Structural Break Tests Robust to Regression Misspecification

Abi Morshed, Alaa; Andreou, E.; Boldea, Otilia

Document version:
Early version, also known as pre-print

Publication date:
2016

CentER

Discussion Paper

No. 2016-019

STRUCTURAL BREAK TESTS ROBUST TO REGRESSION MISSPECIFICATION

By

Alaa Abi Morshed, Elena Andreou,
Otilia Boldea

2 May 2016

ISSN 0924-7815
ISSN 2213-9532

TILBURG UNIVERSITY

Structural break tests robust to regression misspecification*

Alaa Abi Morshed † Elena Andreou ‡ Otilia Boldea §

April 19, 2016

Abstract

Structural break tests developed in the literature for regression models are sensitive to model misspecification. We show - analytically and through simulations - that the sup Wald test for breaks in the conditional mean and variance of a time series process exhibits severe size distortions when the conditional mean dynamics are misspecified. We also show that the sup Wald test for breaks in the unconditional mean and variance does not have the same size distortions, yet benefits from similar power to its conditional counterpart. Hence, we propose using it as an alternative and complementary test for breaks. While the conditional tests based on dynamic regression models detect breaks in the mean and variance of the US unemployment growth and interest rate growth series around the Great Moderation, the evidence for these breaks disappears when using the unconditional tests. Therefore, there is no evidence of long-run mean or volatility shifts in unemployment growth and interest rate growth.

Keywords: structural change, sup Wald test, dynamic misspecification

JEL classification: C01; C12

*W are grateful to the seminar participants at the Tilburg Econometrics Seminars in 2014 and 2015 and to the conference participants at the NESG 2014 and IAAE 2015 for useful comments.

†CentER and Department of Econometrics and Operation Research, Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands. E-mail: A.AbiMorshed@uvt.nl

‡Department of Economics and CEPR, University of Cyprus, P.O. Box 20537, 1678 Nicosia, Cyprus. E-mail: elena.andreou@ucy.ac.cy

§CentER and Department of Econometrics and Operation Research, Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands. E-mail: o.boldea@uvt.nl

1 Introduction

There is a large literature on alternative structural break tests, as well as empirical evidence that many economic indicators went through periods of structural change. Most structural break tests are developed for the slope parameters of a regression model - see inter alia Andrews (1993), Andrews and Ploberger (1994), Ploberger and Kramer (1992), Bai and Perron (1998).

Macroeconomic variables may often exhibit long-run mean shifts, that is, structural breaks in their unconditional mean. Mean shifts in unemployment rates, interest rates, GDP, inflation and other macroeconomic variables may signal permanent changes in the structure of the economy and are therefore themselves of interest to practitioners. Nevertheless, very few papers test for unconditional mean shifts; instead, most of the literature refers to "mean shifts" as breaks in the short-run conditional mean and proceed with the usual regression based tests for break-points - see inter alia Vogelsang (1997,1998), Perron and Yabu (2009) and McKitrick and Vogelsang (2014).

In this paper, we show, analytically and through simulations, that tests for conditional mean breaks are severely oversized when the functional form is misspecified, leading to detection of spurious breaks. Their unconditional counterparts are not plagued by the same issues and we propose using them instead, or in conjunction with, the conditional mean break tests.

The sensitivity of the conditional mean break tests to functional form misspecifications has been documented earlier. Chong (2003) focused on cases with iid, conditionally homoskedastic errors. Bai et al (2008) focused on models with measurement error and proposed a new break-point test that corrects for measurement error. Another strand of literature focuses on trend breaks rather than mean breaks. It shows that dynamic misspecification of the conditional mean may result in severe size distortions and non-monotonic power - see inter alia Vogelsang (1997,1998,1999), Kejriwal (2009), Perron and Yabu (2009) and McKitrick and Vogelsang (2014). The theoretical and simulation results in these papers are backed up by the empirical studies in Altansukh et al. (2012) and Bataa et al. (2013), who expose the undesirable effects of misspecifying the conditional mean seasonalities, outliers, dynamics and heteroskedasticity in practice.

In this paper, we analyze the effects of static and dynamic misspecification on conditional mean break tests. We focus on the sup Wald test of Andrews (1993) because it is widely used in applied work. We prove that its asymptotic distribution is nonstandard and highly data-dependent when the number of lags is underspecified. Our analysis focuses on stationary weakly dependent and heteroskedastic processes, generalizing the iid homoskedastic results derived in Chong (2003).[1]

[1] Nonstationary processes with a trend break and unit root errors, whose first-differences exhibit mean shifts with stationary errors, have been analyzed in many papers. But, as Vogelsang (1998, 1999) shows, to recover monotonic power, testing the first-differenced series for a mean shift is better than testing the level series for a trend shift.

Most of the literature proposed correcting for dynamic misspecification by better lag selection procedures - Vogelsang and Perron (1998), Perron and Yabu (2009) - or by fixed bandwidth asymptotics - Vogelsang (1998), Sayginsoy and Vogelsang (2011), Cho and Vogelsang (2014). In this paper, we propose testing for breaks in the unconditional mean instead. Breaks in the conditional mean are not equivalent, yet closely related, to breaks in the conditional mean, as long as the conditional mean is correctly specified. Aue and Horváth (2012), among others, illustrate tests for both type of breaks in a recent comprehensive review on structural break tests. We run an extensive simulation study to show that for most common static and dynamic misspecifications in the conditional mean, the unconditional mean break test, corrected for autocorrelation, yields correctly-sized or under-sized tests, compared to over-sized conditional mean tests. Moreover, the power of both tests is very similar, especially as the sample size increases.[2] Similar results hold for the unconditional versus conditional break tests in variance. Therefore, the approach of testing first for a break in the unconditional mean and variance of the variable of interest is not only complementary to the regression approach, but is robust to alternative sources of misspecification.

There is a plethora of empirical evidence for breaks in the conditional mean and volatility of many US macroeconomic time series during the early mid 1980s, associated with the Great Moderation (see for example, McConnell and Perez-Quiros (2000), Stock and Watson (2002), Sensier and van Dijk (2004), Bataa et al. (2013)). Most studies employ dynamic regression models to detect such breaks. Focusing on unemployment and interest rates, we show that the unconditional mean and volatility tests detect no breaks. Thus, the breaks detected by the conditional tests are either spurious because of size distortions, or they do not result in long-run breaks in the mean or volatility of unemployment growth or interest rate growth.

This paper is organized as follows: Section 2 defines the unconditional break tests in mean and variance and derives their asymptotic properties in a unified framework. Section 3 defines the conditional structural break tests in mean and variance. It contains asymptotic results for the conditional break tests under correct specification and misspecification. Section 4 presents the simulation evidence comparing the size and power of the conditional and unconditional break tests. Section 5 illustrates the difference between these alternative structural break tests approaches with two empirical applications on the US civilian unemployment rate and the short-term real interest rates. A final section concludes. All the proofs are relegated to the Appendix.

[2]The only case where our test has comparatively low power to the conditional mean test is in a correctly specified dynamic model with an intercept very close to zero. This case is further discussed in the simulation section.

2 Unconditional mean and variance break tests

In this section, we define the unconditional sup Wald test for an unknown break in the mean or variance of a dynamic univariate process.[3]

To our knowledge, a test for an unknown unconditional mean break, adjusted for autocorrelation, is rarely used in the literature.[4] Most papers test for a break in the conditional mean of a series; when they intend to test for an unconditional mean break, they routinely test for a trend break or an intercept break instead, after specifying a conditional mean - see e.g. Stock and Watson (2002). Such approaches have the disadvantage that they are highly dependent on the correct specification of the conditional mean. They also do not shed light on unconditional mean shifts, which may not be equivalent to conditional mean shifts. Therefore, in this paper, we propose using UM break tests complementarily to CM break tests, to uncover long-run mean shifts in the presence of potential static and dynamic misspecification.

We denote the unconditional mean by UM, and the unconditional variance by UV henceforth. In contrast to UM breaks, UV breaks are routinely tested in applications, for example to uncover the Great Moderation break. It is common to test for a break in the absolute value of the demeaned data, as a proxy for testing a variance break - see McConnell and Perez-Quiros (2000), Stock and Watson (2002) and Sensier and van Dijk (2004). We call these tests UA (unconditional absolute deviation) break tests. One can also use the squared demeaned data to test for a variance break, as in Pitarakis (2004) and Qu and Perron (2007). We call these UV break tests, because they test directly for a variance break.[5]

Below, we state the null asymptotic distributions of UM, UA and UV break tests under fairly general assumptions on the data. These distributions are not dependent on regressor, functional form, or seasonality misspecifications, simply because a conditional mean is not specified. The only misspecification that affects the null asymptotic distribution of these tests are UM breaks for the UA and UV tests, or UV breaks for the UM test. Fortunately, this misspecification is easy to correct; we discuss this correction at

[3]Throughout the paper, we use the sup Wald test definition in Andrews (1993); alternative definitions of the sup Wald test are available, but they are not equivalent to the original sup Wald test in Andrews (1993) and should not be confused with it.

[4]Even though UM tests are not routinely used, they are a special case of the HAC-adjusted conditional break-point test in e.g. Bai and Perron (1998), when the only regressor is an intercept. Also, a CUSUM (cumulative sum) variant of this test for iid data is in Pitarakis (2004). As shown in the Appendix, proof of Theorem 1, for unconditional break tests, there is an explicit asymptotic relationship between the CUSUM test and the sup Wald test. However, as the Appendix shows, the conclusion of the two tests based on asymptotic critical values is in general different. Since there is strong evidence for the non-monotonic power of the CUSUM test - see e.g. Vogelsang (1999), the paper focuses on the sup Wald test instead.

[5]Note that a break in the expected absolute value of a demeaned series is not the same as a variance break only under certain conditions.

the end of this section.

The true model takes the general form:

$$y_t = \mu_1 \mathbf{1}[t \leq T_{UM}] + \mu_2 \mathbf{1}[t > T_{UM}] + u_t, \tag{1}$$

where μ_1, μ_2 are deterministic, the break $T_{UM} = [T\lambda_{UM}]$ is an unknown, fixed fraction of the sample $0 < \lambda_{UM} < 1$, and u_t satisfies the assumption below.

Assumption A 1.

(i) $\mathbf{E}(u_t) = 0$ *and* $\mathbf{Var}\left(T^{-1/2}\sum_{t=1}^{[T\lambda]} u_t\right) = \lambda\ \mathbf{v}_u$*;*

(ii) for some $d > 4$, $\sup_t \mathbf{E}|u_t|^d < \infty$ *and* $\{u_t\}$ *is* $\mathcal{L}_2$*-near epoch dependent of size* $c_m = O(m^{-1})$ *on* $\{g_t\}$*, i.e.* $\left\|u_t - \mathbf{E}[u_t|\mathcal{G}_{t-m}^{t+m}]\right\|_2 \leq d_m$ *with* $d_m = O(m^{-1})$ *where* $\mathcal{G}_{t-m}^{t+m} = \sigma(g_{t-m}, \ldots, g_{t+m})$*, and* $\{g_t\}$ *is either* ϕ*-mixing of size* $m^{-d/(2(d-1))}$ *or* α*-mixing of size* $m^{-d/(d-2)}$*.*[6]

With these assumptions, y_t can exhibit very general dependence - ARMA, GARCH, nonlinear dependence - but it cannot have unit roots or UV breaks.

For a UM break, the null and alternative hypotheses are:

$$H_0^{UM} : \mu_1 = \mu_2 \qquad vs. \qquad H_A^{UM} : \mu_1 \neq \mu_2.$$

For a UA break, let $a_t = \mathbf{E}|y_t - \overline{y}|$, and test:

$$H_0^{UA} : a_t = a_u \qquad vs. \qquad H_A^{UA} : a_t = a_{u,1}\ \mathbf{1}[t \leq T_{UA}] + a_{u,2}\ \mathbf{1}[t > T_{UA}], a_{u,1} \neq a_{u,2}.$$

For a UV break, let $v_{ut} = \mathbf{E}(y_t - \overline{y})^2$, and test:

$$H_0^{UV} : v_t = v_u \qquad vs. \qquad H_A^{UV} : v_t = v_{u,1}\ \mathbf{1}[t \leq T_{UV}] + v_{u,2}\ \mathbf{1}[t > T_{UV}], v_{u,1} \neq v_{u,2}.$$

Under the alternative hypotheses, all breaks $T_{UM} = [T\lambda_{UM}]$, $T_{UV} = [T\lambda_{UV}]$, $T_{UA} = [T\lambda_{UA}]$ are assumed to occur at unknown fixed fractions $0 < \lambda_{UM}, \lambda_{UV}, \lambda_{UA} < 1$ of the sample.

The UM test is defined below. It is a special case of the Andrews (1993) sup Wald test when the only regressor is an intercept, and when the variance is estimated under the null of no break. Therefore, it is not new; nevertheless, to our knowledge, it is rarely used in the empirical literature in the form defined below:

$$UM_T^* = \sup_{\lambda \in [\epsilon, 1-\epsilon]} UM_T(\lambda), \qquad UM_T(\lambda) = T(\overline{y_{1\lambda}} - \overline{y_{2\lambda}})^2 / \widehat{\mathbf{v}}_{u\lambda}, \tag{2}$$

$\epsilon > 0$ is a small cut-off, typically $\epsilon = 0.15$ in applications, $\overline{y_{i\lambda}} = T_{i\lambda}^{-1}\sum_{i\lambda} y_t$ for $i = 1, 2$, $T_{1\lambda} = [T\lambda]$, $T_{2\lambda} = T - [T\lambda]$, $\sum_{1\lambda} = \sum_{t=1}^{T_{1\lambda}}$, $\sum_{2\lambda} = \sum_{t=T_{2\lambda}+1}^{T}$ and $\widehat{\mathbf{v}}_{u\lambda}$ is a HAC consistent estimator of $\mathbf{v}_{u\lambda} = \mathbf{AVar}(\sqrt{T}(\overline{y_{1\lambda}} - \overline{y_{2\lambda}})) = \mathbf{v}_u / [\lambda(1-\lambda)]$ under H_0^{UM}.

[6] Here, $\|\cdot\|_2 = (\mathbf{E}\|\cdot\|^2)^{1/2}$ stands for the $\mathcal{L}_2$-norm, and $|\cdot|$ stands for the Euclidean norm.

For the HAC estimator $\widehat{\mathbf{v}}_{u\lambda}$, it is crucial to calculate it over the full sample, i.e. under the null H_0^{UM}. If we use sub-sample estimators in its computation - i.e. we estimate the variances $T^{1/2}(\overline{y_{1\lambda}} - \mu)$ and $T^{1/2}(\overline{y_{2\lambda}} - \mu)$ separately - we need a separate bandwidth for each. Since the bandwidth estimation is only accurate in large samples, for those λ's that are close to ϵ and $1-\epsilon$, such an estimation would be highly inaccurate, resulting in high size distortions.[7]

Thus, we define:

$$\widehat{\mathbf{v}}_{u\lambda} = \frac{1}{\lambda(1-\lambda)} \sum_{j=-b_T+1}^{b_T-1} k\left(\frac{j}{b_T}\right)\widehat{\gamma}_j, \qquad \widehat{\gamma}_j = \begin{cases} \frac{1}{T-1}\sum_{t=j+1}^{T}(y_t-\overline{y})(y_{t-j}-\overline{y}), & j \geq 0 \\ \widehat{\gamma}_{-j} & j < 0 \end{cases},$$

where $\overline{y} = T^{-1}\sum_{t=1}^{T} y_t$, $k(x) = (1-|x|)\,\mathbf{1}[|x| \leq 1]$ is throughout the paper the Bartlett kernel, with the optimal data-dependent bandwidth in Newey and West (1994).[8] Specifically, we let $b_T = \min[T, \widehat{\eta}\ T^{\frac{1}{3}}]$, where $\widehat{\eta} = 1.1447(\widehat{f}^{(1)}/\widehat{f}^{(0)})^{\frac{2}{3}}$, and $\widehat{f}^{(1)} = 2\sum_{j=1}^{\tau} j\widehat{\gamma}_j$, $\widehat{f}^{(0)} = \widehat{\gamma}_0 + 2\sum_{j=1}^{\tau}\widehat{\gamma}_j$, with $\tau = [(T/100)^{2/9}]$. The lag truncation parameter τ governs how many auto-covariances should be used in forming the nonparametric estimates $\widehat{f}^{(1)}$ and $\widehat{f}^{(0)}$, which estimate the spectral density at frequency one and zero.[9] Therefore, $\widehat{f}^{(1)}$, $\widehat{f}^{(0)}$, and $\widehat{\eta}$ are computed over the full sample.

The UA and UV tests are denoted by UA_T^* and UV_T^*. They are computed as UM_T^*, but with y_t replaced by $\widehat{a}_t = |y_t - \overline{y}|$ for UA, $\widehat{v}_t = (y_t - \overline{y})^2$ for UV, and $\widehat{\mathbf{v}}_{u\lambda}$ replaced by the HAC consistent estimator of the asymptotic variance of $\widehat{a}_t$ or $\widehat{v}_t$.

Define the distribution:

$$\mathcal{G}_p = \sup_{\lambda\in[\epsilon,1-\epsilon]} \frac{[B_p(\lambda) - \lambda B_p(1)]'[B_p(\lambda) - \lambda B_p(1)]}{\lambda(1-\lambda)},$$

where $B_p(\cdot)$ is a $p \times 1$ vector of independent standard Brownian motions, for some $p \geq 1$. As Theorem 1 shows, $\mathcal{G}_1$ is the null asymptotic distributions of the UM, UA and UV break tests. Although the distribution of various break point tests under different (more restrictive) assumptions is available, an explicit proof for the UM, UV and UA tests under A1 is not available in a unified setting to our knowledge, and we provide it in the Appendix.

