle 2020, to mention a few). Yet, it is not until recently that

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

<sup>© 2024</sup> The Author(s). Journal of Applied Econometrics published by John Wiley & Sons Ltd.

methods designed to estimate and test for breaks in such relationships have become available to practitioners. Most of these methods are constrained to single break environments (see, e.g., Antoch et al. 2019; Baltagi, Feng, and Kao 2016; Baltagi, Kao, and Liu 2017; Karavias, Naravan, and Westerlund 2023; Zhu, Sarafidis, and Silvapulle 2020). Boldea, Drepper, and Gan (2020), Kaddoura and Westerlund (2022), and Oian and Su (2016) allow for multiple breaks and are therefore more general in this regard. However, this generality comes at a cost in terms of the allowable unobserved heterogeneity, which is important as unattended heterogeneity can be mistaken for breaks (and vice versa). In particular, while Qian and Su (2016) allow for individual fixed effects, Boldea, Drepper, and Gan (2020) and Kaddoura and Westerlund (2022) assume that the unobserved heterogeneity is made up of a special type of interactive random effects that basically can be ignored in the estimation. Li, Qian, and Su (2016) allow for more general interactive effects and multiple breaks that are dealt with using a version of the (group fused) least absolute shrinkage and selection operator (LASSO) method. As far as we are aware, this is the most general approach available at the moment. The use of the LASSO does, however, make for a rather complicated estimation problem that is not only nonlinear but that also involves a number of tuning parameters. In particular, researchers need to specify a penalty tuning parameter, an adaptive weight, a shrinking threshold, and the number of factors, and performance can be very sensitive to these choices. The LASSO of Li et al. is also not equipped to deal with models in which only a subset of coefficients are breaking, and it does not allow for the construction of breakpoint confidence intervals. The same is true for the LASSO considered by Kaddoura and Westerlund (2022). The paper that is closest to ours in terms of the techniques used is that of Karavias, Narayan, and Westerlund (2023). As already pointed out, however, they only allow for one break, and the extension to multiple breaks is nontrivial.

In the present paper, we take the above observations as a source of inspiration for developing a fully fledged panel extension of BP98's time series methodology. The model that we consider for this purpose, laid out in Section 2, is very general and allows for not only multiple structural breaks but also interactive fixed effects. As we explain in detail in Section 3, in contrast to Li, Qian, and Su (2016), the estimation of the interactive fixed effects, the breaks, and the model parameters is not carried out jointly but sequentially in two steps, which enables us to separate the breaks from the effects. We begin by estimating and removing the interactive effects, which is done using a version of the computationally convenient common correlated effects (CCE) approach of Pesaran (2006). With the interactive fixed effects gone, in the second step, we apply a panel version of BP98's methodology to the resulting cleaned up regression. The validity of the new methods is established under the condition that both the number cross-sectional units, N, and the number time periods, T, tend to infinity such that T/N tends to zero. According to our Monte Carlo study reported in the supporting information appendix, however, the proposed methods work well even in small samples where this condition is not met.

While the US economy has been struck by several major events in the past 15 years, the 2007–2008 global financial crisis and the 2020 COVID-19 outbreak have been particularly disruptive. In both cases, the Federal Reserve's policy response was the use of a series of large-scale asset purchase programs, or "quantitative easing" (QE) rounds, whose usefulness to this date is debated. In Section 4, we evaluate whether these rounds achieved their primary goal of spurring credit flow in the economy. According to the Lucas critique, effective QE policies should cause breaks in banks' lending behavior. These can be detected using the new methodology. The sample that we use covers 3557 banks across 64 quarters, from 2005Q3 to 2021Q3. According to the results, bank lending has suffered from a number of breaks that can be attributed to the Federal Reserve's QE interventions. However, a majority of these did not have the intended expansionary effect, but rather banks took the opportunity to build up reserves in order to meet their capitalisation requirements. The main exception occurs towards the end of the sample, a period that includes the COVID-19 outbreak. At this point in time, the health of the banking sector had improved substantially, and the subsequent interventions were large enough for banks to be able to expand lending while at the same time cover their capitalisation needs. The net effect over the whole sample period is that lending has increased, but just barely.

Section 5 concludes. All proofs and theoretical results of secondary nature are provided in the supporting information appendix, which also contains our Monte Carlo study. In a companion paper, Ditzen, Karavias, and Westerlund (2022) present a new Stata command, called xtbreak, for implementing the new methodology.

## 2 | Model and Assumptions

## 2.1 Model

Consider the following panel data model with k structural breaks:

$$y_{i,l} = x'_{i,l}\beta + w'_{i,t}\delta_1 + e_{i,l} \text{ for } t = 1, ..., T_1,$$
  

$$y_{i,l} = x'_{i,l}\beta + w'_{i,t}\delta_2 + e_{i,l} \text{ for } t = T_1 + 1, ..., T_2,$$
  

$$\vdots$$
  

$$y_{i,l} = x'_{i,l}\beta + w'_{i,t}\delta_{k+1} + e_{i,t} \text{ for } t = T_k + 1, ..., T,$$
(2.1)

where  $T_1, \ldots, T_k$  are the dates of the structural breaks. We also define  $T_0 = 0$  and  $T_{k+1} = T$ . The dependent variable  $y_{i,t}$  and the error  $e_{i,t}$  are scalars, the regressors  $x_{i,t}$  and  $w_{i,t}$  are  $p_x \times 1$  vectors and  $p_w \times 1$  vectors, respectively, and  $\beta$  and  $\delta_1, \ldots, \delta_{k+1}$  are conformable vectors of coefficients. The coefficients of  $x_{i,t}$  are unaffected by the breaks, while those of  $w_{i,t}$  are affected by the breaks. It is possible that all coefficients break, in which case we define  $x'_{i,t}\beta = 0$ . It is also possible that different coefficients break at different times by allowing subsets of  $\delta_i$  to remain constant across regimes. We follow the bulk of the previous literature and assume that  $\beta$  and  $\delta_1, \ldots, \delta_{k+1}$  are equal across the cross-section. One way to relax this assumption is to follow Pesaran (2006) and to assume that the individual coefficients are randomly distributed. This will not affect the validity of our methodology, which can then be considered as estimating the mean of those coefficients. We also assume that the breaks are discrete and that their timing is the same for all cross-sectional units, which is again standard and reasonable in low-frequency panel data sets.

In terms of the empirical study of Section 4,  $y_{i,t}$  is one of three measures of lending by bank *i* in quarter *t*. The regressors in  $w_{i,t}$ , whose coefficients are allowed to be breaking, are bank holdings in securities believed to be affected by QE. The regressors included in  $x_{i,t}$  are control variables like returns on assets (ROA), total assets, equity, and cost of deposits, among others.

Unobserved heterogeneity in  $e_{i,t}$  is allowed through the following interactive fixed effects specification:

$$e_{i,t} = f'_t \gamma_i + \varepsilon_{i,t}, \qquad (2.2)$$

where  $f_t$  is a  $m \times 1$  vector of unobserved common factors with  $\gamma_i$  being the associated vector of factor loadings and  $\varepsilon_{i,t}$  is an idiosyncratic error. The interactive effects are here given by  $f'_t \gamma_i$ .<sup>1</sup> Because they are common to all cross-sectional units, the factors are a source of strong cross-section dependence.

The model in (2.4) does not allow the breaks to affect the loadings in  $\gamma_i$ . This condition is not necessary. Breaks in  $\gamma_i$  that take place at the same dates as those in  $\delta_j$  can be permitted without affecting any of our results. Breaks that only affect  $\gamma_i$  or that take place at different dates than those in  $\delta_j$  can also be accommodated without change but that will cause the number of unknown factors to increase. In the empirical study of Section 4, we explain how to accommodate observed common factors with breaking loadings.

In many empirical applications, it is likely that  $f_t$  is correlated with the regressors. In order to capture this, we assume that

$$x_{i,t} = \Gamma'_{x,i} f_t + u_{x,i,t},$$
 (2.3)

$$w_{i,t} = \Gamma'_{w,i} f_t + u_{w,i,t}, \qquad (2.4)$$

where the factor loading matrices  $\Gamma_{x,i}$  and  $\Gamma_{w,i}$  are  $m \times p_x$  and  $m \times p_w$ , respectively, while the idiosyncratic errors  $u_{x,i,t}$  and  $u_{w,i,t}$  are  $p_x \times 1$  and  $p_w \times 1$ , respectively. The presence of  $f_t$  here is reasonable because regressors are often co-moving, both among themselves and across the cross-section.

A prominent example of an unobserved factor that affects both bank lending and the right-hand side control variables is the stimulus provided by the Federal Reserve through conventional monetary policy, as measured by the difference between the federal funds rate and the optimal interest rate predicted by models such as the Taylor rule. This variable is unobserved and is therefore typically not included. Other factors include concurrent policies such as the Troubled Asset Relief Program, house price growth, the average maturity structure of the Federal Reserve's holdings of Treasury securities, and aggregate demand and supply shocks from current and anticipated macroeconomic conditions. It is natural to assume that banks respond heterogeneously to each of these factors, which in (2.2)-(2.4) is captured by the unit-specific factor loadings  $\gamma_i$ ,  $\Gamma_{w,i}$  and  $\Gamma_{x,i}$ .

The formal assumptions that we will be working under are laid out in Section 2.2. However, before we take the assumptions, it is useful to first stack (2.1)-(2.4) over time and to also introduce the CCE estimator that we will be using. Let us therefore denote by diag(A, B) the block-diagonal matrix that takes the matrices A and B as the upper left and lower right block, respectively. Let

 $\mathcal{T}_k = \{T_1, \dots, T_k\}$  be the set of breakpoints. We can now define the following matrices:

$$W_i(\mathcal{T}_k) = \operatorname{diag}(W_{1,i}, \dots, W_{k+1,i}),$$
$$U_{w,i}(\mathcal{T}_k) = \operatorname{diag}(U_{w,1,i}, \dots, U_{w,k+1,i}),$$
$$F(\mathcal{T}_k) = \operatorname{diag}(F_1, \dots, F_{k+1}),$$

where  $W_{j,i} = (w_{i,T_{j-1}+1}, \ldots, w_{i,T_j})'$  is  $(T_j - T_{j-1}) \times p_w$ ,  $U_{w,j,i} = (u_{w,i,T_{j-1}+1}, \ldots, u_{w,i,T_j})'$  is  $(T_j - T_{j-1}) \times p_w$ , and  $F_j = (f_{T_{j-1}+1}, \ldots, f_{T_j})'$  is  $(T_j - T_{j-1}) \times m$  for  $j = 1, \ldots, k + 1$ . Hence,  $W_{j,i}, U_{w,j,i}$ , and  $F_j$  stacks the time observations within regime j. The matrices  $W_i(\mathcal{T}_k)$ ,  $U_{w,i}(\mathcal{T}_k)$ , and  $F(\mathcal{T}_k)$  stack these regime-specific stacks on the diagonal, which means that their dimensions must be  $T \times (k+1)p_w$ ,  $T \times (k+1)p_w$ , and  $T \times (k+1)m$ , respectively. We also introduce the  $T \times 1$  vectors  $Y_i = (y_{i,1}, \ldots, y_{i,T})'$  and  $e_i = (e_{i,1}, \ldots, e_{i,T})'$ , the  $T \times p_x$  matrix  $X_i = \left(x'_{i,1}, \ldots, x'_{i,T}\right)'$ , the  $(k+1)p_w \times 1$  vector  $\delta = \left(\delta'_1, \ldots, \delta'_{k+1}\right)'$ , the  $T \times m$  matrix  $F = (f_1, \ldots, f_T)'$ , and the  $T \times p_x$  matrix  $U_{x,i} = (u_{x,i,1}, \ldots, u_{x,i,T})'$ . In this notation, the time-stacked version of (2.1)-(2.4) becomes

$$Y_i = X_i \beta + W_i(\mathcal{T}_k)\delta + e_i, \qquad (2.5)$$

$$e_i = F\gamma_i + \varepsilon_i, \qquad (2.6)$$

$$X_i = F\Gamma_{x,i} + U_{x,i}, \qquad (2.7)$$

$$W_{i}(\mathcal{T}_{k}) = F(\mathcal{T}_{k})(I_{k+1} \otimes \Gamma_{w,i}) + U_{w,i}(\mathcal{T}_{k}).$$
(2.8)