Theorem 1.

Let the model be as in (1), *and let A1 hold. Then: (i) under* H_0^{UM}, $UM_T^* \Rightarrow \mathcal{G}_1$; *(ii) under* H_0^{UM} *and* H_0^{UA}, $UA_T^* \Rightarrow \mathcal{G}_1$; *(iii) under* H_0^{UM} *and* H_0^{UV}, $UV_T^* \Rightarrow \mathcal{G}_1$.[10]

Note that the distributions are non-standard, but critical values are available in e.g. Andrews (1993) and Bai and Perron (1998).

[7] Simulation evidence for this statement is available from the authors upon request.

[8] Additional simulations not reported here show that the fixed optimal bandwidth proposed in Andrews (1991) leads to worse performance of the UM break test.

[9] The weights mentioned in Newey and West (1994) are set equal to one as usual for scalar cases.

[10] Here, "$\Rightarrow$" indicates weak convergence in the Skorohod metric.

Theorem 1 assumes no UV breaks for UM tests, via imposing A1(i), and no UM breaks for UV/UA tests, via imposing H_0^{UM}. If there is a UM break (H_A^{UM} holds instead of H_0^{UM}), as shown in Pitarakis (2004), we can obtain $\widehat{v}_t$ and $\widehat{a}_t$ via subsample demeaning, and Theorem 1(ii)-(iii) will hold. That is, we let $\widehat{v}_t = (y_t - \overline{y}_t)^2$, $\widehat{a}_t = |y_t - \overline{y}_t|$, and $\overline{y}_t = \widehat{T}_{UM}^{-1} \sum_{t=1}^{\widehat{T}_{UM}} y_t \mathbf{1}[t \leq \widehat{T}_{UM}] + (T - \widehat{T}_{UM})^{-1} \sum_{t=\widehat{T}_{UM}+1}^{T} y_t \mathbf{1}[t > \widehat{T}_{UM}]$, where $\widehat{T}_{UM}$ is the Bai and Perron (1998) OLS break-point estimator of T_{UM} in (1). If there is a UV break, the asymptotic distribution of the UM test is affected, but one can employ the fixed-regressor bootstrap in Hansen (2000) to correct for this. The correction for the UV tests via sub-sample demeaning is necessary and employed in our empirical analysis in Section 5.

3 Conditional mean and variance break tests

3.1 Correct specification

Unlike unconditional break tests, regression-based break tests are pervasive in empirical work, despite their sensitivity to misspecification (this sensitivity is discussed in Section 3.2). The most common regression specification is of the linear form:

$$y_t = x_t'\theta_1 \mathbf{1}[t \leq T_{CM}] + x_t'\theta_2 \mathbf{1}[t > T_{CM}] + \epsilon_t, \tag{3}$$

where $T_{CM} = [T\lambda_{CM}]$, $0 < \lambda_{CM} < 1$, x_t is a $p \times 1$ vector of regressors that includes an intercept and possibly lagged dependent variables, and ϵ_t are scalar errors.

We denote by CM, CA and CV the conditional mean, conditional absolute deviation and the conditional variance, where the word "conditional" simply refers to specifying the conditional mean in (3). To derive the asymptotic distribution of the CM, CA and CV break tests, we need additional assumptions on the joint dependence of regressors and errors.

Assumption A 2.

(i) $\mathbf{E}(x_t\epsilon_t) = 0$, $\mathbf{AVar}(T^{-1/2}\sum_{t=1}^{[T\lambda]} x_t\epsilon_t) = \lambda\mathbf{V}$ *and* $T^{-1}\sum_{t=1}^{[T\lambda]} x_t x_t' \xrightarrow{p} \lambda Q$, *two positive definite (pd) matrices of constants;*

(ii) for some $d > 4$, $\sup_t \|x_t\epsilon_t\|_d < \infty$ *and* $\{x_t\epsilon_t\}$ *is* $\mathcal{L}_2$*-near epoch dependent of size* $c_m = O(m^{-1})$ *on* $\{h_t\}$, *and* $\{h_t\}$ *is either* ϕ*-mixing of size* $m^{-d/(2(d-1))}$ *or* α*-mixing of size* $m^{-d/(d-2)}$.

The null and alternative hypotheses of the conditional tests are:

$$\begin{aligned}
&H_0^{CM}: \theta_1 = \theta_2 && vs. && H_A^{CM}: \theta_1 \neq \theta_2,\\
&H_0^{CA}: a_{\epsilon t} = a_\epsilon && vs. && H_A^{CA}: a_{\epsilon t} = a_{\epsilon 1}\ \mathbf{1}[t \leq T_{CA}] + a_{\epsilon 2}\ \mathbf{1}[t > T_{CA}],\ a_{\epsilon 1} \neq a_{\epsilon 2},\\
&H_0^{CV}: v_{\epsilon t} = v_\epsilon && vs. && H_A^{CV}: v_{\epsilon t} = v_{\epsilon 1}\ \mathbf{1}[t \leq T_{CV}] + v_{\epsilon 2}\ \mathbf{1}[t > T_{CV}],\ v_{\epsilon 1} \neq v_{\epsilon 2},
\end{aligned}$$

where $a_{\epsilon t} = \mathbf{E}|\epsilon_t|$, $v_{\epsilon t} = \mathbf{Var}(\epsilon_t)$, $T_{CA} = [T\lambda_{CA}], T_{CV} = [T\lambda_{CV}]$, $0 < \lambda_{CA}, \lambda_{CV} < 1$.

The corresponding sup Wald test for a CM break is defined in e.g. Andrews (1993):

$$CM_T^* = \sup_{\lambda \in [\epsilon, 1-\epsilon]} CM_T(\lambda), \qquad CM_T(\lambda) = T(\widehat{\theta}_{1\lambda} - \widehat{\theta}_{2\lambda})' \widehat{\mathbf{V}}_\lambda^{-1} (\widehat{\theta}_{1\lambda} - \widehat{\theta}_{2\lambda}),$$

where $\widehat{\theta}_{1\lambda}, \widehat{\theta}_{2\lambda}$ are the OLS estimators of θ in (3) in subsamples $\{1, \ldots, T_{1\lambda}\}$ and $\{T_{1\lambda} + 1, \ldots, T\}$, and $\widehat{\mathbf{V}}_\lambda$ is a consistent estimator of $\mathbf{AVar}(T^{1/2}(\widehat{\theta}_{1\lambda} - \widehat{\theta}_{2\lambda}))$ under H_0^{CM}.

For the conditional test, the asymptotic variance $\mathbf{AVar}(T^{1/2}(\widehat{\theta}_{1\lambda} - \widehat{\theta}_{2\lambda}))$ is routinely estimated over subsamples - i.e. separately for $T^{1/2}(\widehat{\theta}_{1\lambda} - \theta^0)$ and $T^{1/2}(\widehat{\theta}_{2\lambda} - \theta^0)$, or under the alternative. If a HAC estimator under the alternative would be used, the same problems would arise as for the unconditional test: there would be size distortions due to inaccurate bandwidth estimation for λ close to the beginning or the end of the sample. However, in most studies, the conditional mean specification in (3) is assumed to be correct, in which case all lags of the dependent variable are included as regressors, and correcting for autocorrelation is no longer necessary. If this is the case, the variance can be estimated under the alternative without further size distortions. Thus, as in most empirical studies, we use variance estimators that are not autocorrelation-robust in all the simulations except those where the model is static. In a static model, the researcher might suspect that the errors are autocorrelated, and a HAC estimator is justified.

For the theory section, we consider two potential estimators for $\mathbf{AVar}(T^{1/2}(\widehat{\theta}_{1\lambda} - \widehat{\theta}_{2\lambda}))$, under homoskedasticity or heteroskedasticity. The one under homoskedasticity is:

$$\begin{aligned} &\widehat{\mathbf{V}}_\lambda = \widehat{\mathbf{V}}_{1\lambda} + \widehat{\mathbf{V}}_{2\lambda}, \widehat{\mathbf{V}}_{i\lambda} = \widehat{\mathbf{v}}_{\epsilon,i\lambda} \left(T^{-1} \textstyle\sum_{i\lambda} x_t x_t'\right)^{-1}, \widehat{\mathbf{v}}_{\epsilon,i\lambda} = T_{i\lambda}^{-1} \textstyle\sum_{i\lambda} \widehat{\epsilon}_t^2, (i = 1, 2), \\ &\widehat{\epsilon}_t = y_t - x_t'\widehat{\theta}_{1\lambda} \mathbf{1}\{t \leq T_{1\lambda}\} - x_t'\widehat{\theta}_{2\lambda} \mathbf{1}\{t > T_{1\lambda}\}. \end{aligned} \tag{4}$$

Under heteroskedasticity,

$$\widehat{\mathbf{V}}_\lambda = \widehat{\mathbf{V}}_{1\lambda} + \widehat{\mathbf{V}}_{2\lambda}, \ \widehat{\mathbf{V}}_{i\lambda} = \left(T^{-1} \textstyle\sum_{i\lambda} x_t x_t'\right)^{-1} \left(T^{-1} \textstyle\sum_{i\lambda} \widehat{\epsilon}_t^2 x_t x_t'\right) \left(T^{-1} \textstyle\sum_{i\lambda} x_t x_t'\right)^{-1}, (i = 1, 2). \tag{5}$$

We define the CA and CV tests as the UA and the UV tests, but with $\widehat{a}_t, \widehat{v}_t$ replaced by $\widehat{a}_{\epsilon t} = |\widehat{\epsilon}_t|, \widehat{v}_{\epsilon t} = \widehat{\epsilon}_t^2$, and with $\widehat{\epsilon}_t = y_t - x_t'\widehat{\theta}$ the residuals from estimating (3) under the null H_0^{CM}. We emphasize that the name "conditional" refers exclusively to pre-specifying the conditional mean in (3), and not the conditional variance of y_t or ϵ_t. Therefore, the tests in this paper should not be confused with the conditional variance tests proposed in e.g. Andreou and Ghysels (2002), who write down a model for the conditional variance of ϵ_t.

Theorem 2 states the asymptotic distribution of the CM, CA and CV break tests.

Theorem 2.

Let the model be as in (3)*, and let A2 hold. Then: (i) under* H_0^{CM}, $CM_T^* \Rightarrow \mathcal{G}_p$*; (ii) under* H_0^{CM} *and* H_0^{CA}, $CA_T^* \Rightarrow \mathcal{G}_1$*; (iii) under* H_0^{CM} *and* H_0^{CV}, $CV_T^* \Rightarrow \mathcal{G}_1$.

Note that the distributions are similar to the unconditional break tests, but there are more degrees of freedom used up by the conditional break tests.

As for the UA and UV tests, the asymptotic distributions of the CA and CV tests are not valid if there is a CM break; in that case, as Pitarakis (2004) shows, the CM break at T_{CM} can be pre-estimated by $\widehat{T}_{CM}$ along with the slope parameters $\widehat{\theta}_1, \widehat{\theta}_2$ before and after the break, via the methods in Bai and Perron (1998). Then we can redefine $\widehat{\epsilon}_t = y_t - x_t'\widehat{\theta}_1 \mathbf{1}[t \leq \widehat{T}_{CM}] - x_t'\widehat{\theta}_2 \mathbf{1}[t > \widehat{T}_{CM}]$ in the computation of $\widehat{a}_{\epsilon t}, \widehat{v}_{\epsilon t}$, obtaining the same asymptotic distributions as stated in Theorem 2. Under the alternative H_A^{CV}, the asymptotic null distribution of the CM test is not valid, but as for the unconditional break tests, it can be bootstrapped via the fixed regressor bootstrap in Hansen (2000).

3.2 Dynamic Misspecification

Unlike the unconditional break tests, all the conditional break tests are highly dependent on the correct specification of the functional form, including seasonality and dynamics. Bataa et al. (2013) and Altansukh et al. (2012) empirically show the effects of misspecifying the conditional mean seasonalities, outliers, dynamics and heteroskedasticity on the conditional break tests. Chong (2003) and Bai et al (2008) theoretically show that misspecification of the functional form leads to different null asymptotic distributions for the CM break tests. They focus on iid errors and static misspecifications, although some of their theoretical results apply to dynamic misspecification as well. The impact of dynamic misspecification of (3) on conditional break tests has been analyzed by Vogelsang and Perron (1998), Vogelsang (1999), Perron and Yabu (2009), inter alia. But all these studies correct for omitted autocorrelation in the errors by either better selection of lags in the regression equation, or directly correcting the error variance via HAC estimators. The first correction is successful if the method used indeed selects the number of lags correctly. The second correction is not always valid if the regression model is already dynamic, as omitted autocorrelation in the errors often violates the exogeneity assumption A2(i), so a HAC variance estimator does not fix the dynamic misspecification problems.

To our knowledge, the effect of misspecifying the regressors or number of lags on CM break tests has not been studied before under general dependence and conditionally heteroskedastic data as allowed for in A3.[11] We prove below that the asymptotic distribution of the CM break test is data-dependent and different than that stated in Theorem 2. Therefore, in the presence of dynamic misspecification, the critical values of the CM tests will be incorrect[12], while the critical values for the UM break test are correct. So the UM break test provides a valuable tool for assessing stability of the process y_t in the presence of dynamic misspecification.

To formalize the results under dynamic misspecification, let $x_t = \mathbf{vec}(x_{t(1)}, x_{t(2)})$ and

[11] The result in Theorem 3 has to our knowledge only been derived by Chong (2003) for iid, conditionally homoskedastic data.

[12] The simulation section shows that the CM tests are severely oversized with dynamic misspecification.

$\theta = \mathbf{vec}(\theta_{(1)}, \theta_{(2)})$, where $x_{t(1)}, \theta_{(1)}$ are $p_1 \times 1$, $x_{t(2)}, \theta_{(2)}$ is $p_2 \times 1$, and $p_1 + p_2 = p$.[13] The true model is (3), but we mistakenly regress y_t only on $x_{t(1)}$ (which we assume includes the intercept). Thus, we underspecify the number of regressors; in particular, we are interested in the effects of underspecifying the number of lags.[14] Partition $Q = \begin{bmatrix} Q_{(1)} & Q_{(12)} \\ Q'_{(12)} & Q_{(2)} \end{bmatrix}$, with $Q_{(1)}, Q_{(2)}$ of dimensions $p_1 \times p_1$ and $p_2 \times p_2$, respectively.

Assumption A 3. *Let $k_t = x_{t(1)}\epsilon_t$, $L_t = x_{t(1)}x'_{t(2)} - Q_{(12)}$, $M_t = x_{t(1)}x'_{t(1)} - Q_{(1)}$, and $w_t = \mathbf{vech}(k_t, L_t, M_t)$, where $\mathbf{vech}(A, B)$ selects, in order, the unique elements and the first occurrence of the repeating elements in $\boldsymbol{vec}(A, B)$.*

(i) $\mathbf{E}(w_t) = 0$, $\mathbf{AVar}(T^{-1/2}\sum_{t=1}^{[T\lambda]} w_t) = \lambda \mathbf{H}$, *a pd matrix of constants;*

(ii) for some $d > 4$, $\sup_t \|w_t\|_d < \infty$ and $\{w_t\}$ is $\mathcal{L}_2$-near epoch dependent of size $d_m = O(m^{-1})$ on either an ϕ-mixing process of size $m^{-d/(2(d-1))}$ or an α-mixing process of size $m^{-d/(d-2)}$.

(iii) $Q_{(12)} \neq O_{p_1 \times p_2}$, where $O_{p_1 \times p_2}$ is the $p_1 \times p_2$ null matrix;

(iv) $T^{-1}\sum_{1\lambda} w_t w'_t \xrightarrow{p} \lambda\Omega$, a pd matrix of constants.

A3(iii) states that the omitted regressors are correlated with the included regressors, as is the case when the number of lags is underspecified. The rest of the statements in A3 are standard. Let $r = p_1(1 + p_2 + (p_1 + 1)/2)$, the dimension of w_t. Then, under A3, the functional central limit theorem in Wooldridge and White (1988, Theorem 2.11) can be applied to yield $T^{-1/2}\sum_{1\lambda} w_t \Rightarrow H^{1/2}B_r(\lambda)$. To state the asymptotic distribution of the CM break test under misspecification, let $\mathcal{B}_r(\lambda) = B_r(\lambda) - \lambda B_r(1)$, $s = p_1(1 + p_1 + p_2)$, and $\mathcal{B}^*_s(\lambda) = B^*_s(\lambda) - \lambda B^*_s(1)$, where $B^*_s(\lambda)$ is constructed from $B_r(\lambda)$ by repeating its elements exactly in the positions where $w^*_t = \mathbf{vec}(k_t, L_t, M_t)$ repeats the elements of $w_t = \mathbf{vech}(k_t, L_t, M_t)$. Similarly, let $\mathbf{H}^{*1/2}$ and Ω^* be positive semidefinite matrices constructed from $H^{1/2}$ and Ω - which were defined in A3 - so that $\mathbf{AVar}(T^{-1/2}\sum_{t=1}^{[T\lambda]} w^*_t) = \lambda \mathbf{H}^{*1/2}\mathbf{H}^{*1/2\prime}$, and $T^{-1}\sum_{1\lambda} w^*_t w^{*\prime}_t \xrightarrow{p} \lambda\Omega^*$. With this notation, the asymptotic distribution of the CM test is stated in Theorem 3.