The fact that *F* enters (2.6)–(2.8) means that (2.5) cannot be consistently estimated by ordinary least squares (OLS) even if  $\mathcal{T}_k$  is known, because the  $X_i$  and  $W_i(\mathcal{T}_k)$  are correlated with  $e_i$  through *F*. Consistent estimation of  $\beta$  and  $\delta$  is therefore not possible without properly controlling for *F*. While there are other approaches that can be used for this purpose, here, we opt for a version of the CCE approach of Pesaran (2006), which is simple, robust, and with excellent small-sample properties (see, e.g., Westerlund and Urbain 2015). In particular, since in a model with breaks the cross-sectional average of the dependent variable is uninformative in this regard, unlike in standard CCE, the estimator that we use here employs the cross-sectional averages of the regressors only to estimate (space spanned by) *F*. Let us collect these regressors in the  $T \times (p_x + (k + 1)p_w)$  matrix  $Z_i(\mathcal{T}_k) = (X_i, W_i(\mathcal{T}_k))$ . Define the following  $T \times T$  the projection error matrix:

$$M_{\overline{Z}(\mathcal{T}_k)} = I_T - \overline{Z}(\mathcal{T}_k)(\overline{Z}(\mathcal{T}_k)'\overline{Z}(\mathcal{T}_k))^{-1}\overline{Z}(\mathcal{T}_k)', \qquad (2.9)$$

where  $\bar{A} = N^{-1} \sum_{i=1}^{N} A_i$  is the cross-sectional average of any generic variable  $A_i$ . In the supporting information appendix, we show that premultiplication by this matrix eliminates the factors asymptotically. Let us therefore denote by  $\tilde{A} = \tilde{A}(\mathcal{T}_k) = M_{\overline{Z}(\mathcal{T}_k)}A$  the "defactored" version of any *T*-rowed matrix A.<sup>2</sup> In this notation, the model to be estimated is given by

$$\tilde{Y}_{i} = \tilde{X}_{i}\beta + \tilde{W}_{i}(\mathcal{T}_{k})\delta + \tilde{e}_{i}.$$
(2.10)

This model can be stacked over the cross-section, giving

$$\tilde{Y} = \tilde{X}\beta + \tilde{W}(\mathcal{T}_k)\delta + \tilde{E}, \qquad (2.11)$$

where  $\tilde{Y} = \left(\tilde{Y}'_1, \ldots, \tilde{Y}'_N\right)'$ ,  $\tilde{X} = \left(\tilde{X}'_1, \ldots, \tilde{X}'_N\right)'$ ,  $\tilde{W}(\mathcal{T}_k) = \left(\tilde{W}_1(\mathcal{T}_k)', \ldots, \tilde{W}_N(\mathcal{T}_k)'\right)'$ , and  $\tilde{E} = \left(\tilde{e}'_1, \ldots, \tilde{e}'_N\right)'$  are all *NT*-rowed. The CCE estimator of  $\delta$  that we will be considering in this paper is simply the pooled OLS estimator obtained from (2.11);

$$\hat{\delta}(\mathcal{T}_k) = (\tilde{W}(\mathcal{T}_k)' M_{\tilde{X}} \tilde{W}(\mathcal{T}_k))^{-1} \tilde{W}(\mathcal{T}_k)' M_{\tilde{X}} \tilde{Y}, \qquad (2.12)$$

where  $M_{\tilde{X}} = I_{NT} - \tilde{X}(\tilde{X}'\tilde{X})^{-1}\tilde{X}'$ .

#### 2.2 Assumptions

We begin this section with some additional notation. If *A* is a matrix, tr(A) and rank(A) signify its trace and rank, respectively, and  $||A|| = \sqrt{tr(A'A)}$  signifies its Frobenius norm. We write A > 0 to signify that *A* is positive definite. The symbols  $\rightarrow_d$ ,  $\rightarrow_p$  and  $\rightarrow_w$  signify convergence in distribution, convergence in probability, and weak convergence, respectively, and  $\lfloor x \rfloor$  signifies the integer part of *x*. We use *N*,  $T \rightarrow \infty$  to indicate that the limit has been taken while passing both *N* and *T* to infinity. We use w.p.a.1 to denote with probability approaching one. Here and throughout the remainder of this paper,  $\mathcal{T}_{k^0}^0 = \{T_1^0, \ldots, T_{k^0}^0\}$  will be used to denote the set of true breakpoints with  $k^0$  being the true number of breaks. If  $k^0 = 0$ , we define  $\mathcal{T}_{k^0}^0 = \emptyset$ .

## Assumption 2.1. (Breaks).

- (i)  $T_j^0 = \lfloor \lambda_j^0 T \rfloor$  for  $j = 1, ..., k^0 + 1$ , where  $\lambda_0^0 = 0 < \lambda_1^0 < ... < \lambda_k^0 < \lambda_{k^0+1}^0 = 1$ .
- (ii)  $0 < ||\delta|| < \infty$  and  $\delta_{i+1} \neq \delta_i$  whenever  $k^0 > 0$ .

## Assumption 2.2. (Errors).

- (i)  $u_{i,t} = \left(u'_{x,i,t}, u'_{w,i,t}\right)'$  is a covariance stationary process that is independent across *i* with absolutely summable autocovariances,  $E(u_{i,t}) = 0_{(p_x+p_w)\times 1}$ ,  $E(u_{i,t}u'_{i,t}) = \Sigma_{u,i}$ , and  $E(||u_{i,t}||^4) < \infty$ .
- (ii)  $\varepsilon_{i,t}$  is a covariance stationary process that is independent across *i* with absolutely summable autocovariances,  $E(\varepsilon_{i,t}) = 0$ ,  $E\left(\varepsilon_{i,t}^2\right) = \sigma_{\varepsilon,i}^2$ ,  $E\left(\varepsilon_i\varepsilon_i'\right) = \Sigma_{\varepsilon,i}$  and  $E\left(\varepsilon_{i,t}^4\right) < \infty$ .
- (iii)  $\varepsilon_{i,t}$  and  $u_{j,s}$  are independent for all i, j, s and t.

#### Assumption 2.3. (Factors).

- (i)  $T^{-1}F'F > 0$  w.p.a.1 for all *T*.
- (ii)  $E(||f_t||^2) < \infty$  for all t.
- (iii)  $f_t$  is independent of  $\varepsilon_{i,s}$  and  $u_{i,s}$  for all *i*, *s* and *t*.

#### Assumption 2.4. (Loadings).

- (i)  $\operatorname{rank}(\overline{\Gamma}_x) = m \le p_x$  and  $\operatorname{rank}(\overline{\Gamma}_w) = m \le p_w$  for all *N*.
- (ii)  $\gamma_i$ ,  $\Gamma_{w,i}$  and  $\Gamma_{x,i}$  are nonrandom such that  $\|\gamma_i\| < \infty$ ,  $\|\Gamma_{w,i}\| < \infty$  and  $\|\Gamma_{x,i}\| < \infty$ .

**Assumption 2.5.** (Invertibility).  $(NT)^{-1}\tilde{X}'\tilde{X} > 0$  and  $(NT)^{-1}\tilde{W}(\mathcal{T}_k)'M_{\tilde{X}}\tilde{W}(\mathcal{T}_k) > 0$  w.p.a.1 for all N and T.

**Assumption 2.6.** (Moments). The following holds as  $N, T \rightarrow \infty$ :

$$\begin{array}{l} \text{(i)} \quad \frac{1}{N\Delta T_{j}^{0}} \sum_{i=1}^{N} \sum_{t=T_{j-1}^{0}+1}^{T_{j}^{0}} \breve{u}_{w,i,t} \breve{u}_{w,i,t}' \to_{p} \Omega_{j}, \\ \text{(ii)} \quad \frac{1}{N\Delta T_{j}^{0}} \sum_{i=1}^{N} \sum_{t=T_{j-1}^{0}+1}^{T_{j}^{0}} \sum_{t=T_{j-1}^{0}+1}^{T_{j}^{0}} \varepsilon_{i,t} \varepsilon_{i,l} \breve{u}_{w,i,t} \breve{u}_{w,i,l}' \to_{p} \Phi_{j}, \\ \text{(iii)} \quad \frac{1}{\sqrt{N\Delta T_{j}^{0}}} \sum_{i=1}^{N} \sum_{t=T_{j-1}^{0}+1}^{T_{j-1}^{0}+1} \varepsilon_{i,t} \breve{u}_{w,i,t} \to_{w} \Phi_{j}^{1/2} B_{j}(s), \end{array}$$

where  $j = 1, ..., k^0 + 1$ ,  $\Delta T_j^0 = T_j^0 - T_{j-1}^0$ ,  $s \in [0,1]$ ,  $B_i(s)$  is an  $p_w \times 1$  vector standard Brownian motion on the same interval, and  $\check{u}_{w,i,t} = u_{w,i,t} - \sum_{n=1}^N \sum_{l=T_{j-1}^0+1}^{T_j^0} u_{w,n,l} a_{i,n,t,l}$  with  $a_{i,n,t,l} = u'_{x,n,l} \left( \sum_{i=1}^N U'_{x,i} U_{x,i} \right)^{-1} u_{x,i,t}$ .

Assumption 2.1 requires that the breaks are distinct and hence that each regime increases with T. This is standard in the type of large-T panel data that we are considering (see, e.g., Baltagi, Feng, and Kao 2016). In pure time series, it is even necessary, as enough observations are needed to consistently estimate the slope coefficients in each regime. The panel data structure brings more (cross-sectional) information to the table, and this enables consistent estimation of the breakpoints even if the regimes are not expanding (as we show in the supporting information appendix). Testing for breaks is, however, more demanding in this regard and is not possible without Assumption 2.1.