Theorem 3. *Let A2-A3 and H_0^{CM} hold, $\delta = Q_{(1)}^{-1}Q_{(12)}\theta_{(2)}$, $\xi = \boldsymbol{vec}(1, \theta_{(2)}, -\delta)$ and $A = \mathbf{H}^{*1/2\prime}[Q_{(1)}^{-1} \otimes \xi]\left\{[Q_{(1)}^{-1} \otimes \xi']\Omega^*[Q_{(1)}^{-1} \otimes \xi]\right\}^{-1}[Q_{(1)}^{-1} \otimes \xi']\mathbf{H}^{1/2}$.*
(i) If CM^*_T is constructed under heteroskedasticity,

$$CM^*_T \Rightarrow \sup_{\lambda} \frac{\mathcal{B}'_{p_1(p_2+1)}(\lambda)\ A\ \mathcal{B}_{p_1(p_2+1)}(\lambda)}{\lambda(1-\lambda)}.$$

[13]We extend the $\mathbf{vec}(A, B)$ notation to denote stacking in a vector all columns of A, then all columns of B, one by one, in order, even when A, B do not have the same number of rows, and we let $\mathbf{vec'}(A, B) = [\mathbf{vec}(A, B)]'$.

[14]Overspecifying the number of lags or regressors is not a problem, as the coefficients on the additional regressors or lags will converge to zero.

(ii) Let $\nu = \sigma_\epsilon^2 + \theta'_{(2)}[Q_{(2)} - Q'_{(12)}Q_{(1)}^{-1}Q_{(12)}]\theta_{(2)}$. If CM_T^* is constructed under homoskedasticity, then the result in (i) holds, with $A = \nu^{-1}\mathbf{H}^{*1/2\prime}[Q_{(1)}^{-1} \otimes (\xi\xi')]\mathbf{H}^{*1/2}$.

The theorem above shows that the asymptotic distribution of the CM test is nonstandard and highly dependent on the data and the unknown number of lags omitted. Thus, under dynamic misspecification (with $Q_{(12)} \neq 0$), the usual critical values from Theorem 2 no longer apply. Note that Theorem 3 in Chong (2003) is a special case of our result, when the errors ϵ_t are iid and conditionally homoskedastic, and the variance estimation is done under homoskedasticity. Allowing for conditional heteroskedasticity, our Theorem 3 demonstrates that the size distortions of the CM test are dependent on several parameters of the data generating process, and that correcting for heteroskedasticity does not help in overcoming this problem.

4 Simulation results

The objective of the simulation analysis is to compare the size and power of unconditional moments, UM/UV, break tests to their conditional moments, CM/CV, counterparts, under correct regression model specification, and under static and dynamic misspecification. We evaluate the size and power of the tests for alternative model specifications, sample sizes, as well as structural break sources and sizes.[15] We consider sample sizes $T = 100, 200, 500, 1000$ with a break in the middle of the sample, $T_0 = [0.5T]$ and four data generating processes (DGPs). We also considered alternative break points and our results are robust to $T_0 = [0.25T]$ and $T_0 = [0.75T]$. For all simulations, we use the critical values reported in Andrews (2003). For DGPs with static errors, we calculate CM_T^* with $\hat{\mathbf{V}}_\lambda$ as described in (4). For DGPs with AR(1) errors, the CM_T^* test employs the Newey-West HAC estimator for $\hat{\mathbf{V}}_\lambda$.

The results are organized as follows: the sizes of all tests are reported in Tables and the size-adjusted powers in Figures. Tables 1,3 and Figures 1-3,7 are for correctly specified models, and Tables 2,4 and Figures 4-6,8-9 are for misspecified models. Tables 1-2 and Figures 1-6 refer to mean tests, and Tables 3-4 and Figures 7-9 refer to variance tests.

We consider four DGPs, some of which we analyze under both correct specification and misspecification. The first DGP is a simple AR(1) model with iid errors:

$$DGP1 : y_t = \alpha_t + \beta_t y_{t-1} + \epsilon_t, \epsilon_t \sim iid\ \mathcal{N}(0,1), (t = 1, \dots, T).$$

All simulations are performed in Matlab for 10000 replications and for the AR models we use zero as the starting value and 100 burn-in observations. Under the null, $\alpha_t = \alpha = 1$,

[15] The unconditional mean and variance sup Wald tests require a long-run variance estimator. We report the Newey and West (1994) HAC estimator with the data dependent bandwidth therein and the Bartlett kernel, as explained in detail in Section 2. The Andrews (1991) fixed bandwidth HAC estimator leads to slightly worse performance across all tests and designs; results are available upon request from the authors.

and the persistence ranges from $\beta_t = \beta \in [0.1, 0.7]$. Under the alternative, there is one break either in the intercept, with $\alpha_t = 1\ \mathbf{1}_{t\leq T_0} + \delta_\alpha\ \mathbf{1}_{t>T_0}$, and $\delta_\alpha \in (0, 2]$, or in the slope, with $\beta_t = 0.1\ \mathbf{1}_{t\leq T_0} + \delta_\beta\ \mathbf{1}_{t>T_0}$, and $\delta_\beta \in (0, 0.6]$.

For DGP1, we estimate only the correctly specified dynamic model. The size of the CM and UM break tests (under the null) are reported in the top panel of Table 1. Using the 5% critical values we find that the UM test exhibits slightly better size for small sample sizes of $T = 100$ relative to the CM test which yields size around 10%. For large sample sizes of $T > 500$, both tests approach the nominal level, as expected. Under the alternative, we plot the size-adjusted power functions in Figure 1. When the break occurs in the slope parameter, the UM and CM tests have similar power as the sample size grows. The CM test performs only mildly better for moderate changes in the AR slope parameter (with maximum relative gains in power of 10% for $T = 100$). On the other hand, when the break is in the intercept, the UM test has better power in small sample sizes (of $T = 100, 200$), with up to 20% gains vis-a-vis the CM test.[16]

The second DGP is an AR(4) model with iid errors:

$$DGP2 : y_t = \alpha_t + \beta_t'\mathbf{vec}(y_{t-1}, \ldots, y_{t-4}) + \epsilon_t,\ \ \epsilon_t \sim iid\ \mathcal{N}(0, 1), (t = 1, \ldots, T).$$

We set $\beta_t' = (\beta_{t,1}, 0.2, 0.15, 0.075)$ to represent the memory decaying pattern encountered in many time series in economics. Under the null, we set $\alpha_t = \alpha = 1$ and vary $\beta_{t,1} = \beta \in [0.1, 0.3]$. Under the alternative, there is one break either in the intercept, given by $\alpha_t = 1\ \mathbf{1}_{t\leq T_0} + \delta_\alpha\ \mathbf{1}_{t>T_0}$, and $\delta_\alpha \in (0, 2]$, or in the slope, given by $\beta_{t,1} = 0.1\ \mathbf{1}_{t\leq T_0} + \delta_\beta\ \mathbf{1}_{t>T_0}$, and $\delta_\beta \in (0, 0.2]$.

For DGP2, we only analyze the impact of dynamic misspecification: the true DGP is an AR(4) model, but we estimate an $AR(1)$ or an $AR(2)$ model instead. The top two panels of Table 2 show that underestimating the number of lags causes severe size distortions of the CM test, of up to 60% even for small levels of forgone persistence. This effect does not die out even for large sample sizes of $T = 1000$. In contrast, the UM test is not so severely oversized especially for large samples; the size distortions reach a maximum of 13% for large samples of $T = 1000$. The reason UM behaves better in terms of size under dynamic misspecification is due to the HAC correction. What is interesting though from our simulation results is that although the HAC estimator may be less reliable in small samples, under misspecification the size of the UM test is relatively better than that of the CM test for small samples. Under the alternative, the size-adjusted power curves of the UM and CM test with dynamic misspecification are reported in Figure 4. They show that underestimating the AR dynamics does not affect the size-adjusted power of either the CM or UM tests for any sample size or break magnitude.

[16]Note that for DGP1, the UM of y_t is equal to $\alpha_t/(1 - \beta_t)$. If $\alpha_t/(1 - \beta_t)$ is close to zero regardless of t, the UM test will, by design, have little power for a break in the slope β_t. Therefore, if a slope break is the only break of interest, it should be tested directly via the CM test for partial structural change in slopes.

The third DGP is a static model with iid errors:

$$DGP3: y_t = \alpha_t + \beta_t X_t + \epsilon_t,\ \epsilon_t \sim iid\, \mathcal{N}(0,1), X_t \sim iid\, \mathcal{N}(1,1), X_t \perp \epsilon_s, (t,s = 1,\dots T).$$

Under the null, we set $\alpha_t = \alpha = 1$ and vary $\beta_t = \beta \in [0.1, 0.9]$. Under the alternative, there is one break either in the intercept, where $\alpha_t = 1\ \mathbf{1}_{t \leq T_0} + \delta_\alpha\ \mathbf{1}_{t > T_0}$, and $\delta_\alpha \in (0,2]$, or in the slope, where $\beta_{t,1} = 0.1\ \mathbf{1}_{t \leq T_0} + \delta_\beta\ \mathbf{1}_{t > T_0}$, and $\delta_\beta \in (0, 0.9]$.

For DGP3, we analyze both correctly specified and misspecified models. As expected, if we estimate the correctly specified static model in DGP3, the size of both the CM and UM break tests is close to the nominal size, as shown by the second panel in Table 1. The corresponding size-adjusted power curves in Figure 2 are again similar for the two tests, especially as T increases.

However, if instead we estimate an AR(1) model, the results in the third panel of Table 2 show that the UM test is undersized for small sample sizes and that its size improves for $T > 500$. In contrast, the CM test is oversized for small samples. As for the power, the two tests have similar power for large samples, as shown in Figure 5. For smaller samples, misspecifying the regressors compromises the power of the CM test which can be up to around 20% smaller than that of the UM test when $T = 100$.

The fourth DGP is a static model with AR(1) errors:

$$DGP4: y_t = \alpha_t + \beta_t X_t + \epsilon_t,\ \epsilon_t = 0.6\epsilon_{t-1} + \nu_t,\ \nu_t \sim iid\, \mathcal{N}(0,1), (t = 1, \dots, T).$$

For comparison purposes, in DGP4 the X_t and the null and alternatives are generated following DGP3. For DGP4, we analyze both correctly specified and misspecified models. Under correct specification, the last panel of Table 1 shows that the UM test is correctly sized for all sample sizes, whereas the CM test is oversized even for large sample sizes. The size of the CM tests can reach up to 10% even when $T = 1000$ (and the nominal size is 5%).

Furthermore, we consider a nonlinear misspecification by estimating the model with X_t^2 instead of X_t, similar to Chong (2003). The nonlinear misspecification yields oversized CM tests across all sample sizes. The last panel of Table 2 shows that, even for $T = 1000$, the traditional CM test yields size of around 13%. In addition, the size-adjusted power of the CM test is lower than that of the UM test. Figure 6, show that the UM test has relatively better power especially for small sample sizes of $T < 200$. When $T = 100$ the UM test has an increasing power function which is almost 70% higher than that of the CM test. For large samples of $T = 1000$, the size-adjusted power curves of the two tests are very similar.

We now turn to examine the size and power of tests for breaks in the variance of the residuals of the regression models by comparing the UV and CV tests.[17] We consider the same DGPs as before, but we set $\alpha_t = 1$, $\beta_t = 0.5$. For DGP 1-3, we let $\epsilon_t \sim iid\, \mathcal{N}(0, \sigma_t)$, and for DGP4, we let $\nu_t \sim iid\, \mathcal{N}(0, \sigma_t)$. Under the null hypotheses, we fix $\sigma_t = \sigma \in [1, 2.6]$.

[17]The results are very similar for the UA and CA tests and they are available upon request.

Under the alternative, we set $\sigma_t = 1\ \mathbf{1}[t \leq T_0] + \delta_\sigma \mathbf{1}[t > T_0]$, and let $\delta_\sigma \in (0, 1.6]$. As before, we estimate both correctly specified and misspecified models.

When the estimated model is correctly specified, as considered in DGP1 and DGP4, the size of both CV and UV tests are close to the nominal size for $T \geqslant 200$, as shown in Table 3. However, the powers of these two tests differ. Figure 7 shows that for DGP1, the CV test has better power for small sample sizes across all break sizes, including small breaks, as T increases. For DGP4, Figure 7 shows that the power curves of the CV tests and UV tests are the same.

If instead, a misspecified model is estimated for DGP1-DGP4, the CV test appears to enjoy good size properties, shown in Table 4. The exception is the oversizing reported in the top panel of Table 4, which is due to underestimating the lag order; in this case the size does not improve as the sample increases. Our analysis shows that misspecifying the dynamics of the conditional mean of the regression model yields an oversized CV test. For the power, the results are less clear-cut: the power of the CV test seems larger than that of the UV test in the presence of dynamic misspecification (Figure 8), but smaller under nonlinear misspecification (Figure 9). Nevertheless, this difference disappears in large samples.

Summarizing, the simulation results show that under correct model specification, the UM/UV and CM/CV have similar size and power. In contrast, under static nonlinear and dynamic misspecifications, the CM/CV tests are severely oversized, having both finite and large sample distortions. While the UM/UV tests may also occasionally exhibit mild size distortions, they feature similar power properties as the CM/CV tests, especially in larger samples. Therefore, the UM/UV tests can be a valuable tool for detecting breaks, because in applied work misspecification is likely to occur and bias the CM/CV break test results.[18]

[18]Other types of model misspecifications may also affect the size and power of the **(CM and CV)** structural break tests. Analyzing them is beyond the scope of this paper, but further results regarding these misspecifications can be found in Chong (2003), Pitarakis (2004), among others.

Table 1: Size of the UM/CM tests in correctly specified models

DGP	Trim	Model		UM_T^*				CM_T^*			
		α	β	T=100	200	500	1000	100	200	500	1000
DGP1 - AR(1) model, iid errors	15%	1	0.1	0.034	0.040	0.047	0.052	0.067	0.056	0.050	0.050
			0.2	0.039	0.047	0.054	0.056	0.078	0.060	0.050	0.051
			0.3	0.040	0.048	0.053	0.057	0.082	0.063	0.056	0.052
			0.4	0.044	0.055	0.056	0.056	0.091	0.064	0.058	0.048
			0.5	0.052	0.052	0.062	0.060	0.107	0.068	0.063	0.050
			0.6	0.061	0.058	0.070	0.065	0.120	0.079	0.055	0.058
			0.7	0.082	0.072	0.075	0.069	0.141	0.092	0.062	0.058
DGP3 - static model, iid errors	15%	1	0.1	0.067	0.052	0.060	0.050	0.075	0.063	0.059	0.041
			0.2	0.073	0.053	0.049	0.043	0.086	0.052	0.054	0.047
			0.3	0.067	0.056	0.054	0.052	0.083	0.062	0.049	0.049
			0.4	0.084	0.055	0.049	0.052	0.091	0.058	0.042	0.052
			0.5	0.061	0.053	0.058	0.051	0.070	0.056	0.061	0.051
			0.6	0.058	0.058	0.046	0.054	0.080	0.056	0.044	0.052
			0.7	0.073	0.056	0.056	0.055	0.081	0.059	0.054	0.048
			0.8	0.063	0.060	0.047	0.048	0.079	0.059	0.051	0.048
			0.9	0.069	0.062	0.055	0.050	0.092	0.065	0.050	0.052
DGP 4 - static model, AR(1) errors	15%	1	0.1	0.063	0.055	0.067	0.064	0.115	0.107	0.090	0.098
			0.3	0.065	0.061	0.066	0.062	0.113	0.111	0.090	0.100
			0.5	0.063	0.060	0.063	0.059	0.122	0.109	0.091	0.106
			0.7	0.060	0.061	0.070	0.060	0.111	0.113	0.104	0.109
			0.9	0.064	0.061	0.065	0.065	0.109	0.106	0.086	0.099

Table 2: Size of the UM/CM tests in misspecified models

DGP	Estimated Model	Trim	Model α	β	UM_T^* T=100	200	500	1000	CM_T^* 100	200	500	1000
DGP2 - AR(4), iid errors	AR(1)	15%	1	0.1	0.157	0.114	0.114	0.088	0.427	0.463	0.486	0.508
				0.2	0.191	0.131	0.132	0.102	0.493	0.509	0.531	0.560
				0.3	0.231	0.171	0.174	0.130	0.554	0.579	0.606	0.616
DGP2 - AR(4), iid errors	AR(2)	15%	1	0.1	0.157	0.113	0.114	0.088	0.419	0.371	0.336	0.332
				0.2	0.190	0.131	0.131	0.103	0.455	0.378	0.344	0.336
				0.3	0.230	0.170	0.173	0.129	0.496	0.421	0.374	0.358
DGP3 - static model, iid errors	AR(1)	15%	1	0.1	0.023	0.031	0.044	0.047	0.067	0.054	0.053	0.046
				0.2	0.025	0.032	0.042	0.044	0.067	0.054	0.047	0.049
				0.3	0.030	0.036	0.043	0.046	0.070	0.056	0.048	0.056
				0.4	0.027	0.033	0.041	0.046	0.070	0.053	0.048	0.052
				0.5	0.024	0.036	0.044	0.046	0.065	0.055	0.051	0.051
				0.6	0.028	0.032	0.044	0.045	0.074	0.053	0.049	0.051
				0.7	0.027	0.033	0.041	0.046	0.067	0.055	0.049	0.048
				0.8	0.028	0.031	0.043	0.049	0.073	0.054	0.050	0.051
				0.9	0.022	0.032	0.046	0.045	0.065	0.053	0.053	0.052
DGP4 - static model, AR(1) errors	X_t^2 instead of X_t	15%	1	0.1	0.058	0.060	0.064	0.062	0.207	0.165	0.111	0.121
				0.3	0.064	0.061	0.064	0.062	0.204	0.158	0.120	0.125
				0.5	0.063	0.062	0.063	0.062	0.211	0.162	0.121	0.131
				0.7	0.062	0.060	0.069	0.061	0.209	0.160	0.127	0.128
				0.9	0.065	0.061	0.068	0.064	0.216	0.173	0.137	0.143

Table 3: Size of the UV/CV break tests in correctly specified models

DGP	Trim	Model		UV_T^*				CV_T^*			
		α	σ	T=100	200	500	1000	100	200	500	1000
DGP1 - AR(1) model, iid errors	15%	1	1	0.042	0.052	0.051	0.059	0.0753	0.0570	0.050	0.048
			1.2	0.042	0.047	0.051	0.050	0.0730	0.0533	0.050	0.048
			1.4	0.039	0.046	0.050	0.056	0.0754	0.0584	0.050	0.052
			1.6	0.040	0.043	0.052	0.054	0.0745	0.0517	0.053	0.047
			1.8	0.040	0.045	0.052	0.055	0.0735	0.0543	0.053	0.051
			2.0	0.040	0.047	0.052	0.054	0.0775	0.0572	0.053	0.052
			2.2	0.041	0.044	0.054	0.057	0.0768	0.0592	0.050	0.051
			2.4	0.043	0.049	0.054	0.055	0.0774	0.0595	0.051	0.048
			2.6	0.043	0.047	0.052	0.052	0.0738	0.0581	0.049	0.048
DGP4 - static model, AR(1) errors	15%	1	1	0.041	0.051	0.052	0.057	0.076	0.067	0.059	0.060
			1.2	0.041	0.049	0.053	0.055	0.075	0.067	0.060	0.058
			1.4	0.042	0.048	0.053	0.058	0.076	0.066	0.060	0.060
			1.6	0.039	0.047	0.056	0.054	0.071	0.064	0.063	0.056
			1.8	0.041	0.048	0.056	0.053	0.077	0.066	0.063	0.057
			2.0	0.043	0.048	0.055	0.057	0.076	0.062	0.063	0.061
			2.2	0.039	0.049	0.058	0.052	0.073	0.064	0.063	0.056
			2.4	0.042	0.053	0.056	0.054	0.075	0.072	0.061	0.057
			2.6	0.045	0.048	0.053	0.057	0.072	0.061	0.058	0.060

Table 4: Size of the CV/UV tests in misspecified models

DGP	Estimated Model	Trim	Model		UV_T^*				CV_T^*			
			α	σ	T=100	200	500	1000	100	200	500	1000
DGP2-AR(4) model, iid errors	AR(1)	15%	1	1	0.057	0.064	0.075	0.063	0.110	0.107	0.124	0.121
				1.2	0.061	0.063	0.072	0.065	0.112	0.102	0.118	0.129
				1.4	0.059	0.068	0.072	0.065	0.116	0.112	0.119	0.129
				1.6	0.062	0.061	0.072	0.063	0.116	0.100	0.118	0.125
				1.8	0.055	0.056	0.073	0.067	0.113	0.104	0.120	0.125
				2	0.057	0.064	0.068	0.062	0.119	0.106	0.115	0.123
				2.2	0.060	0.063	0.077	0.064	0.118	0.099	0.119	0.122
				2.4	0.057	0.063	0.075	0.067	0.108	0.108	0.121	0.125
				2.6	0.062	0.065	0.074	0.065	0.111	0.113	0.116	0.125
DGP1-AR(1) model, iid errors	AR(4)	15%	1	1	0.042	0.052	0.051	0.058	0.075	0.054	0.047	0.047
				1.2	0.042	0.046	0.050	0.051	0.075	0.052	0.048	0.046
				1.4	0.040	0.046	0.050	0.054	0.070	0.054	0.049	0.050
				1.6	0.040	0.043	0.052	0.055	0.079	0.052	0.049	0.046
				1.8	0.040	0.046	0.051	0.055	0.076	0.051	0.049	0.050
				2	0.040	0.046	0.052	0.053	0.075	0.056	0.050	0.050
				2.2	0.040	0.045	0.053	0.057	0.073	0.051	0.047	0.049
				2.4	0.043	0.049	0.054	0.056	0.074	0.053	0.048	0.046
				2.6	0.043	0.048	0.053	0.052	0.076	0.055	0.047	0.046
DGP3-static model, iid errors	AR(1)	15%	1	1	0.030	0.033	0.041	0.047	0.077	0.057	0.051	0.051
				1.2	0.030	0.032	0.040	0.048	0.076	0.058	0.048	0.053
				1.4	0.029	0.032	0.043	0.047	0.075	0.055	0.055	0.051
				1.6	0.028	0.032	0.041	0.044	0.072	0.054	0.049	0.050
				1.8	0.028	0.033	0.039	0.049	0.071	0.054	0.047	0.053
				2	0.029	0.037	0.040	0.048	0.074	0.061	0.049	0.054
				2.2	0.029	0.033	0.040	0.043	0.080	0.057	0.049	0.045
				2.4	0.026	0.036	0.039	0.044	0.069	0.058	0.048	0.049
				2.6	0.029	0.032	0.045	0.045	0.074	0.055	0.056	0.050
DGP4-static model, AR(1) errors	X_t^2 instead of X_t	15%	1	1	0.044	0.050	0.053	0.055	0.075	0.064	0.059	0.058
				1.2	0.044	0.049	0.053	0.052	0.074	0.063	0.058	0.055
				1.4	0.043	0.049	0.050	0.055	0.073	0.064	0.058	0.058
				1.6	0.041	0.048	0.057	0.053	0.071	0.063	0.063	0.057
				1.8	0.043	0.049	0.052	0.057	0.070	0.064	0.060	0.059
				2	0.046	0.049	0.053	0.055	0.076	0.065	0.058	0.059
				2.2	0.043	0.049	0.056	0.051	0.069	0.064	0.063	0.054
				2.4	0.046	0.052	0.053	0.049	0.075	0.067	0.060	0.052
				2.6	0.042	0.049	0.052	0.056	0.075	0.060	0.057	0.058

Figure 1: DGP1 - Size-adjusted power for a correctly specified AR(1) model with iid errors

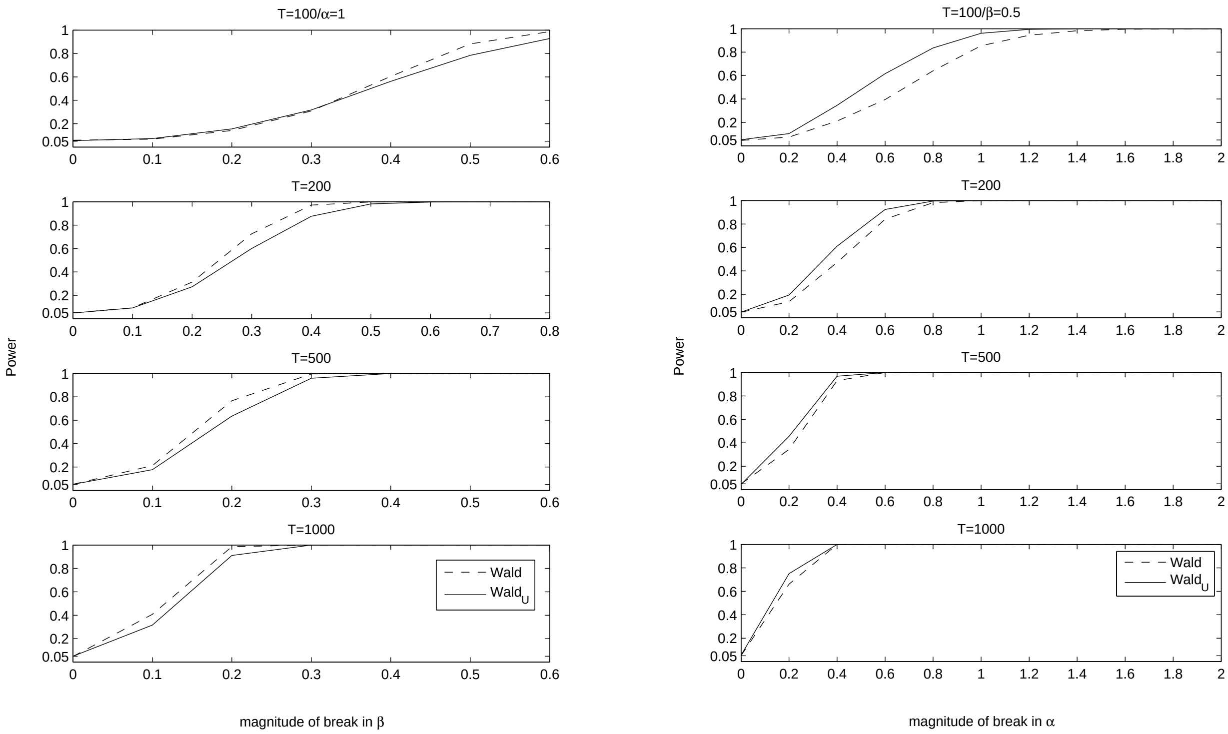
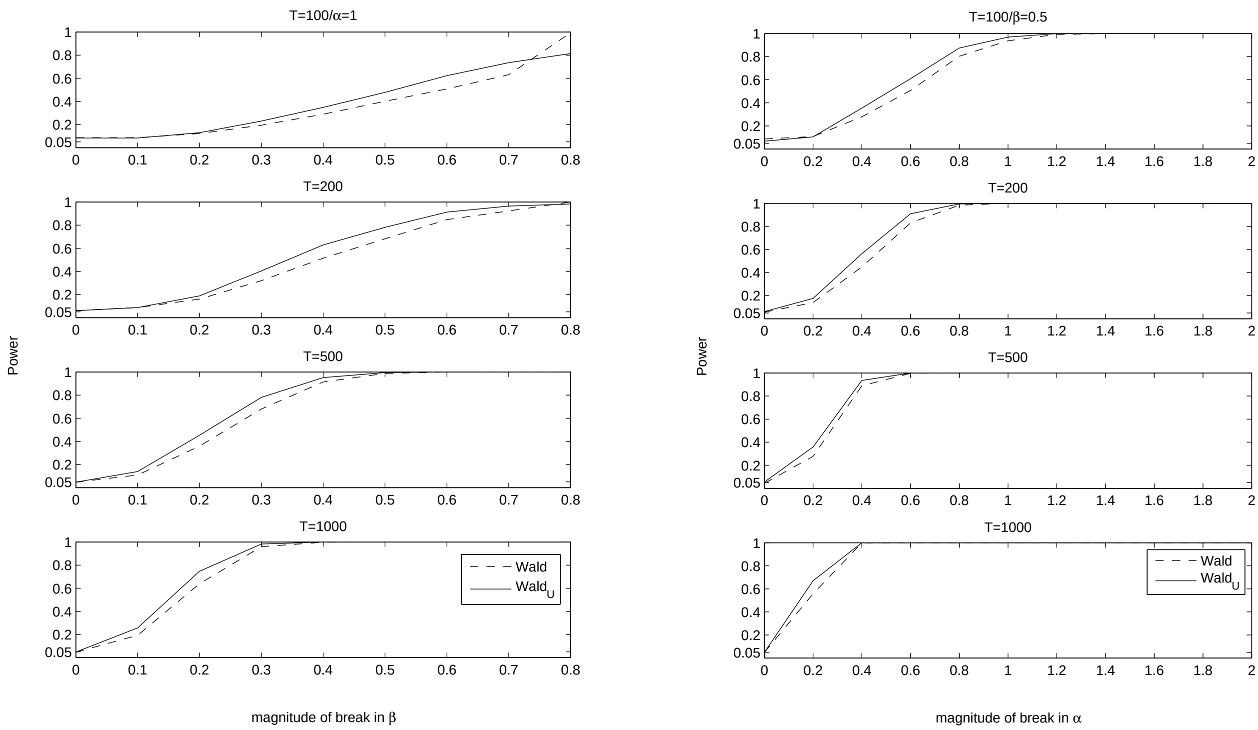


Figure 2: DGP3 - Size-adjusted power for a correctly specified static model with iid errors



*Here, $Wald$ is the CM test, and $Wald_u$ the UM test.

Figure 3: DGP4 - Size-adjusted power for a correctly specified static model with AR(1) errors

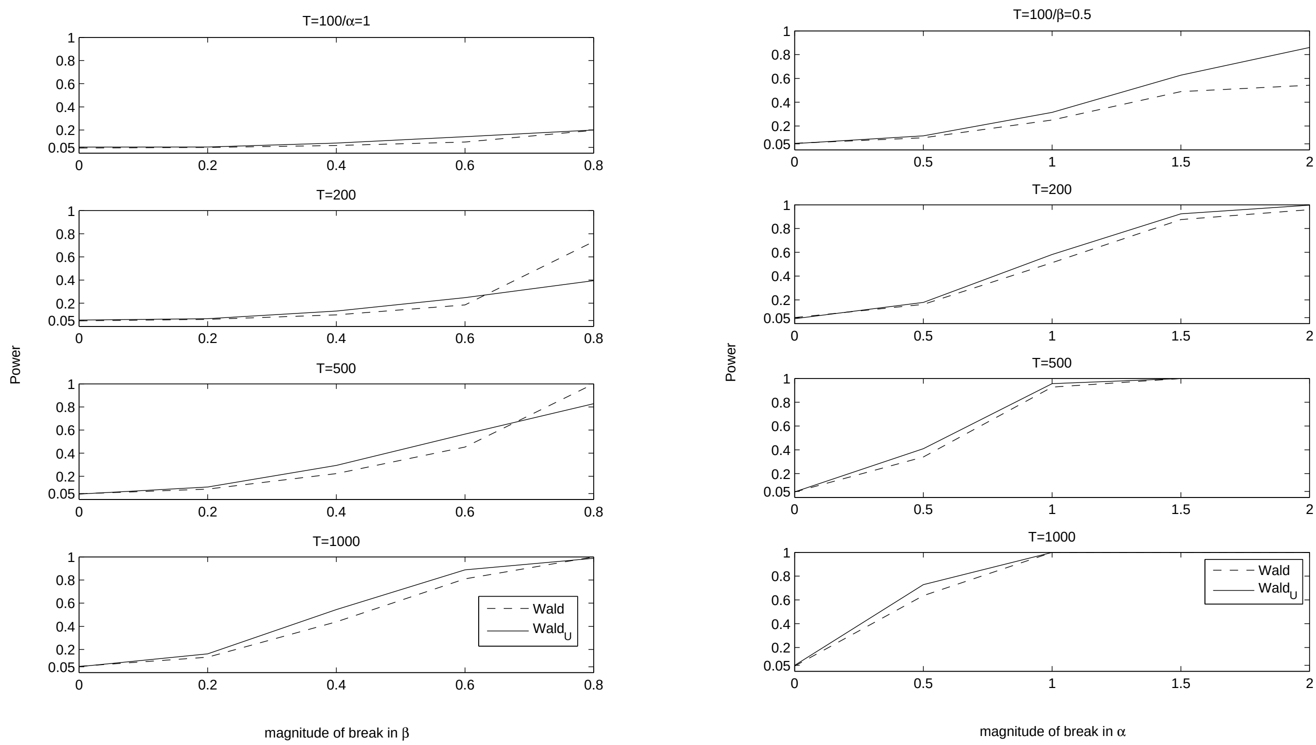
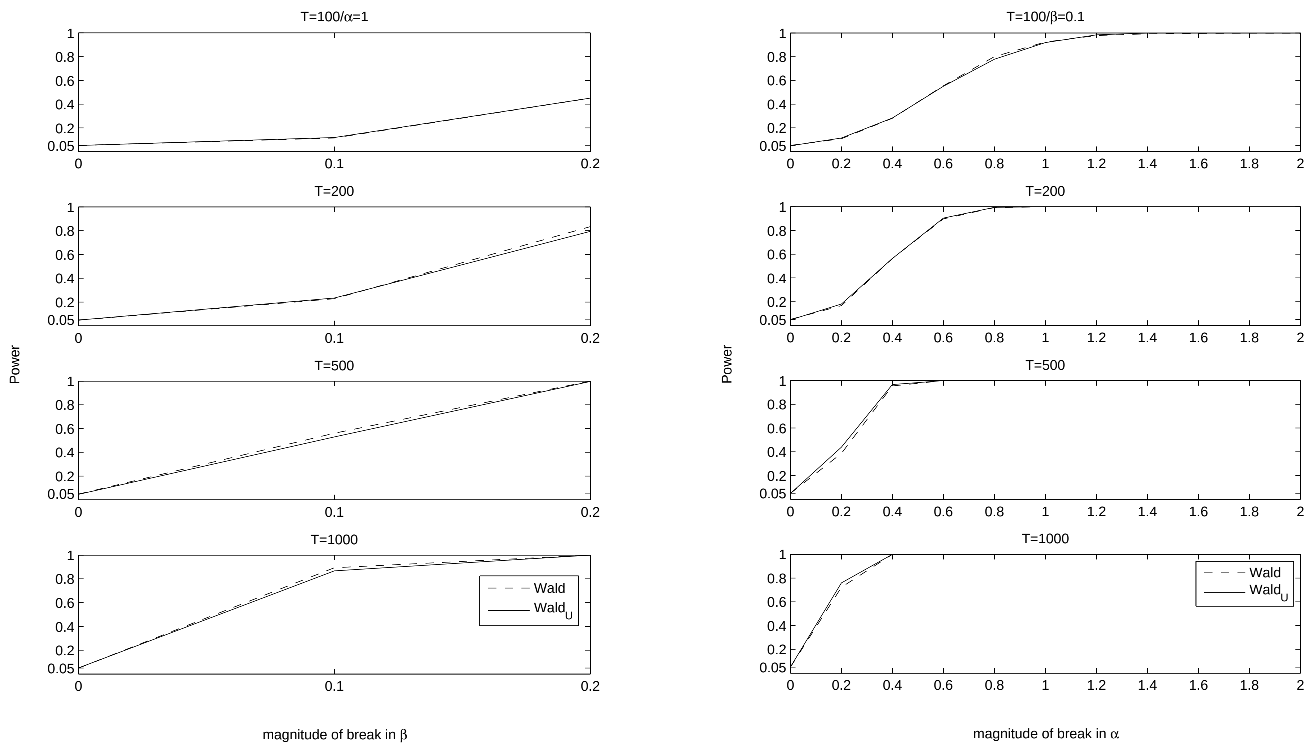


Figure 4: DGP2 - Size-adjusted power for dynamic misspecification: estimating an AR(1) model instead of an AR(4)



*Here, $Wald$ is the CM test, and $Wald_u$ the UM test.