Assumptions 2.2-2.4 are standard in the CCE literature (see Pesaran 2006) and are not particularly restrictive. For example, while Assumption 2.2 (iii) do rule out lagged dependent variables in  $x_{i,t}$  and  $w_{i,t}$ ,  $y_{i,t}$  is still allowed to be serially correlated through both  $f_t$  and  $\varepsilon_{i,t}$ . This assumption also does not rule out endogeneity, as  $x_{i,t}$  and  $w_{i,t}$  can be correlated with  $e_{i,t}$  through  $f_t$ . Assumption 2.3 holds if  $f_t$  is stationary, which can be restrictive. Assumption 2.3 is not necessary, though and can be relaxed to allow for generally trending factors without changes to the methodology by the same arguments as in Westerlund (2018). Assumption 2.4 (i), which dictates that the number of factors cannot be larger than the number of averages employed by CCE, is also stronger than necessary. It can be relaxed if some of the factors are observed. Observed factors can be appended to  $\overline{Z}(\mathcal{T}_{\mu})$ and projected out with the effect that these factors do not have to satisfy Assumption 2.4 (i). In Section 4, we elaborate on this point.

Assumption 2.5 is a noncollinearity condition. It demands that the regressors have enough variation across both the cross-section and time after projecting out all variation that can be explained by the factors, which is a standard requirement in the interactive effects literature (see Pesaran 2006).

Assumption 2.6 requires that the moments of  $u_{w,i,t}$  after projecting out the part of the variation that is due to  $u_{x,i,t}$  are constant within regimes. This is needed for our asymptotic distribution theory but not for the consistency of the estimated breakpoints.

(3.11)

# 3 | The New Methods and their Asymptotic Properties

## 3.1 | Testing for Structural Breaks

This section presents tests for three hypotheses, labelled "(A)"–"(C)", which are useful for establishing whether or not there are any breaks and for determining their number. Some of these are stated in terms of the set of permissible break dates, in which all breaks are distinct and bounded away from the sample endpoints. This set is given by

$$\mathcal{T}_{k,\epsilon} = \{(T_1, \ \dots, \ T_k) : \Delta T_{j+1} \ge \epsilon T, T_1 \ge \epsilon T, T_k \le (1-\epsilon)T\},$$

$$(3.1)$$

where  $\Delta T_j = T_j - T_{j-1}$  and  $\epsilon > 0$  is a user-defined trimming parameter, the choice of which will be discussed later.

## Hypotheses:

(A) The null hypothesis is that there are no breaks, and the alternative hypothesis is that there are at most k structural breaks, which may or may not be equal to  $k^0$ , and where the dates of the breaks can be either known or unknown. If the breakpoints are known, the null and alternative hypotheses can be stated in the following way:

$$H_0^A : \delta_1 = \ldots = \delta_{k+1}, \tag{3.2}$$

$$H_1^A: \delta_n \neq \delta_j \text{ for some } n \neq j \in \{1, \dots, k+1\}.$$
(3.3)

If, on the other hand, the breakpoints are unknown, the null hypothesis is the same but the alternative changes to

$$H_1^A : \bigcup_{\mathcal{T}_k \in \mathcal{T}_{k,\epsilon}} \{\delta_n \neq \delta_j \text{ for some } n \neq j\}.$$
(3.4)

(B) The null hypothesis is that there are no breaks, and the alternative is that there is an unknown number of breaks, where the number of breaks is bounded from above by some prescribed value  $k_{max}$ . Formally,

$$H_0^B: \bigcup_{k=2}^{k_{max}+1} \{\delta_1 = \dots = \delta_k\},$$
 (3.5)

$$H_1^B : \bigcup_{k=1}^{k_{max}} \left\{ \bigcup_{\mathcal{T}_k \in \mathcal{T}_{k,\varepsilon}} \delta_n \neq \delta_j \text{ for some } n \neq j \right\}.$$
(3.6)

(C) The null hypothesis is that there are k breaks and the alternative is that there are k + 1 breaks. The value of k is specified by the researcher. Formally,

$$H_0^C : \bigcup_{i=1}^{k+1} \left\{ \delta_i = \delta_{i+1} \text{ and } \delta_j \neq \delta_k \text{ for any } (j,k) \neq (i,i+1) \right\},$$
(3.7)

$$H_1^C: \bigcap_{i,j \in \{1,\dots,k+1\} \atop i \neq j} \left\{ \delta_i \neq \delta_j \right\}.$$
(3.8)

We will consider four test statistics, two for (A) and one for each of (B) and (C). These can be seen as panel extensions of the time series statistics developed by BP98.

To test  $H_0^A$  versus  $H_1^A$  when the dates of the breaks are known, define the  $kp_w \times (k+1)p_w$  matrix  $R = I_k \otimes (I_{p_w}, -I_{p_w})$ , which is such that  $(R\delta)' = (\delta'_1 - \delta'_2, \dots, \delta'_k - \delta'_{k+1})$ . The appropriate *F*-statistic to use for testing this hypothesis is given by  $F(\mathcal{F}_{k^0}^0)$ , where

$$F(\mathcal{T}_k) = \frac{N(T - p_x - (k+1)p_w) - p_x - (k+1)p_w}{kp_w}$$

$$\hat{\delta}(\mathcal{T}_k)' R' (R\hat{V}_{\hat{\delta}}R')^{-1}R\hat{\delta}(\mathcal{T}_k).$$
(3.9)

The degree of freedom correction used in the normalization of  $\hat{\delta}(\mathcal{T}_k)' R' (R\hat{V}_{\hat{\delta}}R')^{-1}R\hat{\delta}(\mathcal{T}_k)$  is not necessary but we keep it since it leads to slightly better small sample performance than if  $NT/(kp_w)$  is used. The  $(k+1)p_w \times (k+1)p_w$  matrix  $\hat{V}_{\hat{\delta}}$  is an estimator of the asymptotic covariance matrix of  $\hat{\delta}(\mathcal{T}_k)$  (with the dependence on  $\mathcal{T}_k$  suppressed), which is given by

$$\hat{V}_{\hat{\delta}} = \hat{\Omega}^{-1} \hat{\Phi} \hat{\Omega}^{-1}.$$
 (3.10)

Here

 $\hat{\Omega}$ 

$$= (NT)^{-1} \tilde{W}(\mathcal{T}_k)' M_{\tilde{X}} \tilde{W}(\mathcal{T}_k),$$

$$\hat{\Phi} = \hat{\Lambda}_0 + \sum_{l=1}^{L} \left( 1 - \frac{l}{L+1} \right) (\hat{\Lambda}_l + \hat{\Lambda}_l'), \qquad (3.12)$$

$$\hat{\Lambda}_{l} = \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=l+1}^{T} \hat{\varepsilon}_{i,t} \hat{\varepsilon}_{i,t-l} \breve{w}_{i,t} \breve{w}'_{i,t-l}, \qquad (3.13)$$

where *L* is a user-specified bandwidth. The  $1 \times p_w$  vector  $\check{w}'_{i,t}$  is the *t*-th row of the  $T \times p_w$  matrix  $\check{W}_i = (\check{w}_{i,1}, \ldots, \check{w}_{i,T})'$ , which in turn is the *i*-th block of the  $NT \times p_w$  matrix  $M_{\tilde{X}} \bar{W}(\mathcal{F}_k) =$  $(\check{W}'_1, \ldots, \check{W}'_N)'$ . The scalar  $\hat{\varepsilon}_{i,t}$  is the *t*-th row of the  $T \times 1$  vector  $\hat{\varepsilon}_i = (\hat{\varepsilon}_{i,1}, \ldots, \hat{\varepsilon}_{i,T})'$ , which is the *i*-th block of the  $NT \times 1$  vector  $\hat{\varepsilon} = (\hat{\varepsilon}'_1, \ldots, \hat{\varepsilon}'_N)' = M_{\tilde{X}}(\tilde{Y} - \tilde{W}(\mathcal{F}_k)\hat{\delta}(\mathcal{F}_k))$ .

The following theorem provides the asymptotic distribution of  $F(\mathcal{T}_k^0)$ . The theorem is stated in terms of the following process, which is a multiple break generalization of the squared tied-down Bessel process that is otherwise so common in the break testing literature (see, e.g., Andrews 1993):

$$Q(\Lambda_k) = \frac{1}{kp_w} \sum_{j=1}^k \frac{[\lambda_j B(\lambda_{j+1}) - \lambda_{j+1} B(\lambda_j)]' [\lambda_j B(\lambda_{j+1}) - \lambda_{j+1} B(\lambda_j)]}{\lambda_j \lambda_{j+1} (\lambda_{j+1} - \lambda_j)},$$
(3.14)

with  $B(\lambda_j)$  being an  $p_w \times 1$  vector standard Brownian motion on [0,1] and  $\Lambda_k = \{\lambda_1, \dots, \lambda_k\}$ . The true value of  $\Lambda_k$  is denoted  $\Lambda_{k^0}^0 = \{\lambda_1^0, \dots, \lambda_{k^0}^0\}$ . We also define the supremum of the above process:

$$\sup Q(k) = \sup_{\Lambda_k \in \Lambda_{k,\epsilon}} Q(\Lambda_k), \qquad (3.15)$$

where

$$\Lambda_{k,\epsilon} = \{ (\lambda_1, \dots, \lambda_k) : \Delta \lambda_{j+1} \ge \epsilon, \lambda_1 \ge \epsilon, \lambda_k \le 1 - \epsilon \}$$
(3.16)

with  $\Delta \lambda_j = \lambda_j - \lambda_{j-1}$ . The dependence of  $Q(\Lambda_k)$  and  $\sup Q(k)$  on  $p_w$  and  $\epsilon$  is suppressed in order to avoid cluttering the notation.

**Theorem 3.1.** Suppose that Assumptions 2.1–2.6 are met and that  $H_0^A$  holds. Then, as  $N, T \to \infty$  with  $T/N \to 0$ ,

$$F(\mathcal{T}_{k^{0}}^{0}) \to_{w} Q(\Lambda_{k^{0}}^{0}) =_{d} F_{k^{0}p_{w}, N(T-p_{x}^{-}(k^{0}+1)p_{w}) - p_{x}^{-}(k^{0}+1)p_{w}}, \quad (3.17)$$

where  $\rightarrow_w$  and  $=_d$  signify weak convergence and equality in distribution, respectively.

The requirement that  $T/N \rightarrow 0$  does not seem to be very important in practice. Indeed, as we demonstrate in our Monte Carlo study reported in the supporting information appendix, the proposed methods perform well even when *N* and *T* are similar in size (see, e.g., Pesaran 2006; Westerlund and Urbain 2015, for similar results).

If  $\mathcal{T}_k^0$  is unknown, which is the empirically most plausible scenario, the following statistic may be used:

$$\sup F(k) = \sup_{\mathcal{T}_k \in \mathcal{T}_{k,\epsilon}} F(\mathcal{T}_k).$$
(3.18)

This test is feasible if *k* is "small". If *k* is "large" it becomes computationally very costly to find the set  $\mathcal{T}_k$  that maximizes  $F(\mathcal{T}_k)$ . Grid search requires  $O(T^k)$  least squares operations. By contrast, by using the efficient breakpoint estimation algorithm presented in Section 3.2, we can limit the number of operations to  $O(T^2)$  for any *k*. The basic idea is to first apply this algorithm to obtain  $\hat{\mathcal{T}}_k = \{\hat{T}_1, \dots, \hat{T}_k\}$ . Analogously to  $T_0, T_{k+1}$ , and  $\mathcal{T}_0$ , we define  $\hat{T}_0 = 1$ ,  $\hat{T}_{k+1} = T$ , and  $\hat{\mathcal{T}}_0 = \emptyset$ . Given  $\hat{\mathcal{T}}_k$ , we compute sup  $F(k) = F(\hat{\mathcal{T}}_k)$  and use this as our test statistic. The asymptotic distribution of sup F(k) is a direct consequence of Theorem 3.1, and it is reported in Corollary 3.1 below. However, before we state the corollary, we present the proposed tests of hypotheses (B) and (C).