Figure 5: DGP3 - Size-adjusted power for dynamic misspecification: estimating an AR(1) model instead of a static model with iid errors

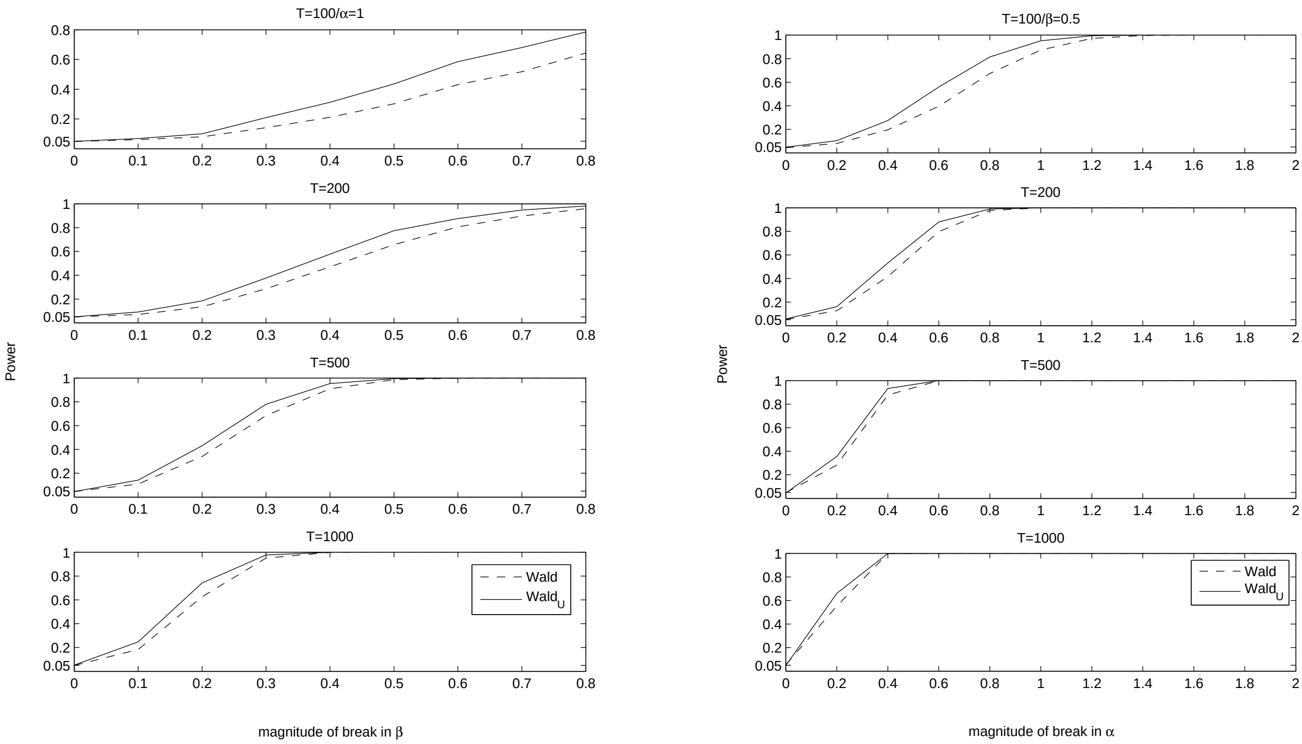
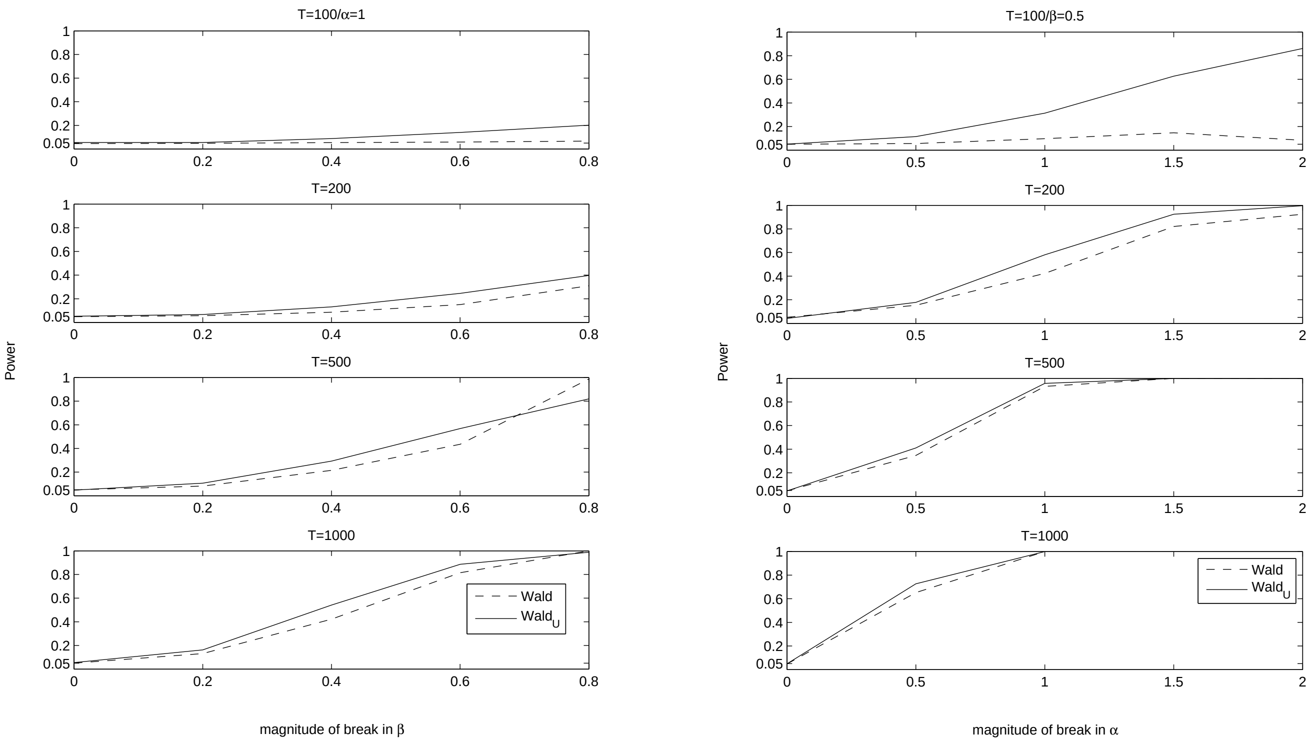


Figure 6: DGP4 - Size-adjusted power for nonlinear misspecification: regressing on X_t^2 instead of X_t in a static model with AR(1) errors



*Here, $Wald$ is the CM test, and $Wald_u$ the UM test.

Figure 7: DGP1 (left), DGP4 (right) - Size-adjusted power for a correctly specified model: with an AR(1) lag and iid errors (left), or with static regressors and AR(1) errors (right)

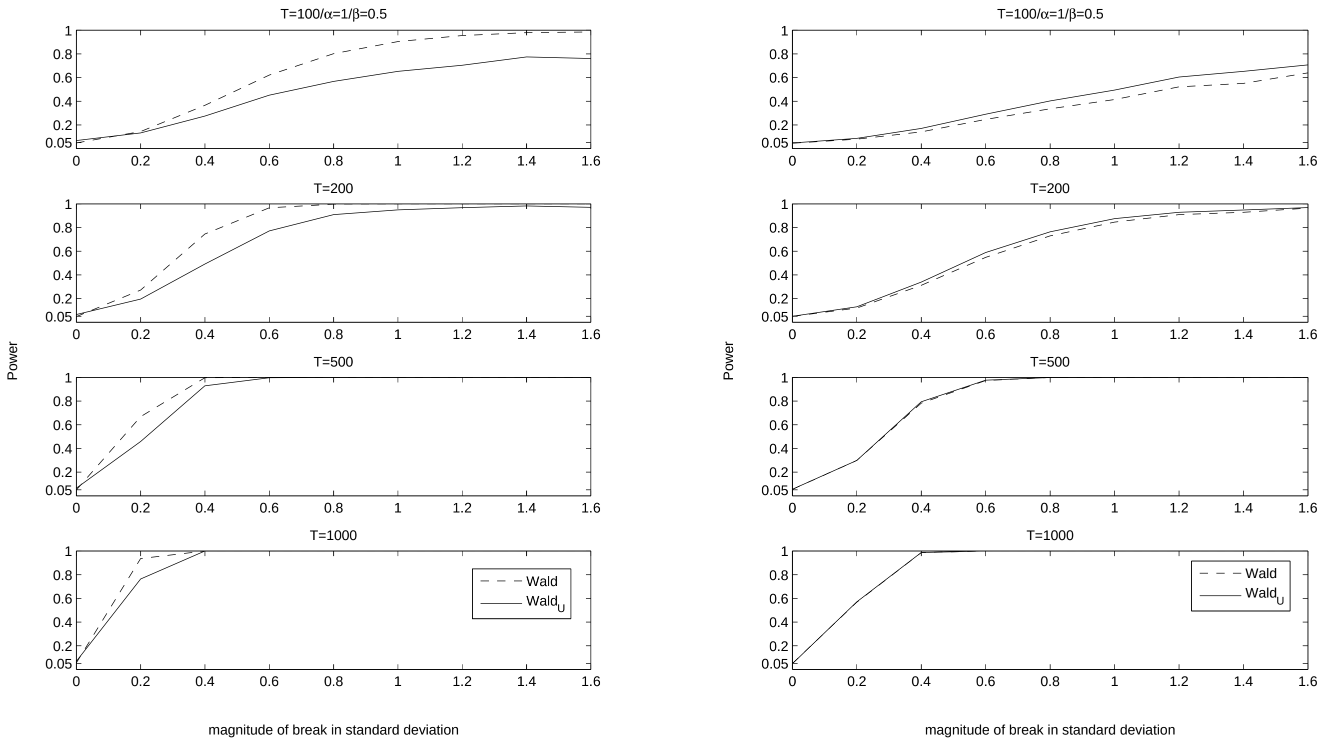
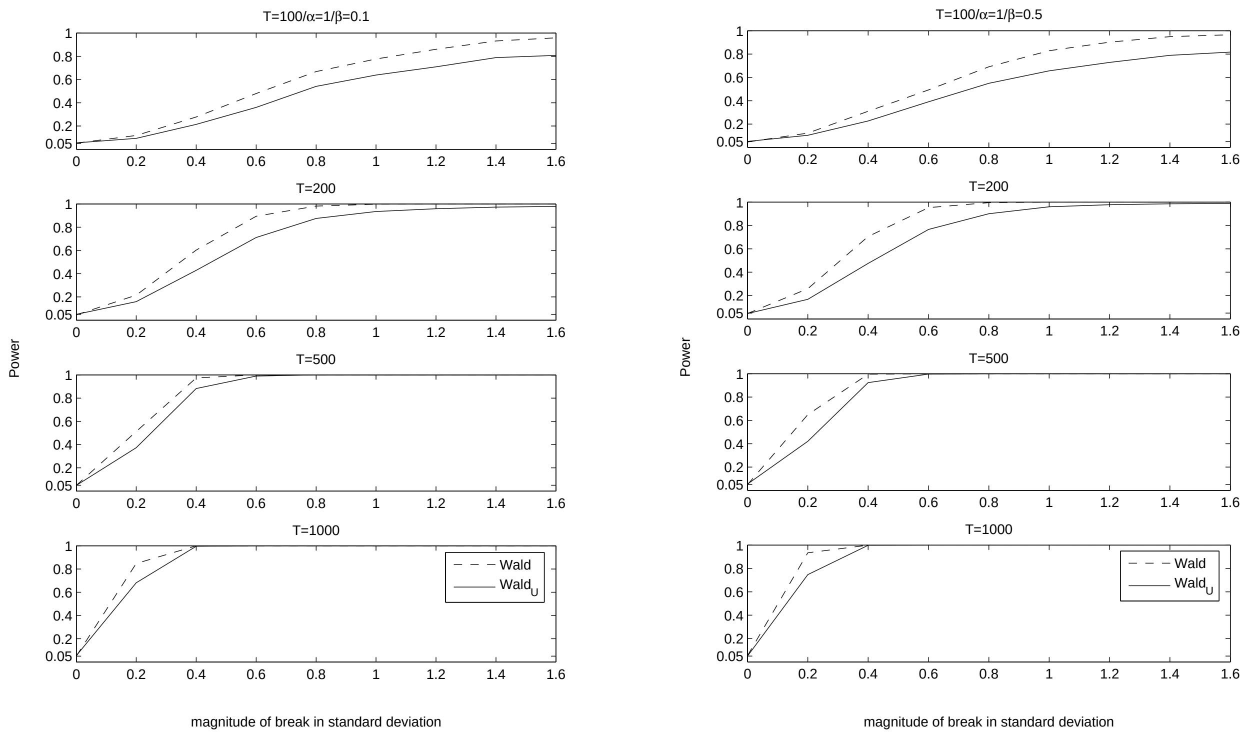
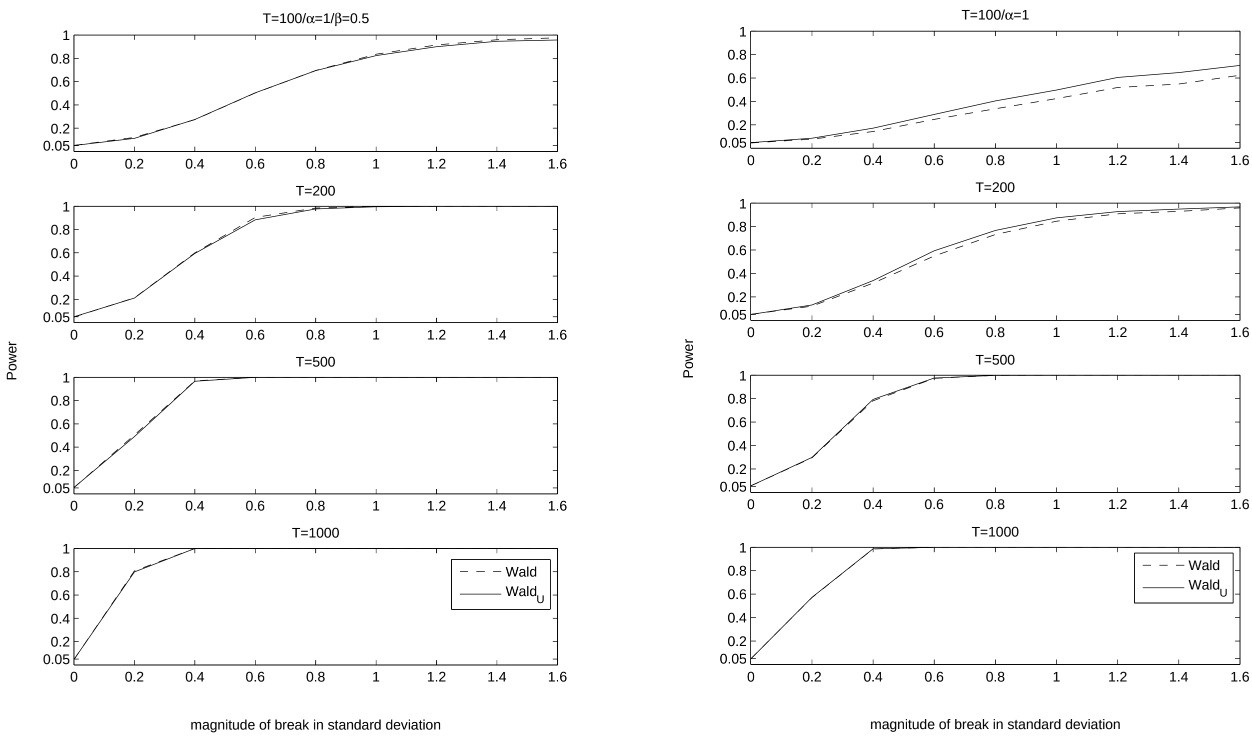


Figure 8: DGP2 (left), DGP1 (right): Size-adjusted power for dynamic misspecification: an AR(4) model (iid errors) misspecified as an AR(1) (left) and vice versa (right)



*Here, $Wald$ is the CV test and $Wald_U$ the UV test.

Figure 9: DGP3 (left), DGP4 (right): Size-adjusted power for misspecification: a static model with iid errors misspecified as an AR(1) model (left), and nonlinear misspecification: a static model with AR(1) errors where the regressors are X_t^2 instead of X_t (right)



*Here, $Wald$ is the CV test and $Wald_U$ the UV test.

5 Empirical Illustrations

This section illustrates the implications of the UM/UV and CM/CV breaks tests for two macroeconomic series. The first is the monthly US civilian unemployment rate and the second the monthly short term real interest rates. These variables are also examined for breaks in the conditional mean and volatility in Stock and Watson (2002) and Sensier and van Dijk (2004), among others, at quarterly frequency. We employ the analysis on a monthly sample from 1960:1-2014:10 with $T = 658$, taken from the FRED database at the Federal Reserve Bank of St. Louis. Given our focus on breaks tests for stationary processes, we transform these series by taking their first differences, based on the evidence from a number of unit root tests. Similar transformations are employed in Stock and Watson (2002), and *from now on, when we refer to unemployment and interest rates, we mean the first differences in unemployment rates and interest rates, unless stated otherwise*. We apply the breaks tests using both 5% and 10% trimming, to detect potential breaks due to the recent economic crisis. For all the tables, we use the critical values in Andrews (2003).