Testing  $H_0^B$  versus  $H_1^B$  can be done using the following panel version of BP98's "weighted double maximum" statistic:

WDmax 
$$F(k_{max}) = \max_{1 \le k \le k_{max}} \frac{c_{\alpha,1}}{c_{\alpha,k}} \sup F(k),$$
 (3.19)

where  $c_{\alpha,k}$  is the critical value of sup F(k) at significance level  $\alpha$  and k breaks.<sup>3</sup>

The test of  $H_0^C$  versus  $H_1^C$  can be carried out using the following statistic:

$$F(k+1|k) = \sup_{1 \le j \le k+1} \sup_{T_+ \in \widehat{\mathcal{T}}_{j,e}} F(T_+ | \widehat{\mathcal{T}}_k), \qquad (3.20)$$

where  $\hat{\mathcal{T}}_k$  is a set of k estimated (or known) break dates stipulated under the hull hypothesis,  $T_+$  is the additional (k + 1)-th break under the alternative, and

$$\hat{\mathcal{T}}_{j,\epsilon} = \left\{ T_+ : \hat{T}_{j-1} + \Delta \hat{T}_j \epsilon \le T_+ \le \hat{T}_j - \Delta \hat{T}_j \epsilon \right\}$$
(3.21)

is the set of permissible breaks in between the estimated (j - 1)-th and *j*-th breaks. Hence, F(k + 1|k) is testing the null of *k* breaks versus the alternative that there is an additional break

somewhere within the regimes stipulated under the null. Finally,  $F(T_+|\hat{\mathcal{T}}_k)$  is given by

$$F(T_{+}|\hat{\mathcal{T}}_{k}) = \frac{N(T - p_{x} - (k+2)p_{w}) - p_{x} - (k+2)p_{w}}{p_{w}} \times \hat{\delta}\left(\left\{\hat{\mathcal{T}}_{k}, T_{+}\right\}\right)' R_{j}' \left(R_{j}\hat{V}_{\hat{\delta}}R_{j}'\right)^{-1} R_{j}\hat{\delta}\left(\left\{\hat{\mathcal{T}}_{k}, T_{+}\right\}\right)$$
(3.22)

where  $\hat{\delta}(\{\hat{\mathscr{T}}_k, T_+\})$  comes from a regression with k + 1 breaks at dates  $\{\hat{\mathscr{T}}_k, T_+\}$ . The same is true for  $\hat{V}_{\hat{\delta}}$ . The matrix  $R_j$  is given by  $R_j = (0_{p_w \times 1}, \dots, 0_{p_w \times 1}, I_{p_w}, -I_{p_w}, 0_{p_w \times 1}, \dots, 0_{p_w \times 1})$ , where  $I_{p_w}$  sits in the *j*-th position, while  $-I_{p_w}$  sits in the (j + 1)-th, so that  $(R_j \delta)' = \delta'_j - {\delta'_{j+1}}$ .<sup>4</sup>

**Corollary 3.1.** Suppose that Assumptions 2.1–2.6 are met and that the null hypothesis of each test holds. Then, as  $N, T \rightarrow \infty$  with  $T/N \rightarrow 0$ ,

- (a)  $\sup F(k) \to_w \sup Q(k)$ ,
- (b) WDmax  $F(k_{max}) \rightarrow_w \max_{1 \le k \le k_{max}} \frac{c_{a,1}}{c_{a,k}} \sup Q(k),$
- (c)  $P[F(k+1|k) \le x] \to (P[\sup Q(1) \le x])^{k+1}$ .

The asymptotic results given in Corollary 3.1 are the same as the ones given in Propositions 6 and 7 of BP98. This is convenient because it means that appropriate critical values are already available. Critical values for sup F(k) and WDmax  $F(k_{max})$  are reported for  $\epsilon = 0.05$ ,  $k \in \{1, ..., 9\}$  and  $p_w \in \{1, ..., 10\}$  in Table 1 of BP98, and in Table 2, they report critical values for F(k + 1|k). Bai and Perron (2003a, Table 1) report response surface regressions for all tests that are valid for more values  $\epsilon$ , k and  $p_w$ .

The F(k + 1|k) test can be applied sequentially for k = 0,1,... to estimate the number of breaks. In this case, we start by testing the null of no breaks against the alternative of a single break using F(1|0). If the null is accepted, we set  $\hat{k} = 0$  and terminate the procedure. If, however, the null is rejected, we estimate the breakpoint, denoted  $\hat{T}_1$ , and split the sample in two at  $\hat{T}_1$ . We then test for the presence of a break in each of the two subsamples using F(2|1). If no breaks are found, we set  $\hat{k} = 1$  and stop, whereas if breaks are detected, we estimate their location and split the sample again. This process continues until the test fails to reject or until the maximum permissible number of breaks is reached. This number is a function of the trimming parameter  $\epsilon$  and is given by  $|1/\epsilon| - 2$ , where |x| denotes the integer part of x.

A problem with the sequential approach just described is that it does not account for the multiplicity of the testing problem. It will therefore reject too often. In order to prevent this from happening the significance level of each test in the sequence,  $\alpha$ , say, should be set as a decreasing function of the sample size.

**Theorem 3.2.** Suppose that Assumptions 2.1–2.6 are met. Suppose also that F(k + 1|k) is applied sequentially at significance level  $\alpha$  such that  $NT\alpha \rightarrow K > 0$ . Then, as  $N, T \rightarrow \infty$  with  $T/N \rightarrow 0$ ,

$$P(\hat{k} = k^0) \to 1.$$
 (3.23)

iable Definition			
Dependent variables			
Total loans	Log of total loans		
RE loans	Log of real estate loans		
C&I loans	Log of commercial and industrial loans		
Regressors			
MBS	Mortgage-backed securities over total assets		
non-MBS	Total securities held minus MBS securities over total assets		
Total assets	Log of total assets		
Equity	Equity over total assets		
ROA	Return on assets		
Bank cash	Balance sheet cash flow over total assets		
Cost of deposits	Interest expense of deposits over total assets		
Net income	Net income over total assets		
Inflation	The first difference of US inflation		
GDP	US per capita GDP growth		

Note: All regressors are lagged one period. The data are sourced from the Federal Deposit Insurance Corporation and FRED.

Variable	Mean	STD	Min	Max	CD	CIPS	
Total loans	11.709	1.482	6.590	20.773	10124.055***	-4.19***	
RE loans	11.273	1.610	3.638	20.008	9489.100***	-4.41***	
C&I loans	9.508	1.709	0.693	19.684	5728.323***	-4.53***	
MBS	0.075	0.092	0.000	0.820	623.285***	-3.88***	
non-MBS	0.161	0.125	0.000	0.926	2089.694***	-4.43***	
Total assets	12.228	1.399	8.004	21.914	14258.548***	-4.66***	
Equity	0.111	0.032	-0.026	0.730	2095.189***	-4.58***	
ROA	0.103	0.082	-3.083	2.548	1593.754***	-5.18***	
Bank cash	0.082	0.081	0.000	0.939	4513.740***	-5.34***	
Cost of deposits	0.006	0.006	0.000	0.123	19083.858***	-4.82***	
Net income	0.006	0.006	-0.209	0.113	9888.248***	-5.16***	
Inflation	0.002	0.770	-3.847	1.931	_	_	
GDP	0.887	1.792	-9.793	8.184	_	_	

TABLE 2Descriptive statistics.

*Note:* "Mean", "STD", "Min", and "Max" refer to the sample average, the sample standard deviation, the sample minimum value, and the sample maximum value. While the column labelled "CD" reports the results obtained by applying Pesaran's (2021) test for cross-sectional correlation, the column labelled "CIPS" reports the results obtained by applying Pesaran's (2007) panel test for a unit root to the defactored variables. The CD and CIPS test results for GDP and inflation growth are not reported as these variables do not vary by bank. The superscripts "\*", "\*\*", and "\*\*\*" denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Theorem 3.2 requires that  $\alpha$  converges to zero. In empirical practice, however, just as *N* and *T* are never infinite,  $\alpha$  does not need to be zero. Our simulation results show that  $\alpha = 0.05$  works satisfactorily, as do the results of Bai (1999).

## 3.2 | Breakpoint Estimation

In the previous section, we were concerned with testing for the existence of breaks and with estimating their number,  $k^0$ . Once  $k^0$  has been estimated, interest turns to the location of the breaks,

which is the topic of this section. According to Theorem 3.2, knowing  $\hat{k}$  is as good as knowing  $k^0$ , at least asymptotically. In this section, we therefore treat  $k^0$  as known.

The problem of estimating the breakpoints is not independent of that of testing for breaks. In fact, the  $\hat{\mathcal{T}}_k$  that minimizes the sum of squared residuals (for a given k) is the same as the one that maximizes  $F(\mathcal{T}_k)$ . The breakpoint estimator that we will employ is therefore given by

$$\hat{\mathcal{T}}_{k^0} = \arg\min_{\mathcal{T}_{k^0} \in \mathcal{T}_{k^0,c}} SSR(\mathcal{T}_{k^0}) = \arg\max_{\mathcal{T}_{k^0} \in \mathcal{T}_{k^0,c}} F(\mathcal{T}_{k^0}), \quad (3.24)$$

$$SSR(\mathcal{T}_k) = \hat{\varepsilon}'\hat{\varepsilon} = \left(\tilde{Y} - \tilde{W}(\mathcal{T}_k)\hat{\delta}(\mathcal{T}_k)\right)' M_{\tilde{X}}\left(\tilde{Y} - \tilde{W}(\mathcal{T}_k)\hat{\delta}(\mathcal{T}_k)\right)$$
(3.25)

Consider the model for  $\tilde{Y}$  in (2.11). If  $\beta = 0_{p_{x} \times 1}$  was known, then  $\tilde{Y} = \tilde{W}(\mathcal{T}_{k^0})\delta + \tilde{E}$  is a pure structural change model, and for such a model,  $\mathcal{T}_{\nu^0}$  and  $\delta$  can be estimated using the dynamic programming algorithm of Bai and Perron (2003b), "BP03" henceforth. While initially proposed as an efficient way to minimize the sum of squared residuals in the pure time series context, the algorithm can be extended in a straightforward manner to the current more general context, as it is just a way to compare possible combinations of breakpoints to achieve a minimum global sum of squared residuals. The efficiency of the algorithm comes from recognizing that with T times series observations the total number of possible sample splits is T(T+1)/2 for any k and is therefore  $O(T^2)$ . With  $\beta$  known but not necessarily zero, then  $\tilde{X}\beta$  can be subtracted from  $\tilde{Y}$  and the dynamic programming algorithm can be applied to  $\tilde{Y} - \tilde{X}\beta = \tilde{W}(\mathcal{T}_{k^0})\delta + \tilde{E}$ , which is again a pure structural change model. With  $\beta$  unknown, the estimation can be carried out in an iterative fashion, which we will now describe.