The UM and CM tests for unemployment using AR(p) models with p=1,4,12 are reported in Table 5. The UM test provides empirical evidence of no structural breaks in the US unemployment. In contrast, the CM tests for the AR(1) model find a break in the conditional mean in the late 1960s, associated with women joining the labor force. One may argue that these two results can be reconciled if both the intercept and the slope of the dynamic model have undergone a structural change, in a way that does not affect the long-run structural unemployment rate.[19] On the other hand, our simulations show that the CM test is severely oversized when underestimating the lag order of a dynamic model, even for large sample sizes of $T = 500$ and 1000. These results may imply that the 1960s break detected by CM could be spurious, and do not represent a true break in the long-run mean of unemployment (in first differences).

The evidence of no structural change in the unconditional mean is also consistent with the findings of Stock and Watson (2002), who estimate AR(4) models for the quarterly difference in the US unemployment rate and find no break in the conditional mean, over a shorter period from 1959-2001. Therefore, we also analyze the quarterly first difference in the US civilian unemployment for an extended sample period (up to 2014) in Table 6. Both UM and CM tests indicate no mean break. Moreover, selecting the number of breaks via Information Criteria (IC)[20], for both the monthly and quarterly unemployment (in Tables 5 and 6, respectively), we find additional empirical support of no shift in both the conditional and unconditional mean. The monthly and quarterly IC results hold for our

[19] Recall that we are analyzing the first-difference and not the level of the unemployment rate.

[20] We use the modified BIC proposed by Hall et.al (2013), which employ a modified penalty function in which each break is equivalent to the estimation of three individual regression coefficients, instead of the BIC, which tends to underestimate the number of breaks. Additionally, we use the LWZ criterion which does not require any modified penalty.

Table 5: Structural breaks in the mean of unemployment

Moments/Models	Test	Trimming	Statistic Value	Critical value	Break fraction	Break date
Unconditional Mean sup Wald tests:						
$E(y_t)$	UM_T^*	10% 5%	1.711 3.577	9.11 9.71	0.417 0.928	- -
Conditional Mean sup Wald tests:						
AR(1)	CM_T^*	10% 5%	24.506* 24.506*	12.17 12.80	0.179 0.179	12/1969 12/1969
AR(4)	CM_T^*	10% 5%	18.184 41.365*	18.86 19.57	0.898 0.928	- 11/2010
AR(12)	CM_T^*	10% 5%	36.962* 46.340*	32.76 33.63	0.899 0.942	05/2009 09/2011
Regression model with predictors	CM_T^*	10% 5%	26.992* 26.992*	20.81 21.53	0.882 0.882	05/2008 05/2008
Regression model with macro factor	CM_T^*	10% 5%	11.623 11.623	12.17 12.80	0.786 0.786	- -
Regression model with macro uncertainty factor	CM_T^*	10% 5%	13.545* 15.308*	12.17 12.80	0.288 0.938	05/1975 11/2008

Notes: Unemployment refers to the first-difference in the monthly civilian unemployment rate, seasonally adjusted. Sample period: 01/1960-10/2014. Source: FRED. Superscript * indicates rejection of the null hypothesis. Because of different sample sizes available for different regressions due to constructing lags, different break factions may be indicative of the same break point.

extended sample 1959-2014, as well as the subsample 1959-2001, considered in Stock and Watson (2002). Therefore, both the UM test and the IC results support the hypothesis of no long-run structural change in unemployment (in first differences). These results shed light on the current debate as to whether the recent economic crisis caused a permanent shift in the structural unemployment, for which the evidence in Tables 5 and 6 provide no empirical support.

Next, we test for breaks in the variance of unemployment via the UV and CV tests; these results are reported in Table 7.[21] The UV test yields no evidence of structural change. In contrast, the CV tests, based on AR(p) models, show a structural break in the conditional variance of the unemployment in the mid 1980s, associated with the Great Moderation period. The simulation evidence in Table 4 (top panel) showed that dynamic misspecification, especially underestimation of the number of lags, yields severely over-

[21]The results for UA and CA tests are identical and omitted to save space.

Table 6: Structural breaks in the mean and variance of unemployment

Moments/Models	Test	Trim	Statistic Value	Critical value	Break fraction	Break date
Unconditional Mean sup Wald tests:						
$\mathrm{E}(y_t)$	UM_T^*	10%	1.041	9.11	0.435	-
Conditional Mean sup Wald tests:						
AR(4)	CM_T^*	10%	10.937	18.86	0.403	-
Unconditional Variance sup Wald tests:						
$\mathrm{Var}(y_t)$	UV_T^*	10%	2.400	9.11	0.894	-
Conditional Variance sup Wald tests:						
AR(4)	CV_T^*	10%	7.443	9.11	0.435	-

Notes: See Table 5. The variance tests are corrected for a mean break, when necessary, as in Pitarakis (2004).

sized CV tests even for $T = 1000$, while their power is comparable to the UV tests. This might explain the difference in results between the two tests. Alternatively, the results in Table 7 can be taken to suggest that although there is no break in the unconditional (long-run) variance of unemployment, there is a structural change in the conditional (short-run) variance of the unemployment dynamics, possibly related to the Great Moderation. It is worth mentioning that the mixed empirical evidence on volatility breaks in unemployment provided by the two tests is also found in other studies using quarterly data. Namely, while Stock and Watson (2002) report no evidence of breaks using the UV test for quarterly unemployment, Sensier and van Dijk (2004) find support for a Great Moderation volatility break when using an AR(4) model.

We further analyze the stability of the unemployment via a representative distributed lag regression, using a small set of predictors. The model includes an intercept, the first lag of the dependent variable, and the first lag of each of the predictors. The predictors used in this model, in addition to the lagged dependent variable, are the real economic activity proxied by the industrial production index as in Yashiv (2000), the average hours worked, the unemployment insurance claims and the US real interest rate (e.g. Blanchard and Wolfers, 2000), all taken from the FRED database.[22] All variables are first log differences, except the hours worked. These transformations are based on the evidence from unit root tests, also employed in Stock and Watson (2002).

The CM breaks tests for this model can be found in Table 5. The CM test now indicates a break associated with the recent crisis, more specifically the Lehman Brothers collapse. The CV breaks test, in Table 7, indicates a Great Moderation break. The tests based on the AR and DL models refer to different conditional moments stemming from

[22]The short term ex-post real interest rate is proxied by the 3-month Treasury Bill rate minus the inflation rate based on the Consumer Price index for all Urban consumers (all items).

Table 7: Structural breaks in the variance of unemployment

Moments/Models	Test	Trimming	Statistic Value	Critical value	Break fraction	Break date
Unconditional Variance sup Wald tests:						
Var(y_t)	UV_T^*	10%	5.913	9.11	0.447	-
		5%	5.913	9.71	0.447	-
Conditional Variance sup Wald tests:						
AR(1)	CV_T^*	10%	16.416*	9.11	0.475	02/1986
		5%	16.416*	9.71	0.475	02/1986
AR(4)	CV_T^*	10%	14.442*	9.11	0.473	02/1986
		5%	14.442*	9.71	0.473	02/1986
AR(12)	CV_T^*	10%	14.557*	9.11	0.466	02/1986
		5%	14.557*	9.71	0.466	02/1986
Regression model with predictors	CV_T^*	10%	13.811*	9.11	0.473	02/1986
		5%	13.811*	9.71	0.473	02/1986
Regression model with macro factor	CV_T^*	10%	12.759*	9.11	0.472	07/1984
		5%	12.759*	9.71	0.472	07/1984
Regression model with macro uncertainty factor	CV_T^*	10%	20.147*	9.11	0.496	02/1986
		5%	20.147*	9.71	0.496	02/1986

Notes: See Table 5.

the different conditioning information, which can also explain the difference in the breaks results.

More generally given that the CM/CV tests examine a different null hypothesis than the UM/UV tests (short-run versus long-run breaks), it is not surprising that these two tests yield different results. Besides potential oversizing stemming from static and dynamic misspecification, the CM tests are also subject to the Hansen (2000) critique, according to which these tests are not robust to potential breaks in the unconditional (marginal) distribution of regressors. We investigate whether this critique applies to our empirical analysis, and report the results in Table 8. We find that there are two regressors with a break in their marginal distribution: the unemployment claims series, which has a UV break, and the hours worked series, which has a UM break. Both breaks occur around the Great Moderation. The CM/CV tests are not robust to these breaks, also explaining the difference in results between the unconditional and conditional tests.

We also apply these tests to two distributed lag models of unemployment, driven by two monthly factors respectively: a macro factor, extracted from the mean of a large cross-section of economic and financial US series, and a macro uncertainty factor (the

Table 8: Structural breaks in the regressors

Moments	Test	Trimming	Statistic Value	Critical value	Break fraction	Break date
Hours Worked						
E(y_t)	UM_T^*	10%	14.805*	9.11	0.467	10/1985
		5%	14.805*	9.71	0.467	10/1985
Unemployment Claims						
E(y_t)	UM_T^*	10%	3.729	9.11	0.898	-
		5%	3.729	9.71	0.898	-
Industrial Production						
E(y_t)	UM_T^*	10%	3.816	9.11	0.118	-
		5%	3.816	9.71	0.118	-
3-month real interest rate						
E(y_t)	UM_T^*	10%	5.154	9.11	0.387	-
		5%	5.154	9.71	0.387	-
Macro Factor						
E(y_t)	UM_T^*	10%	2.646	9.11	0.281	-
		5%	2.646	9.71	0.281	-
Macro Uncertainty Factor						
E(y_t)	UM_T^*	10%	5.214	9.11	0.898	-
		5%	6.295	9.71	0.921	-
Hours Worked						
Var(y_t)	UV_T^*	10%	3.605	9.11	0.892	-
		5%	5.778	9.71	0.947	-
Unemployment Claims						
Var(y_t)	UV_T^*	10%	13.708*	9.11	0.430	10/1983
		5%	13.708*	9.71	0.430	10/1983
Industrial Production						
Var(y_t)	UV_T^*	10%	5.756	9.11	0.436	-
		5%	5.756	9.71	0.436	-
3-month real interest rate						
Var(y_t)	UV_T^*	10%	4.514	9.11	0.409	-
		5%	4.514	9.71	0.409	-
Macro Factor						
Var(y_t)	UV_T^*	10%	4.277	9.11	0.898	-
		5%	10.181*	9.71	0.935	01/2009
Macro Uncertainty Factor						
Var(y_t)	UV_T^*	10%	4.969	9.11	0.898	-
		5%	8.317	9.71	0.928	-

Notes: *First series:* Average Weekly Hours of Production and Nonsupervisory Employees: Manufacturing, Hours, Monthly, Seasonally Adjusted. Sample Period: 01/1960-10/2014. Source: FRED. Second series*:* Average Weekly Initial Claims, Unemployment Insurance (Thousands). Sample Period: 01/1960-10/2014. Source: Ludvigson and Ng (2009). *Third series:* Industrial Production Index, Index 2007=100, Monthly, Seasonally Adjusted. Sample Period: 01/1960-10/2014. Source: FRED. *Fourth series:* 3-month real interest rate (computed as the difference between the 3-month Treasury Bill rate and the 3-month (CPI) based inflation rate, FRED). Sample Period: 04/1960-10/2014. The variance tests are corrected for a mean break, if needed, as in Pitarakis (2004).

models include an intercept and the first lag of each factors). Both factors are taken from Jurado et al. (2014) and their use is further motivated in Benigno et al. (2015), inter alia. We examine the stability of the relationship between unemployment and these macro factors; the results are reported in the last two rows of Tables 5 and 7, respectively. We find no break in the dynamic relationship between the unemployment rate and the macro factor (that represents macroeconomic conditions). This result is in contrast to the remaining distributed lag models in Table 5, which are driven by a small set of predictors. An explanation for this mixed evidence can be found in Table 8, which shows that while some of individual predictors in Table 5 are subject to the Hansen (2000) critique, the macro factor is not. Hence, applying the CM tests to factor augmented models yields more reliable inference than the aforementioned regression models based on certain economic variables.

Thus, both the CM test of the factor-augmented regression with the macro factor and the UM test indicate no evidence of mean breaks in unemployment.[23] Turning though to the results in Table 7, there is consistent evidence of a structural change associated with the Great Moderation in the conditional variance of the unemployment rate.

To summarize, we find no long-run breaks in the unemployment mean, in its volatility, or in its relationship to the overall macro conditions approximated by macro factors. However, there is evidence of short-run mean shifts in the unemployment rate related to women entering the labor force, or shifts in the relationship to other macroeconomic variables around the Great Moderation. There is further evidence of conditional variance breaks. It is worth emphasizing that our results are for the first-difference in the unemployment rate, and not the level of this series. There may be long-run shifts in the level of structural unemployment rates, but these can be detected only with tests for nonstationary variables, which are beyond the scope of this paper.

Turning to the analysis of short-term interest rates, several papers find breaks in their conditional mean - see e.g. Garcia and Perron (1996), Stock and Watson (2002) and Sensier and van Dijk (2004). We focus on the first-differenced interest rates; they are computed as the annualized nominal 3-month Treasury Bill minus the 3-month CPI inflation rate, over the sample period 1960:1-2014:10. Tables 9 and 10 show that the UM and UV tests detect no break whilst the corresponding CM and CV tests, based on the AR(p) model, detect a break associated with the oil crises.[24] In particular, the CM and CV tests report breaks around the mid 1970s and early 1980s, respectively. Similar to the unemployment evidence reported earlier, the short-run breaks in the conditional moments do not reflect the long-run unconditional moment structural change detected in

[23]It is worth mentioning that the macro uncertainty factor provides only weak and mixed empirical evidence of a break in the relationship between the unemployment and macro uncertainty, as shown in the last row of Table 5.

[24]In Table 8 we use the real interest rates whereas Tables 9 and 10 we use the annualized real interest rates because the latter features as the primary series in many other empirical analyses. The qualitative results are the same whether the real interest rates are annualized or not.

the US interest rates.

Table 9: Structural breaks in the mean of real interest rates

Moments/Models	Test	Trimming	Statistic Value	Critical value	Break fraction	Break date
Unconditional Mean sup Wald tests:						
E(y_t)	UM_T^*	10%	5.900	9.11	0.386	-
		5%	5.900	9.71	0.386	-
Conditional Mean sup Wald tests:						
AR(1)	CM_T^*	10%	23.940*	12.17	0.366	04/1980
		5%	23.940*	12.80	0.366	04/1980
AR(2)	CM_T^*	10%	44.190*	14.69	0.262	09/1974
		5%	44.190*	15.36	0.262	09/1974
AR(3)	CM_T^*	10%	46.807*	16.91	0.261	09/1974
		5%	46.807*	17.54	0.261	09/1974
AR(4)	CM_T^*	10%	43.762*	18.86	0.261	09/1974
		5%	43.762*	19.57	0.261	09/1974
AR(12)	CM_T^*	10%	72.248*	32.76	0.889	11/2008
		5%	72.248*	33.63	0.889	11/2008

Notes: The real interest rate is the first difference in the annualized 3-month real interest rate (computed as the annualized 3-month Treasury Bill rate minus the annualized 3-month (CPI) inflation rate). Sample Period: 04/1960-10/2014. Superscript * indicates rejection of the null hypothesis. Source: FRED.

Table 10: Structural breaks in the variance of real interest rates

Moments/Models	Test	Trimming	Statistic Value	Critical value	Break fraction	Break date
Unconditional Variance sup Wald tests:						
Var(y_t)	UV_T^*	10%	4.823	9.11	0.409	-
		5%	4.823	9.71	0.409	-
Conditional Variance sup Wald tests:						
AR(1)	CV_T^*	10%	20.720*	9.11	0.410	09/1982
		5%	20.720*	9.71	0.410	09/1982
AR(2)	CV_T^*	10%	24.770*	9.11	0.412	11/1982
		5%	24.770*	9.71	0.412	11/1982
AR(3)	CV_T^*	10%	23.219*	9.11	0.410	10/1982
		5%	23.219*	9.71	0.410	10/1982
AR(4)	CV_T^*	10%	23.452*	9.11	0.409	10/1982
		5%	23.452*	9.71	0.409	10/1982
AR(12)	CV_T^*	10%	20.746*	9.11	0.398	08/1982
		5%	20.746*	9.71	0.398	08/1982

Notes: See Table 9.

6 Conclusion

In this paper, we propose an alternative and complementary approach to the sup Wald test for breaks in the conditional mean and variance. We show that the corresponding unconditional mean and variance break tests exhibit not only comparable size and power properties but are also robust to various forms of regression model misspecification. We show that under certain commonly encountered forms of regression model misspecification, the traditional conditional mean break tests suffer from severe oversizing, even for large sample sizes, compared to the unconditional mean break tests which don't suffer from this problem. Moreover, both tests have similar size-adjusted power as the sample size grows. In a comprehensive empirical analysis, we apply these tests to show that there is no evidence of long-run breaks in the US civilian unemployment growth, and US short-term real interest rate growth.

7 References

Altansukh, G., Osborn, D.R., Bratsiotis, G., Becker, R., 2012. Structural breaks in international inflation linkages for OECD countries. Unpublished manuscript.

Andreou, E., Ghysels, E., 2002. Detecting multiple breaks in financial market volatility dynamics. Journal of Applied Econometrics 17, 579-600.

Andrews, D.W.K., 1991. Heteroskedasticity and autocorrelation consistent covariance matrix estimation. Econometrica 59, 817-858.

Andrews, D.W.K., 1993. Tests for parameter instability and structural change with unknown change point. Econometrica 61, 821-856.

Andrews, D.W.K., 2003. Tests for parameter instability and structural change with unknown change point: a corrigendum. Econometrica 71, 395-397.

Andrews, D.W.K., Ploberger, W., 1994. Optimal tests when a nuisance parameter is present only under the alternative. Econometrica 62, 1383-1414.

Aue, A., Horváth, L., 2012. Structural breaks in time series. Journal of Time Series Analysis 34, 1-16.

Bai, J., Perron, P., 1998. Estimating and testing linear models with multiple structural changes. Econometrica 66, 47-78.

Bai, J., Chen, H., Chong, T.-L.T., Wang, X.S., 2008. Generic consistency of the break-point estimators under specification errors in a multiple-break model. Econometrics Journal 11, 287-307.

Bataa, E., Osborn, D., Sensier, M., Van Dijk, D., 2013. Structural breaks in the international dynamics of inflation. The Review of Economics and Statistics 95, 646-659.

Benigno, P., Ricci, L.A., Surico, P., 2015. Unemployment and productivity in the long run: the role of macroeconomic volatility. The Review of Economics and Statistics 97, 698-709.

Blanchard, O., Wolfers, J., 2000. The role of shocks and institutions in the rise of european unemployment: the aggregate evidence. The Economic Journal 110, 1-33.

Cho, C.-K., Vogelsang, T.J., 2014. Fixed-b inference for testing structural change in a time series regression. Working Paper.

Chong, T.-L. T., 2003. Generic consistency of the break-point estimator under specification errors. Econometrics Journal 6, 167-192.

Davidson, J., 1994. Stochastic limit theory : an introduction for econometricians. Oxford University Press, Oxford.

Garcia, R., Perron, P., 1996. An analysis of the real interest rate under regime shifts. The Review of Economics and Statistics 78, 111-125.

Hall, A.R., Han, S., Boldea, O., 2012. Inference regarding multiple structural changes in linear models with endogenous regressors. Journal of Econometrics 170, 281-302.

Hall, A.R., Osborn, D.R., Sakkas, N., 2013. Inference on structural breaks using information criteria. The Manchester School 81, 54-81.

Hansen, B.E., 2000. Testing for structural change in conditional models. Journal of Econometrics 97, 93-115.

Jurado, K., Ludvigson, S.C., Ng, S., 2014. Measuring uncertainty. American Economic Review 105, 1177-1216.

Kejriwal, M., 2009. Tests for a mean shift with good size and monotonic power. Economic Letters 102, 78–82.

Liu, J., Wu, S., Zidek, J.V., 1997. On segmented multivariate regresssion. Statistica Sinica 7, 497-525.

Ludvigson, S.C., Ng, S., 2009. Macro Factors in Bond Risk Premia. The Review of Financial Studies 22, 5027-5067.

McConnell, M.M., Perez-Quiros, G., 2000. Output fluctuations in the United States: what has changed since the early 1980's?. American Economic Review 90, 1464-1476.

McKitrick, R.R., Vogelsang, T.J., 2014. HAC robust trend comparisons among climate series with possible level shifts. Environmetrics 25, 528–547.

Newey, W.K., West, K.D., 1994. Automatic lag selection in covariance matrix estimation. Review of Economic Studies 61, 631-653.

Perron, P., Yabu, T., 2009. Testing for shifts in trend with an integrated or stationary noise component. Journal of Business and Economic Statistics 27, 369-396.

Pitarakis, J.-Y., 2004. Least squares estimation and tests of breaks in mean and variance under misspecification. Econometrics Journal 7, 32-54.

Ploberger, W., Krämer, W., 1992. The CUSUM Test with OLS Residuals. Econometrica 60, 271-285.

Qu, Z., Perron, P., 2007. Estimating and Testing Structural Changes in Multivariate Regressions. Econometrica 75, 459-502.

Sayginsoy, Ö., Vogelsang, T.J., 2011. Testing for a shift in trend at an unknown date: a fixed-b analysis of heteroskedasticity autocorrelation robust OLS-based tests. Econometric Theory 27, 992-1025.

Sensier, M., van Dijk, D., 2004. Testing for volatility changes in U.S. macroeconomic time series. The Review of Economics and Statistics 86, 833-839.

Stock, J.H., Watson, M.W., 2002. Has the business cycle changed and why?. In NBER Macroeconomics Annual 17, 159-218.

Vogelsang, T.J., 1997. Wald-type tests for detecting breaks in the trend function of a dynamic time series. Econometric Theory 13, 818-848.

Vogelsang, T.J., 1998. Trend function hypothesis testing in the presence of serial correlation. Econometrica 66, 123-148.

Vogelsang, T.J., 1999. Sources of nonmonotonic power when testing for a shift in mean of a dynamic time series. Journal of Econometrics 88, 283-299.

Vogelsang, T.J., Perron, P., 1998. Additional tests for a unit root allowing for a break in the trend function at an unknown time. International Economic Review 39, 1073-1100.

Wooldridge, J.M., White, H., 1988. Some invariance principles and central limit theorems for dependent heterogeneous processes. Econometric Theory 4, 210-230.

Yashiv, E., 2000. The determinants of equilibrium unemployment. American Economic Review 90, 1298-1322.

8 Appendix

Proof of Theorem 1:
Part (i). Aue and Horváth (2012) define a CUSUM test for H_0^{UM} versus H_A^{UM} as follows:

$$\mathcal{Z}_T^* = \sup_{\lambda\in[\epsilon,1-\epsilon]} \mathcal{Z}_T(\lambda), \qquad \mathcal{Z}_T(\lambda) = \tfrac{1}{\sqrt{T}} \left(\sum_{t=1}^{[T\lambda]} y_t - \tfrac{[T\lambda]}{T} \sum_{t=1}^{T} y_t \right) / \widehat{v}_u^{1/2},$$

where $\widehat{v}_u$ is a HAC consistent estimator of $v_u = \mathbf{AVar}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^T y_t\right)$ under H_0^{UM} and A1(i). They state that if a functional central limit theorem (FCLT) holds under H_0^{UM} for $\frac{1}{\sqrt{T}}\sum_{t=1}^{[T\lambda]} u_t$, then $\mathcal{Z}_T(\lambda) \Rightarrow [B_1(\lambda) - \lambda B_1(1)] = \mathcal{B}_1(\lambda)$, and so by the continuous mapping theorem (CMT),

$$\mathcal{Z}_T^* \Rightarrow \sup_{\lambda\in[\epsilon,1-\epsilon]} \mathcal{B}_1(\lambda), \tag{6}$$

where $\mathcal{B}_1(\lambda) = B_1(\lambda) - \lambda B_1(1)$ is a scalar independent Brownian bridge. Below, we show that there is a clear connection between the CUSUM and the UM test, so the asymptotic distribution of the second follows from the first.

$$\begin{aligned}
T(\overline{y_{1\lambda}} - \overline{y_{2\lambda}})^2 &= T\left(\tfrac{1}{T_{1\lambda}}\sum_{1\lambda} y_t - \tfrac{1}{T_{2\lambda}}\sum_{2\lambda} y_t\right)^2 = \tfrac{T^3}{T_{1\lambda}^2 T_{2\lambda}^2}\left(\tfrac{T_{2\lambda}}{T}\sum_{1\lambda} y_t - \tfrac{T_{1\lambda}}{T}\sum_{2\lambda} y_t\right)^2 \\
&= \tfrac{T^3}{T_{1\lambda}^2 T_{1\lambda}^2}\left(\sum_{1\lambda} y_t - \tfrac{T_{1\lambda}}{T}\sum_{t=1}^T y_t\right)^2 = \left[\tfrac{1}{\lambda^2(1-\lambda)^2} + \mathbf{o}(1)\right]\left[\tfrac{1}{\sqrt{T}}\left(\sum_{1\lambda} y_t - \tfrac{T_{1\lambda}}{T}\sum_{t=1}^T y_t\right)\right]^2 \\
&\Rightarrow v_u[B_1(\lambda) - \lambda B_1(1)]^2/[\lambda^2(1-\lambda)^2] = v_u\, \mathcal{B}_1^2(\lambda)/[\lambda^2(1-\lambda)^2].
\end{aligned}$$

Since $\mathbf{v}_{u\lambda} = \mathbf{AVar}(\sqrt{T}(\overline{y_{1\lambda}} - \overline{y_{2\lambda}})) = v_u/[\lambda(1-\lambda)]$, $UM_T(\lambda) \Rightarrow \mathcal{B}_1^2(\lambda)/[\lambda(1-\lambda)]$, so:

$$UM_T^* \Rightarrow \sup_{\lambda\in[\epsilon,1-\epsilon]} \mathcal{B}_1^2(\lambda)/[\lambda(1-\lambda)]. \tag{7}$$

Comparing (6) and (7), the two limiting distributions attain their supremum at different λ's, and thus the size of these tests will in general be different.[25] However, underlying the asymptotic theory is the same assumption, that the FCLT holds for $T^{-1/2}\sum_{t=1}^{[T\lambda]} u_t$. A1 guarantees that the FCLT in Wooldridge and White (1988), Theorem 2.11, can be applied for u_t (in fact, we only need $d_m = O(m^{-1/2})$), completing the proof of (i).

Part (ii). Here, we just verify A1 for $|y_t - \overline{y}| - a$ instead of u_t. The rest of the proof is as in part (i) of the proof. A1(i) is straightforward, and we are left to verify A1(ii). Since u_t is $\mathcal{L}_2$-near epoch dependent of size $m^{-1/2}$ on $\{g_t\}$ with positive constants equal to 1 (these constants appear in the near epoch dependent definition in Davidson (1994) but since here they are fixed, they are absorbed into the definition for d_m), it follows that so is $y_t - \overline{y}$, with constants $2\sup_t(1) = 2$. In Theorem 17.12 in Davidson (1994), let $\phi_t(\cdot) = |\cdot|$, a uniform Lipschitz function, with the argument $y_t - \overline{y}$. Then, $y_t - \overline{y}$ is $\mathcal{L}_2$-near epoch dependent of size $m^{-1/2}$.
Part (iii). Here, we just verify A1(ii) for $(y_t - \overline{y})^2 - v_u$, instead of u_t. In Theorem 17.12 in Davidson (1994), under H_0^{UV} and H_0^{UM}, define the function $\phi_t(y_t - \overline{y}) = (y_t - \overline{y})^2 - v_u$.

[25] Also note that the test statistic $\sup_{\lambda\in[\epsilon,1-\epsilon]}\sqrt{UM_T}$ is known in statistics as a "weighted version" of the CUSUM test - see Aue and Horváth (2012), p. 5.

From part (ii) of the proof, $(y_t - \overline{y})$ is a $\mathcal{L}_2$-near epoch dependent process of size m^{-1} on $\{g_t\}$ with constants equal to 2. Below, we show that under H_0^{UM}, ϕ_t is uniform Lipschitz almost surely:

$$
\begin{aligned}
&|\phi_t(y_t - \overline{y}) - \phi_t(y_k - \overline{y})| = |(y_t - \overline{y})^2 - (y_k - \overline{y})^2| \\
&\leq |y_t + y_k - 2\overline{y}| \; |(y_t - \overline{y}) - (y_k - \overline{y})| \leq |u_t + u_k - 2\overline{u}| \; |(y_t - \overline{y}) - (y_k - \overline{y})| \\
&\leq (4 \sup_t |u_t|) \; |(y_t - \overline{y}) - (y_k - \overline{y})| \leq \kappa |(y_t - \overline{y}) - (y_k - \overline{y})|, \text{ almost surely,}
\end{aligned}
$$

for some $\kappa > 0$, by Assumption 1.3 for u_t, where $\overline{u} = T^{-1} \sum_{t=1}^{T} u_t$. Hence, by Theorem 17.12 in Davidson (1994), $(y_t - \overline{y})^2 - v_u$ is also $\mathcal{L}_2$-near epoch dependent of size $m^{-1/2}$. □

Proof of Theorem 2:
(i). Since A2 is a special case of Assumption 8 in Hall, Han and Boldea (2012), the result follows directly from their Theorem 6, setting $x_t = z_t$.
(ii), (iii). Primitive assumptions for CV test can be found in e.g. Qu and Perron (2007) and involve joint mixing assumptions on $\{x_t \epsilon_t\}$ and ϵ_t^2. They mention that these conditions can be replaced by sufficient conditions to yield a FCLT for $\{x_t \epsilon_t\}$ and $\epsilon_t^2 - v_\epsilon$ under the null. By similar reasoning, for the CA test, sufficient conditions to yield a joint FCLT for $\{x_t \epsilon_t\}$ and $|\epsilon_t| - E|\epsilon_t|$ suffice. Since x_t includes an intercept, these conditions can be verified as for the proof of Theorem 1(ii)-(iii). Note that they all require H_0^{CM}. □

Proof of Theorem 3:
Denote, for $i = 1, 2$, $\widehat{Q}_{i\lambda,(1)} = T^{-1} \sum_{i\lambda} x_{t(1)} x'_{t(1)}$, $\widehat{Q}_{i\lambda,(12)} = T^{-1} \sum_{i\lambda} x_{t(1)} x'_{t(2)}$, $\widehat{Q}_{i\lambda,(2)} = T^{-1} \sum_{i\lambda} x_{t(2)} x'_{t(2)}$, where recall that $\sum_{1\lambda} = \sum_{t=1}^{[T\lambda]}$, and $\sum_{2\lambda} = \sum_{[T\lambda]+1}^{T}$. By A2, $\widehat{Q}_{i\lambda,(1)} \xrightarrow{p} \lambda_i Q_{(1)}$, $\widehat{Q}_{i\lambda,(2)} \xrightarrow{p} \lambda_i Q_{(2)}$ and $\widehat{Q}_{i\lambda,(12)} \xrightarrow{p} \lambda_i Q_{(12)}$, where $i = 1, 2$, $\lambda_1 = \lambda$ and $\lambda_2 = 1 - \lambda_1$. Recall that we mistakenly regress y_t only on $x_{t(1)}$; let $\hat{\theta}_{1\lambda}$ and $\hat{\theta}_{2\lambda}$ be the OLS estimators in $\{1, \ldots, [T\lambda]\}$, respectively $\{[T\lambda] + 1, \ldots, T\}$.

$$
\begin{aligned}
\widehat{\theta}_{1\lambda} &= \widehat{Q}_{1\lambda,(1)}^{-1} \textstyle\sum_{1\lambda} x_{t(1)} y_t = \theta_{(1)} + \widehat{Q}_{1\lambda,(1)}^{-1} \widehat{Q}_{1\lambda,(12)} \theta_{(2)} + \widehat{Q}_{1\lambda,(1)}^{-1} T^{-1} \textstyle\sum_{1\lambda} x_{t(1)} \epsilon_t \\
\widehat{\theta}_{2\lambda} &= \widehat{Q}_{2\lambda,(1)}^{-1} \textstyle\sum_{1\lambda} x_{t(2)} y_t = \theta_{(1)} + \widehat{Q}_{2\lambda,(1)}^{-1} \widehat{Q}_{2\lambda,(12)} \theta_{(2)} + \widehat{Q}_{2\lambda,(1)}^{-1} T^{-1} \textstyle\sum_{2\lambda} x_{t(1)} \epsilon_t \\
T^{1/2}(\widehat{\theta}_{1\lambda} - \theta_{(1)}) &= \widehat{Q}_{1\lambda,(1)}^{-1} T^{1/2} \widehat{Q}_{1\lambda,(12)} \theta_{(2)} + \widehat{Q}_{1\lambda,(1)}^{-1} T^{-1/2} \textstyle\sum_{1\lambda} x_{t(1)} \epsilon_t \\
T^{1/2}(\widehat{\theta}_{2\lambda} - \theta_{(1)}) &= \widehat{Q}_{2\lambda,(1)}^{-1} T^{1/2} \widehat{Q}_{2\lambda,(12)} \theta_{(2)} + \widehat{Q}_{2\lambda,(1)}^{-1} T^{-1/2} \textstyle\sum_{2\lambda} x_{t(1)} \epsilon_t \\
T^{1/2}(\widehat{\theta}_{1\lambda} - \widehat{\theta}_{2\lambda}) &= \widehat{Q}_{1\lambda,(1)}^{-1} T^{-1/2} \textstyle\sum_{1\lambda} [x_{t(1)} \epsilon_t + (x_{t(1)} x'_{t(2)} - Q_{(12)}) \theta_{(2)}] \\
&\quad - \widehat{Q}_{2\lambda,(1)}^{-1} T^{-1/2} \textstyle\sum_{2\lambda} [x_{t(1)} \epsilon_t + (x_{t(1)} x'_{t(2)} - Q_{(12)}) \theta_{(2)}] \\
&\quad + (\lambda \widehat{Q}_{1\lambda,(1)}^{-1} - (1 - \lambda) \widehat{Q}_{2\lambda,(1)}^{-1}) \; T^{1/2} Q_{(12)} \theta_{(2)} + o_P(1) \\
&= \widehat{Q}_{1\lambda,(1)}^{-1} T^{-1/2} \textstyle\sum_{1\lambda} (k_t + L_t \theta_{(2)}) - \widehat{Q}_{2\lambda,(1)}^{-1} T^{-1/2} \textstyle\sum_{2\lambda} (k_t + L_t \theta_{(2)}) + o_P(1) \\
&\quad - \lambda(1 - \lambda) \; T^{1/2} \widehat{Q}_{1\lambda,(1)}^{-1} \; [\widehat{Q}_{1\lambda,(1)}/\lambda - \widehat{Q}_{2\lambda,(1)}/(1 - \lambda)] \; \widehat{Q}_{2\lambda,(1)}^{-1} Q_{(12)} \theta_{(2)} + o_P(1)
\end{aligned}
$$

$$= \widehat{Q}^{-1}_{1\lambda,(1)} T^{-1/2} \textstyle\sum_{1\lambda}(k_t + L_t\theta_{(2)}) - \widehat{Q}^{-1}_{2\lambda,(1)} T^{-1/2} \sum_{2\lambda}(k_t + L_t\theta_{(2)}) + o_P(1)$$
$$- \widehat{Q}^{-1}_{1\lambda,(1)} T^{-1/2} [\textstyle\sum_{1\lambda} x_{t(1)}x'_{t(1)} - \lambda \sum_{t=1}^T x_{t(1)}x'_{t(1)}] \widehat{Q}^{-1}_{2\lambda,(1)} Q_{(12)}\theta_{(2)} + o_P(1)$$
$$\equiv I + II - III.$$
$$I + II = Q^{-1}_{(1)}\, T^{-1/2} \, [\textstyle\sum_{1\lambda}[k_t \; L_t]/\lambda - \sum_{2\lambda}[k_t \; L_t]/(1-\lambda)] \, \mathbf{vec}(1, \theta_{(2)}) + o_P(1)$$
$$= \tfrac{1}{\lambda(1-\lambda)} Q^{-1}_{(1)} T^{-1/2} \left[\textstyle\sum_{1\lambda}[k_t \; L_t] - \lambda \sum_{t=1}^T [k_t \; L_t]\right] \mathbf{vec}(1, \theta_{(2)}) + o_P(1).$$