#### Breakpoint estimation algorithm:

- 1. Initiate  $\hat{\beta}$  by treating the coefficients of both  $x_{i,t}$  and  $w_{i,t}$  as subject to structural change. Define  $\tilde{X}(\mathcal{T}_{k^0})$  similarly to  $\tilde{W}(\mathcal{T}_{k^0})$  and apply BP03's dynamic programming algorithm to the pure structural change model  $\tilde{Y} = \tilde{X}(\mathcal{T}_{k^0})\beta + \tilde{W}(\mathcal{T}_{k^0})\delta + \text{error}$ , where  $(\overline{X}(\mathcal{T}_{k^0}), \overline{W}(\mathcal{T}_{k^0}))$  is used in place of  $\overline{Z}(\mathcal{T}_{k^0})$  to estimate the factors. This yields  $\hat{\mathcal{T}}_{k^0}, \hat{\beta}, \text{ and } \hat{\delta}$ .
- 2. Update  $\hat{\beta}$  by fitting  $\tilde{Y} \tilde{W}(\hat{\mathcal{T}}_{k^0})\hat{\delta} = \tilde{X}\beta + \text{error by OLS}$ , this time using  $\overline{X}$  in place of  $\overline{Z}(\mathcal{T}_{k^0})$ .
- 3. Update  $\hat{\mathcal{T}}_{k^0}$  and  $\hat{\delta}$  by applying the dynamic programming algorithm to  $\tilde{Y} \tilde{X}\hat{\beta} = \tilde{W}(\mathcal{T}_{k^0})\delta + \text{error using }\overline{W}(\mathcal{T}_{k^0})$  in place of  $\overline{Z}(\mathcal{T}_{k^0})$ .
- 4. Update  $\hat{\beta}$  and  $\hat{\delta}$  by estimating (2.11) by OLS conditional on  $\hat{\mathcal{T}}_{k^0}$  (and using  $\overline{Z}(\hat{\mathcal{T}}_{k^0})$  to estimate the factors).
- 5. Iterate steps 3 and 4 until convergence.

Notice how the averages change in every step. As mentioned in Section 2.2, if there are known factors, then these should be appended to the averages at every step.

The above algorithm is similar in spirit to the one considered by BP03 in case of a partial structural change model. While convergence to the global minimum is not guaranteed, convergence to a local optimum for the resulting iterated estimator has been shown by Sargan (1964). Tests with both simulated and real data in BP03 and here confirm that convergence is fast, typically with only one iteration needed and rarely a second.

In what follows, we prove that  $\hat{\mathcal{T}}_{k^0}$  is consistent and derive its limiting distribution.

**Theorem 3.3.** Suppose that Assumptions 2.1–2.5 are met. Then, for all  $j = 1, ..., k^0 + 1$ , as  $N, T \to \infty$  with  $T/N \to 0$ ,

$$P\left(\hat{T}_{j} = T_{j}^{0}\right) \to 1.$$
(3.26)

As is well-known, with time series data consistent estimation of the breakpoints is not possible, but only consistent estimation of the break fractions, and this is true also for BP98. By contrast, Theorem 3.3 states that  $\hat{\mathcal{T}}_{k^0}$  is consistent. The accuracy of the estimated breakpoints is therefore greatly enhanced when compared to the time series case. This is important in our empirical study, because policy-induced breaks have proven to be difficult to detect using pure time series techniques (see, e.g., Lubik and Surico 2010).

Define  $\xi_j = \Delta'_j \Omega_{j+1} \Delta_j / \Delta'_j \Omega_j \Delta_j$ ,  $\phi_{1,j} = \Delta_j \Phi_j \Delta_j / \Delta'_j \Omega_j \Delta_j$  and  $\phi_{2,j} = \Delta_j \Phi_{j+1} \Delta_j / \Delta'_j \Omega_{j+1} \Delta_j$ , where  $\Delta_j = \delta_{j+1} - \delta_j$ , and  $\Phi_j$  and  $\Omega_j$  are as in Assumption 2.6. Let  $B_{1,j}(s)$  and  $B_{2,j}(s)$  be two scalar standard Brownian motions on  $[0, \infty)$  that are independent of each other as well as over *j*. Also,  $B_{1,j}(0) = B_{2,j}(0) = 0$ . We now define  $V_j(s)$  such that  $V_j(s) = B_{1,j}(-s) - |s|/2$  if  $s \le 0$  and  $V_j(s) = \sqrt{\xi_j \phi_{2,j} / \phi_{1,j}} B_{2,j}(s) - \xi_j |s|/2$  if s > 0. We now have everything we need in order to state the asymptotic distribution of  $N(\hat{T}_j - T_j^0)$ .

**Theorem 3.4.** Suppose that Assumptions 2.1–2.6 are met. Then, for all  $j = 1, ..., k^0 + 1$ , as  $N, T \to \infty$  with  $T/N \to 0$ ,

$$\frac{(\Delta'_{j}\Omega_{j}\Delta_{j})^{2}}{\Delta_{j}\Phi_{j}\Delta_{j}}N(\hat{T}_{j}-T^{0}_{j}) \to_{w} \arg\max_{s\in[0,\infty)}V_{i}(s).$$
(3.27)

The probability density function of  $\arg \max_{s \in [0,\infty)} V_i(s)$  is known analytically and is given in Bai (1997). The density function depends on  $\sqrt{\xi_j \phi_{2,j}/\phi_{1,j}}$  and  $\xi_j$ , which can be estimated given estimates of  $\Delta_j$ ,  $\Phi_j$ , and  $\Omega_j$ . Such an estimate can be constructed as  $\hat{\Delta}_j = \hat{\delta}_{j+1} - \hat{\delta}_j$  and by specifying  $\hat{\Phi}_j$  and  $\hat{\Omega}_j$  analogously to  $\hat{\Phi}$ and  $\hat{\Omega}$ , respectively, but for the estimated *j*-th subsample spanning the interval  $[\hat{T}_{j-1} + 1, \hat{T}_j]$ . For example, in case of  $\hat{Q}_j$ , we take

$$\hat{\Omega}_{j} = \frac{1}{N\Delta\hat{T}_{j}} \sum_{i=1}^{N} \sum_{t=\hat{T}_{j-1}+1}^{T_{j}} \check{w}_{i,t} \check{w}_{i,t}', \qquad (3.28)$$

where  $\tilde{w}_{i,t}$  is as before. Once  $\hat{\Delta}_j$ ,  $\hat{\Phi}_j$ , and  $\hat{\Omega}_j$  have been obtained, we can construct  $\hat{\xi}_j$ ,  $\hat{\phi}_{1,j}$ , and  $\hat{\phi}_{2,j}$  by plugging in estimates in place of true parameters. Denote by  $c_{\alpha}$  the  $(1 - \alpha/2)$ -th percentile of the probability density function of  $\arg \max_{s \in [0,\infty)} V_i(s)$ . In analogy to Bai (1997), an asymptotically correctly sized  $100(1 - \alpha)\%$ confidence interval for  $T_j^0$  can now be constructed in the following way:

$$\left[\hat{T}_{j}-\left|c_{\alpha}\frac{\hat{\Delta}_{j}^{\prime}\hat{\Phi}_{j}\hat{\Delta}_{j}}{N\left(\hat{\Delta}_{j}^{\prime}\hat{\Omega}_{j}\hat{\Delta}_{j}\right)^{2}}\right]-1,\,\hat{T}_{j}+\left|c_{\alpha}\frac{\hat{\Delta}_{j}^{\prime}\hat{\Phi}_{j}\hat{\Delta}_{j}}{N\left(\hat{\Delta}_{j}^{\prime}\hat{\Omega}_{j}\hat{\Delta}_{j}\right)^{2}}\right]+1\right].$$
(3.29)

Because the break dates are integer valued, the confidence interval is integer valued, too.

Once the presence of breaks has been established and their locations determined,  $\beta$  and  $\delta$  can be estimated by simply applying OLS to (2.11) with  $\mathcal{T}_k$  replaced by  $\hat{\mathcal{T}}_k$ .

# 4 | The Impact of Quantitative Easing on Bank Lending

# 4.1 | Motivation

Since the 2007-2008 global financial crisis and until recently, short-term interest rates have been close to the zero lower bound in many countries, and central banks have therefore been constrained to unconventional monetary policy. The main policy instrument has been QE. The US Federal Reserve, in particular, has implemented at least four major rounds of QE through which they purchased US Treasuries and mortgage-backed securities (MBSs) from the commercial banking sector, with the aim of boosting lending and stimulating economic activity. The most well-known are the "QE1", "QE2", and "QE3" rounds that took place in the aftermath of the global financial crisis. QE1 was announced in November 2008 and lasted until June 2010 and was followed by QE2, which spanned the period November 2010 to June 2011. QE3 began in September 2012 and ended in October 2014. The fourth and most recent QE round is the one initiated in March 2020 as a response to the COVID-19 pandemic, henceforth labelled "OE4".

A common feature of QE1–QE4 is that they are massive in scale. As a reflection of this, the Federal Reserve balance sheet increased from about USD 800 billion in 2007 to over USD 8.5 trillion in 2021. This is clearly visible in Figure 1, which plots the Federal Reserve Treasury and MBS holdings over time, together with the QE dates.

Interestingly enough, however, while massive indeed, so far the empirical evidence regarding the effectiveness of QE has been mixed and far from conclusive. Rodnyansky and Darmouni (2017) were among the first to examine the first three QE rounds jointly. They found that banks with relatively larger holdings of MBSs expanded lending, but only during QE1 and QE3, because QE2 targeted Treasuries which were sparsely held by banks. Similar results have later been reported by Luck and Zimmermann (2020). Kapoor and Peia (2021) also study QE1-QE3. According to their results, however, only QE3 had a strong effect on liquidity creation. This last finding is in turn markedly different from that of Chakraborty, Goldstein, and MacKinlay (2020), who document a negative effect of QE3 on commercial and industrial (C&I) lending. For the most recent QE4 round, there is to the best of our knowledge no evidence at all.

The present paper is the first to consider all four rounds. This is our first contribution. Our second contribution lies in our choice of econometric method. The standard approach in the literature is to exploit differences in exposure to QE policies across banks. The basic idea is to split the sample of banks into a treatment and a control group, where the former is assumed to be relatively more exposed to QE policy. Given that the Federal Reserve bought large quantities of mainly MBSs during the QE rounds, the argument goes on to say that banks with relatively large MBS holdings should benefit more, and hence be more exposed. The effect of QE policy is then estimated via a standard difference-in-difference (DiD) regression in which banks' lending is regressed onto a dummy variable that takes on the value one whenever a bank that belongs to the treatment group has been subject to a particular QE policy and zero otherwise, control variables, and bank and time fixed effects.

While popular, the standard DiD approach to QE evaluation has (at least) two drawbacks. One drawback is that fixed effects are highly restrictive in that they require that in the absence of treatment the difference between the treatment and control groups is constant over time. This is the so-called "parallel trend" condition, which has attracted considerable attention in the QE literature (see, e.g., Di Maggio, Kermani, and Palmer 2020; Luck and Zimmermann 2020; Rodnyansky and Darmouni 2017). The reason is that banks with higher exposure to QE may be lending to firms that experience faster credit demand growth due to improvements in their borrower health. Changes in the demand for credit may therefore cause the treatment and control groups to differ systematically over time even in absence of QE. Demand-side effects is one source of nonparallel trending, but there are others, such as aggregate macroeconomic conditions and omitted variables (see, e.g., Di Maggio, Kermani, and Palmer 2020; Kapoor and Peia 2021, for discussions). This is important because if the parallel trend condition is violated, the fixed effects OLS estimator is no longer consistent.

Another drawback of the DiD approach is that it requires correct specification not only of the timing of the QEs but also of the treatment and control groups, both of which are key in the construction of the treatment dummy. As mentioned earlier, existing studies focus on QE1–QE3, leaving "Operation Twist", the QE1 rollover, QE extensions and other sizeable Federal Reserve purchases outside the QE rounds out of the analysis. Furthermore, the rounds were lengthy and contained periods with varying degree of asset purchases, and it is not clear if their impact began with the announcements (see Luck and Zimmermann 2020).

If it is difficult to get the timing of the QEs right, correct specification of the treatment and control groups is literally impossible, given that strictly speaking there is no control group, since all banks are exposed to QE to some extent (see Chakraborty, Goldstein, and MacKinlay 2020; Luck and Zimmermann 2020). The standard approach is to measure banks' exposure to QE by their MBS holdings, and to define the treatment and control groups as the upper and lower quantiles of the MBS distribution (prior to QE1). This raises (at least) two concerns; some QE policies focused solely on US Treasuries and therefore banks' MBS holdings are likely an imperfect measure of their QE exposure, and the choice of which quantiles to use is arbitrary. The treatment dummy is therefore generally mismeasured, which is a major problem as it renders OLS inconsistent (see, e.g., Mahajan 2006).

The methodology developed in Section 3 overcomes both drawbacks. It allows for interactive fixed effects, which means that it dispenses with the parallel trend condition.<sup>5</sup> It also removes the need for a correctly specified treatment dummy. Instead, we take our measures of banks' exposure to QE and include them as regressors with potentially breaking coefficients, which means that the timing of any QE effect is determined in a data-driven way. And the proposed methodology should be accurate enough to detect such effects, as opposed to existing time series techniques (see, e.g., Lubik and Surico 2010).



FIGURE 1 | Federal reserve's MBS and treasury holdings and QE rounds.

## 4.2 | Data

The analysis is based on bank-level quarterly data taken from the Call Reports of the Federal Deposit Insurance Corporation. The initial sample is the universe of commercial banks, observed from 2005Q3 to 2021Q3. The starting period coincides with the first Federal Reserve Treasury purchases. Following Rodnyansky and Darmouni (2017) and others, we balance our sample, keeping only those banks for which we have observations for the whole time period to control for mergers or acquisitions. Many studies in the literature aggregate their data to bank holding company level. However, whenever tested, the results for the bank level data tend to be the same (see, e.g., Rodnyansky and Darmouni 2017). In the present paper, we therefore base our analysis on this last type of data, which, in addition to being free of aggregation bias, contains a relatively large number of observations. In particular, there are N = 3557 banks that are observed over T = 64 quarters, for a total of 227,648 bank-quarter observations, which is substantially larger than many of the samples employed in the existing literature (see, e.g., Chakraborty, Goldstein, and MacKinlay 2020; Rodnyansky and Darmouni 2017).

We employ the same three dependent variables ( $y_{i,l}$ ) as in Luck and Zimmermann (2020), Kapoor and Peia (2021), and Rodnyansky and Darmouni (2017). They are the logarithm of (i) total loans, (ii) real estate (RE) loans, and (iii) commercial and industrial (C&I) loans. The main regressor of interest, whose coefficient will be allowed to be breaking, is banks' MBS holdings. Unlike most existing studies, however, for reasons mentioned earlier, we do not want to rely solely on MBS holdings as a measure of banks' exposure to QE. Because of the way that they enter the model as regressors with potentially breaking coefficients, the proposed methodology can easily accommodate additional measures. Rodnyansky and Darmouni (2017) consider Treasuries holdings as an alternative to MBS holdings. According to their results, however, this measure is relatively unimportant, which they explain by the fact that banks do not hold many Treasury securities. Chakraborty, Goldstein, and MacKinlay (2020) therefore suggest an alternative, indirect, measure of the effect of Treasury purchases. In particular, they argue that banks with relatively high non-MBS securities holdings should benefit more from Treasury purchases lowering yields on these securities. In this section, we therefore include banks' non-MBS securities holdings as an additional measure of their exposure to QE. As usual in the literature, both measures are normalised by total assets. Also, to reduce the risk of simultaneity, following Chakraborty, Goldstein, and MacKinlay (2020), all regressors are lagged once. Hence, in terms of the model in (2.1),  $w_{i,t}$  is lagged MBS and non-MBS holdings over total assets.

A number of bank-level controls are included to capture differences in the scale and financial position of banks that might affect their lending activity, again following the convention in the literature. They are ROA, total assets, equity over total assets, cash over total assets, cost of deposits, and net income. These are the variables that go into  $x_{i,t}$ . GDP growth and inflation in first differences are also included to control for macroeconomic conditions and are treated as observed common factors, as discussed in Section 2.2.<sup>6</sup> The data for these last two variables are extracted from the FRED database of the Federal Reserve Bank of St. Louis.

Tables 1 and 2 provide the definition of each variable and some descriptive statistics, respectively. One of the descriptives included is Pesaran's (2021) CD test for the presence of cross-sectional correlation. According to the test results, the null hypothesis of no correlation is strongly rejected for all the bank-level variables, which is suggestive of common factors, and hence of nonparallel trending. Thus, as Di Maggio, Kermani, and Palmer (2020) point out, the parallel trend condition is unlikely to

		Total loans	RE loans		C&I loans	
Break	Date	95% CI	Date	95% CI	Date	95% CI
1	2007Q1	[2006Q4, 2007Q2]	2007Q1	[2006Q4, 2007Q2]	2007Q1	[2006Q4, 2007Q2]
2	2009Q1	[2008Q4, 2009Q2]	2009Q1	[2008Q4, 2009Q2]	2009Q1	[2008Q4, 2009Q2]
3	2011Q3	[2011Q2, 2011Q4]	2011Q3	[2011Q2, 2011Q4]	2011Q3	[2011Q2, 2011Q4]
4	2013Q3	[2013Q2, 2013Q4]	2013Q3	[2013Q2, 2013Q4]	2013Q3	[2013Q2, 2013Q4]
5	2014Q4	[2014Q3, 2015Q1]	2014Q4	[2014Q3, 2015Q1]	2014Q4	[2014Q3, 2015Q1]
6	2017Q1	[2016Q4, 2017Q2]	2017Q1	[2016Q4, 2017Q2]	2016Q4	[2016Q3, 2017Q1]
7	2020Q1	[2019Q4, 2020Q2]	2020Q1	[2019Q4, 2020Q2]	2020Q1	[2019Q4, 2020Q2]

Note: "Date" and "95% CI" refer to the estimated breakpoint and the 95% confidence interval, respectively, for each of the seven breaks.

hold when using the type of low-frequency panel data considered here because the infrequent observations capture also changes in the demand of credit or other aggregate shocks. Another descriptive reported in Table 2 is Pesaran's (2007) panel unit root test, which is applied to the defactored variables. The unit root null hypothesis is again strongly rejected for all variables, which we take as evidence in support of stationarity.

# 4.3 | Break Testing and Estimation

We first want to test if there are any breaks present at all. Given that the QEs were of different sizes and contained different policy mixes, we do not want to assume that all of them led to breaks. This means that we want to treat the number of breaks as unknown. For this reason, we employ the WDmax  $F(k_{max})$  statistic, in which the null hypothesis of no breaks is tested against the alternative of up to  $k_{max}$  breaks. Because the test outcome was unaffected by the weighing, we set  $c_{\alpha,1}/c_{\alpha,k} = 1$ . We also set  $k_{max} = 9$  and  $\epsilon = 0.05$ , which means that the critical values can be taken directly from Table I of BP98. The test values for total, RE and C&I loans are 246.877, 97.722 and 28.671, respectively, which are all larger than the appropriate critical value at the most conservative 1% level, 17.61.

The number of breaks turned out to be a difficult object to estimate, which is partly expected given the discussion in BP03. The main problem is, as BP98 point out, that while for relatively small values of k the critical values of the F(k + 1|k) test increase markedly when an additional break is added, for  $k \ge 5$  the critical values are quite flat in k. As a result, the estimated number of breaks can sometimes be large, and this is also what we find. In fact, the number of breaks is always estimated to  $k_{max}$ , regardless of how we set this tuning parameter. These estimates are therefore not reliable.

On the other hand, the exact choice of *k* to use in the estimation of (2.1) should not matter too much, as long as *k* is chosen large enough to cover all relevant breaks. Unreported results suggest that the estimation results for  $\beta$  and  $\delta_1, \ldots, \delta_{k+1}$  are quite sensitive to changes in *k* for small values of *k*, in that the estimate of  $\delta_{k+1}$  is markedly different from that of  $\delta_k$ , and the estimates of  $\beta$  from one value of *k* to another differ, too, which is presumably a reflection of underspecification-induced inconsistency. As *k* increases, however, these differences tend to become smaller

and from k = 7 onward they are almost completely absent. The fact that the estimated coefficients are stable after k = 7 suggests that with this many breaks at least the model is not underspecified.<sup>7</sup> In what follows, we therefore employ seven breaks.

Table 3 reports the estimated breakpoints and the associated 95% confidence interval for each of the seven breaks. The breaks are precisely estimated, having very narrow confidence intervals covering only a few quarters before and after, and this is happening even though some of the regimes are estimated to be quite short. The reason is that in this empirical application N is substantial, which according to Theorem 3.4 is expected to lead to accurate breakpoint estimation.<sup>8</sup> The estimated breakpoints for total and RE loans coincide and are very similar to those for C&I loans.

Of direct interest are the estimated break dates, which all coincide with major (QE) events. The first break takes place at about the same time as the first signs of the global financial crisis became visible. On February 27, 2007, stock prices in China and the USA fell by the most since 2003 as reports of a decline in home prices and durable goods orders led to growth fears, with former Chairman of the Federal Reserve Alan Greenspan predicting a recession. The second and seventh breaks coincide with the start of QE1 and QE4, respectively, which were the largest QE rounds. The third break coincides with the announcement of Operation Twist. The fourth and fifth breaks coincide with the tapering and the end of QE3, respectively. The sixth break is estimated close to the start of Federal Reserve's balance sheet normalisation program.

It is important to note that the above characterization of the estimated breaks is not free of errors. In particular, there may have been other, confounding, factors that have also affected bank lending. As alluded to earlier, however, the events that we single out were among the largest at the time and they are likely to have affected bank lending.

Another major point about the breakpoints reported in Table 3 is that they are not the same as the conventional QE dates employed in the literature. Indeed, while the starting dates of QE1 and QE3 are within the 95% confidence intervals for the second and fourth breaks, respectively, and the end date of QE3 is within the interval for the fifth, the end of QE1 is not included, as is the entire QE2. It therefore seems as that some of the conventional QE

	Total loans		<b>RE loans</b>		C&I loans	
COEF	EST	SE	EST	SE	EST	SE
Regressor with breaking coef	ficient: MBS					
$\delta_1$	0.180	0.209	0.399**	0.161	0.113	0.411
$\Delta_1$	-0.293	0.215	-0.572***	0.168	-0.408	0.425
$\Delta_2$	-0.0233	0.053	-0.046	0.059	$0.291^{**}$	0.125
$\Delta_3$	-0.192***	0.043	-0.185***	0.048	$-0.197^{*}$	0.104
$\Delta_4$	-0.209***	0.057	-0.227**	0.089	-0.555***	0.139
$\Delta_5$	0.329***	0.059	0.354***	0.100	0.782***	0.165
$\Delta_6$	0.033	0.053	0.121	0.074	-0.289**	0.142
$\Delta_7$	0.237***	0.085	0.166	0.103	0.209	0.199
Regressor with breaking coef	ficient: non-MBS					
$\delta_1$	$0.122^{**}$	0.051	0.092	0.060	-0.100	0.114
$\Delta_1$	-0.223***	0.060	-0.299***	0.073	-0.003	0.134
$\Delta_2$	-0.060	0.040	0.010	0.049	-0.137	0.091
$\Delta_3$	$-0.074^{**}$	0.032	$-0.101^{***}$	0.038	-0.033	0.082
$\Delta_4$	-0.259***	0.042	-0.219***	0.056	-0.213*	0.112
$\Delta_5$	0.273***	0.059	0.273***	0.076	0.424***	0.139
$\Delta_6$	-0.063	0.060	-0.059	0.071	-0.160	0.129
$\Delta_7$	$0.290^{***}$	0.055	0.204***	0.064	0.490***	0.132
Interactive effects	Yes		Yes		Yes	s
Observed factors	Yes		Yes		Yes	s
Bank-level controls	Yes		Yes		Yes	s
Rank condition satisfied	Yes		Yes		Yes	S

*Note:* The table reports the results obtained by fitting (4.1) by CCE while allowing the coefficients of both MBS and non-MBS to be breaking. "COEF", "EST", and "SE" refer to the estimated coefficient, the associated point estimate and its standard error, respectively. Here,  $\delta_1$  is the coefficient of the first regime, while  $\Delta_j = \delta_{j+1} - \delta_j$  is the change in the coefficient from regime *j* to *j* + 1. All models are fitted with cross-sectional averages of the regressors to account for unknown interactive effects and fixed effects, GDP and Inflation as observed common factors. The fixed effects are allowed to be breaking, while the coefficients of GDP and Inflation are not. The included bank-level controls are Total assets, Equity, ROA, Bank cash, Cost of deposits, and Net income. The rank condition of Assumption 2.4 has been testes using the rank condition classifier of De Vos et al. (2024). The superscripts "\*", "\*\*", and "\*\*\*" denote statistical significance at the 10%, 5% and 1% levels, respectively.

dates have been labelled as "breaks" when in fact they are not. This finding is largely consistent with studies such as Kapoor and Peia (2021), Rodnyansky and Darmouni (2017), and Luck and Zimmermann (2020), which all point to a relatively stronger impact of QE1 and QE3 on lending. More importantly, a number of the dates that we label as breaks are not among the conventional QE dates.

# 4.4 | Model Estimation

With the estimated breaks in hand, we turn to the estimation of (2.1), which we reparameterize in the following way:

$$y_{i,t} = x'_{i,t}\beta + w'_{i,t}\delta_1 + \sum_{j=1}^{k_0} 1(t > T_j)w'_{i,t}\Delta_j + e_{i,t}, \qquad (4.1)$$

where  $\Delta_j = \delta_{j+1} - \delta_j$  is as in Section 3.2 and 1(A) is the indicator function for the event *A* taking the value one if *A* is true and zero otherwise. The reason for why we use this parametrization is because at the end of a QE intervention, the Federal Reserve's

balance sheet does not return to its pre-intervention levels, and therefor, e price effects remain. Therefore, it is not meaningful to compare say, bank lending behavior after QE3 with bank lending behavior before QE1 to determine the effectiveness of QE3, but rather we should compare post-QE3 behavior to pre-QE3 behavior. The reparameterized model in (4.1) with  $\delta_1, \Delta_1, \ldots, \Delta_{k_0}$  as opposed to  $\delta_1, \ldots, \delta_{k_0+1}$  as coefficients captures this.

The expectation is that QE should cause increased lending by banks (see Chakraborty, Goldstein, and MacKinlay 2020). The intuition goes as follows. The Federal Reserve purchases MBSs and Treasuries that are held by banks. Banks sell some of their holdings in these securities to the Federal Reserve, and this increases their reserves.<sup>9</sup> On the other hand, unsold holdings increase in value because of the increased demand.<sup>10</sup> Both the increased reserves and the higher valued MBS and Treasury holdings improve the financial condition of banks, which means that they can increase their lending. Hence, in terms of (4.1), we expect  $\Delta_j$  to be positive following a QE intervention.



FIGURE 2 | Bank lending, the Tier 1 risk based capital ratio and the estimated structural breaks.

As outlined in Section 2, the bank-level regressors in  $x_{i,t}$  and  $w_{i,t}$ are augmented with their cross-section averages to account for interactive effects, and, as mentioned earlier, GDP and inflation growth are included as observed common factors. The loadings of all these factors are assumed to be time-invariant, which rules out the possibility of policy-induced breaks. In the previous literature, it is quite common to allow QE to affect the overall level of lending (see, e.g., Kapoor and Peia 2021; Rodnyansky and Darmouni 2017). We do more. In particular, a breaking constant is included, which is equivalent to allowing for breaking bank fixed affects. This type of breaking observed factors is easily accommodated. In fact, all one has to do is to organize the breaking factors by regime similarly to  $\overline{W}(\mathcal{T}_k)$ , and append it to  $\overline{Z}(\mathcal{T}_k)$ . All-in-all, we include  $p_x + p_w = 6$  averages and three observed factors, which means that we can allow for up to nine common factors where one is potentially breaking. This should be more than enough to capture the unobserved heterogeneity of bank lending, which we verify using the rank condition classifier of De Vos, Everaert, and Sarafidis (2024).

The estimation results for each loan category are reported in Table 4.<sup>11</sup> We begin by considering the results for total loans. The effect of banks' MBS and non-MBS holdings tend to have the same sign, which we interpret as that Federal Reserve's purchases of MBSs and Treasuries work in the same direction. The main exceptions are in the first and second regimes, which coincide with the precrisis and financial crisis periods, respectively. In particular, while the estimated effects are of the same sign, only banks' non-MBS holdings that matters is expected given Figure 1. The estimated effect of these holdings is positive in the first regime (that is,  $\delta_1$  is estimated to be positive), suggesting that increased Treasury purchases by the Federal Reserve

caused banks to increase lending, which is just as expected given the discussion of the previous paragraph. But then in the second regime the sign changes to negative (i.e.,  $\Delta_1$  is estimated to be negative), which means that the estimated effect of banks' non-MBS holdings on their lending decreases significantly when compared to the first regime. This last effect can be explained by considering the health of the banking sector, as we describe in detail below. The effect of banks' MBS holdings on lending decreases too, although not significantly so. From the third regime onwards, however, the estimated effect is basically the same regardless of the type of holding being considered. We also see that for quite some time following the crisis the estimated effect of banks' MBS and non-MBS holdings on lending is either insignificant or significantly negative, suggesting that the impact of Federal Reserve's purchases of MBSs and Treasuries on lending is either the same as during the crisis or even more negative. This development is brought to a halt at the time of the fifth break (the end of QE3) when the estimated effect of banks' MBS and non-MBS holdings on lending increases significantly when compared to previous quarters, which is consistent with the results of Luck and Zimmermann (2020). This increase continues until the end of the sample period.

As a measure of the overall effect of QE on lending, we look at the difference between the first (pre-intervention) regime and the last (QE4) regime, as given by  $\sum_{j=1}^{k_0} \Delta_j = \delta_{k_0+1} - \delta_1$ . The overall effects for MBS and non-MBS holdings are estimated to 0.165 and 0.007, respectively. This is when we ignore coefficient estimates that are insignificant. If we include all estimates, said effects are -0.118 and -0.115, respectively. Hence, even if in the latter part QE seems to have worked, considering the full sample period QE has not been very effective in spurring credit flow. In order to ease the interpretation of the results reported in Table 4, in Figure 2, we plot the average Tier 1 risk based capital ratio, which is an indicator of banks' financial strength and is regulated by Basel III. Between 2008Q2 and 2014Q4, this ratio is increasing steeply, suggesting that banks retain some of the reserves produced by Federal Reserve's purchases in order to improve their financial position.<sup>12</sup> This explains why in this period lending does not increase. The Tier 1 ratio flattens after 2014Q4, only to increase again in 2020Q1. The fact that in the 2014Q4-2020Q1-period, the Tier 1 ratio is relatively stable means that banks are no longer building up reserves to the same extent as before. Banks therefore have more room to expand lending, and this is reflected in the estimation results. The observed increase in the Tier 1 ratio after 2020Q1 is compensated in part by the scale of Federal Reserve's purchases during QE4. Banks are therefore able to expand lending while at the same time cover their capitalisation needs. The importance of banks' financial position for their ability to lend is consistent with studies such as Kim and Sohn (2017).

The above conclusions for total loans apply also to RE loans. In fact, the estimation results are almost identical, except that the contractionary effect of the crisis is more pronounced for RE than for total loans. The results for C&I loans are quite different, though. Note in particular how in the second (financial crisis) regime banks' MBS and non-MBS holdings are no longer significant and that the positive effect in the sixth (end of QE3) regime is much stronger than for total and RE loans. This last difference corroborates the finding of Luck and Zimmermann (2020) that C&I lending has not been very responsive to QE, except post-QE3 when it increases. It also strengthens our conclusion that it is mainly from the fifth break onwards that Federal Reserve's interventions have led to increased lending. The overall effects of MBS and non-MBS holdings on C&I lending are estimated to 0.032 and 0.701, respectively, suggesting that while Federal Reserve's MBS purchases have again not been very effective, its Treasury purchase has had a substantial effect, which consistent with the findings of Chakraborty, Goldstein, and MacKinlay (2020). It is also consistent with the Figure 1 and the massive increase in Federal Reserve's Treasuries holdings during QE4.

## 5 Conclusions

This paper provides a new econometric methodology that meets the needs of researchers interested in a linear panel data model with interactive effects and possibly multiple structural breaks. The new methods allow researchers to test for the presence of breaks, and, if breaks are detected, to estimate the location of the breaks and construct confidence intervals for the true breakpoints.

The new methodology is employed to study the effects of the US Federal Reserve's QE interventions during 2005–2021 on bank lending. The idea behind QE is that banks should transform the reserves generated by the interventions into loans, which should in turn increase consumption and employment. However, since the composition of banks' balance sheets affect the amount of loans they create, the impact of QE on lending is not a priori clear. Our results suggest that while QE has caused a number of breaks, it is only towards the end of the sample period that these breaks have led to a substantial increase in the effect of banks' MBS and non-MBS holdings on lending.

#### Acknowledgments

The authors would like to thank Theologos Dergiades, Tomas Del Barrio Castro, Esfandiar Maasoumi, Sophocles Mavroeidis, Hashem Pesaran, Francesco Ravazzolo, Barbara Rossi, Elias Tzavalis, conference participants at the Bolzano-Waseda Workshop on Statistics & Time Series analysis, the 2023 Africa Meeting of the Econometric Society, the International Conference on Refined Econometrics and Endogeneity, the 7th International Conference on Applied Theory, Macro and Empirical Finance, and seminar participants at the Free University of Bozen-Bolzano, the Center for Econometrics and Business Analytics (CEBA), the University of the Balearic Islands, and the Universities of Birmingham, Dundee and Sheffield for many valuable comments and suggestions. The authors would like to thank Theologos Dergiades particularly for creating Figures 1 and 2. Ditzen acknowledges financial support from Italian Ministry MIUR under the PRIN project Hi-Di NET - Econometric Analysis of High Dimensional Models with Network Structures in Macroeconomics and Finance (grant 2017TA7TYC). Westerlund would like to thank the Knut and Alice Wallenberg Foundation for financial support through a Wallenberg Academy Fellowship.

#### Data Availability Statement

All data and codes necessary to replicate the results in the empirical application are available in the journal data archive (https://doi.org/10 .15456/jae.2024255.1442474978).

#### **Open Research Badges**

This article has been awarded Open Data Badge for making publicly available the digitally-shareable data necessary to reproduce the reported results. Data is available at http://10.15456/jae.2024255.1442474978.

#### Endnotes

- <sup>1</sup> The interactive effects specification is very general and nests many extant specifications. For example, if m = 1 and  $f_i = 1$ , then  $f'_i \gamma_i = \gamma_i$ , which is the usual one-way error components model, whereas if m = 2,  $f_i = (1, v_i)'$  and  $\gamma_i = (\xi_i, 1)'$ , then  $f'_i \gamma_i = \xi_i + v_i$ , which is the usual two-way error components model. In our paper, *m* and the form of  $f'_i \gamma_i$  are unspecified.
- <sup>2</sup> The dependence of  $\tilde{A}$  on  $\mathcal{T}_k$  is suppressed for notational simplicity.
- <sup>3</sup> The weighting by  $c_{\alpha,1}/c_{\alpha,k}$  here ensures that the marginal *p*-values of the weighted supremum statistics are all equal. This counterweights the decrease in the marginal *p*-value of sup F(k) that comes from increasing *k*, and the resulting loss of power when *k* is large.
- <sup>4</sup> If the errors are homoskedastic and serially uncorrelated, then F(k + 1|k) takes the same form as in equation (10) in BP98.
- <sup>5</sup> In terms of the terminology of the treatment effects literature, the factors represent common trends and the loadings measure the extent to which the impact of these trends is equal, or parallel, across units. This means that unless  $f_i = f$  for all t, the parallel trends condition requires that  $\gamma_i = \gamma_j$  for all pairs *i* and *j*. By leaving  $f_i$  and  $\gamma_i$  essentially unrestricted, we can accommodate very general forms of nonparallel trending behavior.
- <sup>6</sup> Inflation is included in first differences because several papers find it to be non-stationary (see, e.g., Ng and Perron 2001). However, the results do not change much if level inflation is included. The results are available upon request.
- <sup>7</sup> See Moon and Weidner (2015) for a similar approach to the selection of the number of common factors.

- <sup>8</sup> As explained in Section 2.2, while needed for asymptotically valid testing, the expanding regime condition in Assumption 2.1 is not required for accurate breakpoint estimation.
- <sup>9</sup> Increased lending brought about by increases in banks' reserves is called the "mortgage origination channel" by Chakraborty, Goldstein, and MacKinlay (2020) and the "liquidity channel" by Rodnyansky and Darmouni (2017).
- <sup>10</sup> This is what Rodnyansky and Darmouni (2017) refer to as the "net worth channel".
- $^{11}$  The Stata command xtdcce2 by Ditzen (2018) was used to obtain the regression results.
- <sup>12</sup> As an incentive for banks to keep reserves, on October 6 2008, the Board of Governors of the Federal Reserve System announced that it will start to pay interest on reserves.

## References

Andrews, D. W. K. 1993. "Tests for Parameter Instability and Structural Change With Unknown Change Point." *Econometrica* 61: 821–856.

Antoch, J., J. Hanousek, L. Horváth, M. Hušková, and S. Wang. 2019. "Structural breaks in Panel Data: Large Number of Panels and Short Length Time Series." *Econometric Reviews* 38: 828–855.

Bai, J. 1997. "Estimation of a Change Point in Multiple Regression Models." *Review of Economics and Statistics* 79: 551–563.

Bai, J. 1999. "Likelihood Ratio Tests for Multiple Structural Changes." *Journal of Econometrics* 91: 299–323.

Bai, J., and P. Perron. 1998. "Estimating and Testing Linear Models With Multiple Structural Changes." *Econometrica* 66: 47–78.

Bai, J., and P. Perron. 2003a. "Critical Values for Multiple Structural Change Tests." *Econometrics Journal* 6: 72–78.

Bai, J., and P. Perron. 2003b. "Computation and Analysis of Multiple Structural Change Models." *Journal of Applied Econometrics* 18: 1–22.

Baltagi, B. H., Q. Feng, and C. Kao. 2016. "Estimation of Heterogeneous Panels With Structural Breaks." *Journal of Econometrics* 191: 176–195.

Baltagi, B. H., C. Kao, and L. Liu. 2017. "Estimation and Identification of Change Points in Panel Models With Nonstationary or Stationary Regressors and Error Term." *Econometric Reviews* 36: 85–102.

Boldea, O., B. Drepper, and Z. Gan. 2020. "Change Point Estimation in Panel Data With Time-Varying Individual Effects." *Journal of Applied Econometrics* 35: 712–727.

Chakraborty, I., I. Goldstein, and A. MacKinlay. 2020. "Monetary Stimulus and Bank Lending." *Journal of Financial Economics* 136: 189–218.

De Vos, I., G. Everaert, and V. Sarafidis. 2024. "A Method to Evaluate the Rank Condition for CCE Estimators." *Econometric Reviews* 43, no. 2-4: 123–155.

Di Maggio, M., A. Kermani, and C. J. Palmer. 2020. "How Quantitative Easing Works: Evidence on the Refinancing Channel." *Review of Economic Studies* 87: 1498–1528.

Ditzen, J. 2018. "xtdcce2: Estimating Dynamic Common Correlated Effects in Stata." *Stata Journal* 18: 585–617.

Ditzen, J., Y. Karavias, and J. Westerlund. 2022. "Testing and Estimating Structural Breaks in Time Series and Panel Data in Stata." arXiv:2110.14550 [econ.EM].

Kaddoura, Y., and J. Westerlund. 2022. "Estimation of Panel Data Models with Random Interactive Effects and Multiple Structural Breaks when *T* is Fixed." *Journal of Business and Economic Statistics*.

Kapoor, S., and O. Peia. 2021. "The Impact of Quantitative Easing on Liquidity Creation." *Journal of Banking and Finance* 122: 105998. Karavias, Y., P. Narayan, and J. Westerlund. 2023. "Structural Breaks in Interactive Effects Panels and the Stock Market Reaction to COVID-19." *Journal of Business & Economic Statistics* 41, no. 3: 653–666.

Kim, D., and W. Sohn. 2017. "The Effect of Bank Capital on Lending: Does Liquidity Matter?." *Journal of Banking & Finance* 77: 95–107.

Li, D., J. Qian, and L. Su. 2016. "Panel Data Models With Interactive Fixed Effects and Multiple Structural Breaks." *Journal of the American Statistical Association* 111: 1804–1819.

Lubik, T. A., and P. Surico. 2010. "The Lucas Critique and the Stability of Empirical Models." *Journal of Applied Econometrics* 25: 177–194.

Luck, S., and T. Zimmermann. 2020. "Employment Effects of Unconventional Monetary Policy: Evidence From QE." *Journal of Financial Economics* 135: 678–703.

Mahajan, A. 2006. "Identification and Estimation of Regression Models With Misclassification." *Econometrica* 74: 631–665.

Moon, H. R., and M. Weidner. 2015. "Linear Regression for Panel With Unknown Number of Factors as Interactive Fixed Effects." *Econometrica* 83: 1543–1579.

Ng, S., and P. Perron. 2001. "LAG Length selection and the construction of Unit Root Tests With Good Size and Power." *Econometrica* 69: 1519–1554.

Pesaran, H. M. 2021. "General Diagnostic Tests for Cross-Sectional Dependence in Panels." *Empirical Economics* 60: 13–50.

Pesaran, M. H. 2006. "Estimation and Inference in Large Heterogeneous Panels With a Multifactor Error Structure." *Econometrica* 74: 967–1012.

Pesaran, M. H. 2007. "A Simple Panel Unit Root Test in the Presence of Cross-Section Dependence." *Journal of Applied Econometrics* 22: 265–312.

Qian, J., and L. Su. 2016. "Shrinkage Estimation of Common Breaks in Panel Data Models Via Adaptive Group Fused Lasso." *Journal of Econometrics* 191: 86–109.

Rodnyansky, A., and O. M. Darmouni. 2017. "The Effects of Quantitative Easing on Bank Lending Behavior." *Review of Financial Studies* 30, no 11: 3858–3887.

Sargan, J. D. 1964. "Wages and Prices in the United Kingdom: A Study in Econometric Methodology." In *Econometric Analysis for National Economic Planning*, edited by P. E. Hart, G. Mills and J. K. Whitaker, 25–54. London: Butterworths.

Stock, J. H., and M. W. Watson. 1996. "Evidence on Structural Stability in Macroeconomic Time Series Relations." *Journal of Business & Economic Statistics* 14: 11–30.

Stock, J. H., and M. W. Watson. 2003. "Forecasting Output and Inflation: The Role of Asset Prices." *Journal of Economic Literature XLI* 41, no. 3: 788–829.

Westerlund, J. 2018. "CCE In Panels With General Unknown Factors." *Econometrics Journal* 21: 264–276.

Westerlund, J., and J.-P. Urbain. 2015. "Cross-Sectional Averages Versus Principal Components." *Journal of Econometrics* 185: 372–377.

Zhu, H., V. Sarafidis, and M. J. Silvapulle. 2020. "A New Structural Break Test for Panels With Common Factors." *Econometrics Journal* 23: 137–155.

# Supporting Information

Additional supporting information can be found online in the Supporting Information section.