Let $\delta = Q^{-1}_{(1)} Q_{(12)}\theta_{(2)}$, and note:

$$III = \tfrac{1}{\lambda(1-\lambda)} Q^{-1}_{(1)} T^{-1/2} [\textstyle\sum_{1\lambda} x_{t(1)}x'_{t(1)} - \lambda \sum_{t=1}^T x_{t(1)}x'_{t(1)}] Q^{-1}_{(1)} Q_{(12)}\theta_{(2)} + o_P(1)$$
$$= \tfrac{1}{\lambda(1-\lambda)} Q^{-1}_{(1)} T^{-1/2} \left(\textstyle\sum_{1\lambda} M_t - \lambda \sum_{t=1}^T M_t\right) \delta + o_P(1)$$
$$= \tfrac{1}{\lambda(1-\lambda)} Q^{-1}_{(1)} T^{-1/2} \left(\textstyle\sum_{1\lambda} M_t - \lambda \sum_{t=1}^T M_t\right) \delta + o_P(1).$$

With $s_t = [k_t \; L_t \; M_t]$, we have:

$$T^{1/2}(\widehat{\theta}_{1\lambda} - \widehat{\theta}_{2\lambda}) = \tfrac{1}{\lambda(1-\lambda)} Q^{-1}_{(1)} \left(\textstyle\sum_{1\lambda} s_t - \lambda \sum_{t=1}^T s_t\right) \mathbf{vec}(1, \theta_{(2)}, -\delta) + o_P(1).$$

By A3 and the FCLT, $T^{-1/2}\sum_{1\lambda} \mathbf{vec}(s_t) \Rightarrow \mathbf{H^{*1/2}} B^*_s(\lambda)$, and so $T^{-1/2}\sum_{1\lambda} s_t \Rightarrow \mathbf{H^{*1/2}} \mathcal{B}^*_{mat}(\lambda)$, where we denoted

$$\mathcal{B}^*_{mat}(\lambda) = (\mathcal{B}^*_{1:p_1}(\lambda), \mathcal{B}^*_{p_1+1:2p_1}(\lambda), \ldots, \mathcal{B}^*_{p_1p_2+1:p_1(p_2+1)}(\lambda), \mathcal{B}^*_{p_1(p_2+1)+1:p_1(p_2+1)+p_1}(\lambda), \ldots, \mathcal{B}^*_{p_1p+1:p_1(p+1)}(\lambda)),$$

so that $\mathbf{vec}(\mathcal{B}^*_{mat}(\lambda)) = \mathcal{B}^*_s(\lambda)$. Letting $\xi = \mathbf{vec}(1, \theta_{(2)}, \delta)$, we obtain:

$$T^{1/2}(\widehat{\theta}_{1\lambda} - \widehat{\theta}_{2\lambda}) \Rightarrow Q^{-1}_{(1)} \mathbf{H^{*1/2}} \left[B^*_{mat}(\lambda)/\lambda - (B^*_{mat}(1) - B^*_{mat}(\lambda))/(1-\lambda)\right] \mathbf{vec}(1, \theta_{(2)}, \delta)$$
$$= Q^{-1}_{(1)} \mathbf{H^{*1/2}} [B^*_{mat}(\lambda) - \lambda B^*_{mat}(1)]\xi / [\lambda(1-\lambda)]$$
$$= Q^{-1}_{(1)} \mathbf{H^{*1/2}} \left(\mathcal{B}^*_{1:p_1}(\lambda) + \textstyle\sum_{i=1}^{p_2} \mathcal{B}^*_{p_1 i+1:p_1(i+1)} \theta_{i(2)} + \sum_{i=1}^{p_1} \mathcal{B}^*_{p_1(p_2+1)+p_1(i-1)+1:p_1(p_2+1)+p_1 i}(\lambda)\delta_i\right)$$
$$= [Q^{-1}_{(1)} \otimes \xi'] \mathbf{H^{*1/2}} \mathcal{B}^*_{p_1(p+1)},$$

where $\theta_{i(2)}, \delta_i$ are the i^{th} elements of $\theta_{(2)}$, respectively δ.

Part (i). Recall that $\widehat{V}_{i\lambda} = (\sum_{i\lambda} x_{t(1)}x'_{t(1)})^{-1} \widehat{\Omega}_{i\lambda} (\sum_i x_{t(1)}x'_{t(1)})^{-1}$ and $\widehat{\Omega}_{i\lambda} = T^{-1}\sum_{i\lambda} \widehat{\epsilon}^2_t x_{t(1)}x'_{t(1)}$, for $i = 1, 2$. Since $\widehat{\epsilon}_t = \epsilon_t - x'_{t(1)}(\widehat{\theta}_{1\lambda} - \theta_{(1)}) + x'_{t(2)}\theta_{(2)}$, $\widehat{\epsilon}^2_t = \epsilon^2_t + (\widehat{\theta}_{1\lambda} - \theta_{(1)})' x_{t(1)}x'_{t(1)}(\widehat{\theta}_{1\lambda} - \theta_{(1)}) + \theta'_{(2)} x_{t(2)}x'_{t(2)}\theta_{(2)} - 2(\widehat{\theta}_{1\lambda} - \theta_{(1)})'\epsilon_t x_{t(1)} + 2\theta'_{(2)}\epsilon_t x_{t(2)} - 2(\widehat{\theta}_{1\lambda} - \theta_{(1)})' x_{t(1)}x'_{t(2)}\theta_{(2)}$ and so:

$$\widehat{\Omega}_{1\lambda} = T^{-1}\textstyle\sum_{1\lambda} \widehat{\epsilon}^2_t x_{t(1)}x'_{t(1)} = T^{-1}\sum_{1\lambda} \epsilon^2_t x_{t(1)}x'_{t(1)} + T^{-1}\sum_{1\lambda}[x'_{t(1)}(\widehat{\theta}_{1\lambda} - \theta_{(1)})]^2 x_{t(1)}x'_{t(1)}$$
$$+ T^{-1}\textstyle\sum_{1\lambda}[\theta'_{(2)} x_{t(2)}]^2 x_{t(1)}x'_{t(1)} - 2T^{-1}\sum_{1\lambda}(\widehat{\theta}_{1\lambda} - \theta_{(1)})'\epsilon_t x_{t(1)} x_{t(1)}x'_{t(1)}$$
$$+ 2T^{-1}\textstyle\sum_{1\lambda} \epsilon_t x'_{t(2)}\theta_{(2)} x_{t(1)}x'_{t(1)}$$
$$- 2T^{-1}\textstyle\sum_{1\lambda}(\widehat{\theta}_{1\lambda} - \theta_{(1)})' x_{t(1)}x'_{t(2)}\theta_{(2)} x_{t(1)}x'_{t(1)}$$
$$= IV + V + VI - VII + VIII - IX.$$

Partition $\Omega^* = \begin{bmatrix} \Omega^*_{kk} & \Omega^*_{k\ell} & \Omega^*_{km} \\ \Omega^{*'}_{k\ell} & \Omega^*_{\ell\ell} & \Omega^*_{\ell m} \\ \Omega^{*'}_{km} & \Omega^{*'}_{\ell m} & \Omega^*_{mm} \end{bmatrix}$, such that $\Omega^*_{kk}, \Omega^*_{\ell\ell}, \Omega^*_{mm}$ are $p_1 \times p_1$, $(p_1p_2) \times (p_1p_2)$, and $p_1^2 \times p_1^2$ respectively. First, $IV = T^{-1}\sum_{1\lambda} \epsilon^2_t x_{t(1)}x'_{t(1)} = T^{-1}\sum_{1\lambda} k_t k'_t = \lambda\Omega^*_{kk} + o_P(1)$.

Also, from the above, $\widehat{\theta}_{1\lambda} - \theta_{(1)} = \delta + o_P(1)$, where $\delta = Q_{(1)}^{-1} Q_{(12)} \theta_{(2)}$. Thus, because of existence of fourth order moments of x_t by A3, it can be shown that:

$$\begin{aligned}
V &= T^{-1} \textstyle\sum_{1\lambda} [x'_{t(1)}(\widehat{\theta}_{1\lambda} - \theta_{(1)})]^2 x_{t(1)} x'_{t(1)} = T^{-1} \sum_{1\lambda} (x_{t(1)} x'_{t(1)} \delta)(\delta' x_{t(1)} x'_{t(1)}) + o_P(1) \\
&= T^{-1} \textstyle\sum_{1\lambda} (x_{t(1)} x'_{t(1)} - Q_{(1)}) \delta\delta' (x_{t(1)} x'_{t(1)} - Q_{(1)}) + o_P(1) + Q_{(1)} \delta\delta' T^{-1} \sum_{1\lambda} x_{t(1)} x'_{t(1)} \\
&+ T^{-1} \textstyle\sum_{1\lambda} x_{t(1)} x'_{t(1)} \delta\delta' Q_{(1)} - \lambda Q_{(1)} \delta\delta' Q_{(1)} + o_P(1) \\
&= T^{-1} \textstyle\sum_{1\lambda} (x_{t(1)} x'_{t(1)} - Q_{(1)}) \delta\delta' (x_{t(1)} x'_{t(1)} - Q_{(1)}) + \lambda Q_{(1)} \delta\delta' Q_{(1)} + o_P(1) \\
&= T^{-1} \textstyle\sum_{1\lambda} M_t \delta\delta' M'_t + \lambda Q_{(12)} \theta_{(2)} \theta'_{(2)} Q'_{(12)} + o_P(1).
\end{aligned}$$

We have $T^{-1} \sum_{1\lambda} \mathbf{vec}(M_t) \mathbf{vec}'(M'_t) \xrightarrow{p} \lambda \Omega^*_{mm}$ by A3. With $m_{t,ij}, \ell_{t,ij}$ the $(i,j)^{th}$ element of M_t, L_t and δ_i the i^{th} element of δ, we have:

$$\begin{aligned}
\delta' M'_t &= \mathbf{vec}[\textstyle\sum_{n=1}^{p_1} \delta_n m_{t,n1}, \ldots, \sum_{n=1}^{p_1} \delta_n m_{t,np_1}] \\
&= \mathbf{vec}[\delta' \{\mathbf{vec}[M_t]\}_{1:p_1}, \ldots, \delta' \{\mathbf{vec}[M_t]\}_{(p_1^2 - p_1 + 1):p_1^2}] = \mathbf{vec}'[M_t][I_{p_1} \otimes \delta], \\
M_t \delta &= [\textstyle\sum_{n=1}^{p_1} \delta_n M_{t,1n}, \ldots, \sum_{n=1}^{p_1} \delta_n M_{t,p_1 n}] = [I_{p_1} \otimes \delta'] \mathbf{vec}[M'_t], \\
L_t \theta_{(2)} &= [I_{p_1} \otimes \theta'_{(2)}] \mathbf{vec}[L'_t], \\
\theta'_{(2)} L'_t &= \mathbf{vec}'[L_t][I_{p_1} \otimes \theta_{(2)}].
\end{aligned}$$

It follows that:

$$\begin{aligned}
M_t \delta\delta' M'_t &= [I_{p_2} \otimes \delta'] \mathbf{vec}[M'_t] \mathbf{vec}'[M_t][I_{p_1} \otimes \delta], \\
V &= T^{-1} \textstyle\sum_{1\lambda} M_t \delta\delta' M'_t + \lambda Q_{(12)} \theta_{(2)} \theta'_{(2)} Q'_{(12)} + o_P(1) \\
&= \lambda [I_{p_1} \otimes \delta'] \Omega^*_{mm} [I_{p_1} \otimes \delta] + \lambda Q_{(12)} \theta_{(2)} \theta'_{(2)} Q'_{(12)} + o_P(1).
\end{aligned}$$

Similarly, it follows that:

$$\begin{aligned}
VI &= T^{-1} \textstyle\sum_{1\lambda} x_{t(1)} x'_{t(2)} \theta_{(2)} \theta'_{(2)} x_{t(2)} x'_{t(1)} = T^{-1} \sum_{1\lambda} (x_{t(1)} x'_{t(2)} - Q_{(12)}) \theta_{(2)} \theta'_{(2)} (x_{t(2)} x'_{t(1)} - Q'_{(12)}) \\
&+ Q_{(12)} T^{-1} \textstyle\sum_{1\lambda} \theta_{(2)} \theta'_{(2)} x_{t(2)} x'_{t(1)} + T^{-1} \sum_{1\lambda} x_{t(1)} x'_{t(2)} \theta_{(2)} \theta'_{(2)} Q'_{(12)} - \lambda Q_{(12)} \theta_{(2)} \theta'_{(2)} Q'_{(12)} \\
&= T^{-1} \textstyle\sum_{1\lambda} L_t \theta_{(2)} \theta'_{(2)} L'_t + \lambda Q_{(12)} \theta_{(2)} \theta'_{(2)} Q'_{(12)} + o_P(1) \\
&= \lambda [I_{p_1} \otimes \theta'_{(2)}] \Omega^*_{\ell\ell} [I_{p_1} \otimes \theta_{(2)}] + \lambda Q_{(12)} \theta_{(2)} \theta'_{(2)} Q'_{(12)} + o_P(1).
\end{aligned}$$

Also,

$$\begin{aligned}
VII &= 2T^{-1} \textstyle\sum_{1\lambda} (\widehat{\theta}_{1\lambda} - \theta_{(1)})' \epsilon_t x_{t(1)} x_{t(1)} x'_{t(1)} = 2T^{-1} \sum_{1\lambda} (\delta' x_{t(1)} \epsilon_t)\ x_{t(1)} x'_{t(1)} + o_P(1) \\
&= 2T^{-1} \textstyle\sum_{1\lambda} (\delta' x_{t(1)} \epsilon_t)\ (x_{t(1)} x'_{t(1)} - Q_{(1)}) + 2T^{-1} \sum_{1\lambda} (\delta' x_{t(1)} \epsilon_t) Q_{(1)} + o_P(1) \\
&= 2T^{-1} \textstyle\sum_{1\lambda} x_{t(1)} \epsilon_t \delta' (x_{t(1)} x'_{t(1)} - Q_{(1)}) + o_P(1) = 2T^{-1} \sum_{1\lambda} k_t \delta' M'_t + o_P(1) \\
&= T^{-1} \textstyle\sum_{1\lambda} k_t \delta' M'_t + T^{-1} \sum_{1\lambda} \delta' M'_t k_t + o_P(1) = \lambda \Omega^*_{km} [I_{p_1} \otimes \delta] + \lambda [I_{p_1} \otimes \delta'] \Omega^{*\prime}_{km} + o_P(1),
\end{aligned}$$

$$
\begin{aligned}
VIII &= 2T^{-1}\textstyle\sum_{1\lambda}\theta'_{(2)}\epsilon_t x_{t(2)}x_{t(1)}x'_{t(1)} = 2T^{-1}\sum_{1\lambda}x_{t(1)}\epsilon_t\theta'_{(2)}x_{t(2)}x'_{t(1)}\\
&= 2T^{-1}\textstyle\sum_{1\lambda}x_{t(1)}\epsilon_t\theta'_{(2)}(x_{t(2)}x'_{t(1)} - Q'_{(12)}) + o_P(1) = 2T^{-1}\sum_{1\lambda}k_t\theta'_{(2)}L'_t + o_P(1) =\\
&= \lambda\Omega^*_{k\ell}[I_{p_1}\otimes\theta_{(2)}] + \lambda[I_{p_1}\otimes\theta'_{(2)}]\Omega^{*'}_{k\ell} + o_P(1),\\
IX &= 2T^{-1}\textstyle\sum_{1\lambda}(\widehat{\theta}_{1\lambda}-\theta_{(1)})'x_{t(1)}x'_{t(2)}\theta_{(2)}x_{t(1)}x'_{t(1)} = 2T^{-1}\sum_{1\lambda}x_{t(1)}x'_{t(1)}\delta\theta'_{(2)}x_{t(2)}x'_{t(1)} + o_P(1)\\
&= 2T^{-1}\textstyle\sum_{1\lambda}M_t\delta\theta'_{(2)}L'_t + 2Q_{(1)}\delta\theta'_{(2)}T^{-1}\sum_{1\lambda}x_{t(2)}x'_{t(1)} + 2T^{-1}\sum_{1\lambda}x_{t(1)}x'_{t(1)}\delta\theta'_{(2)}Q'_{(12)}\\
&\quad - 2\lambda Q_{(1)}\delta\theta'_{(2)}Q'_{(12)} + o_P(1)\\
&= \lambda[I_{p_1}\otimes\delta']\Omega^{*'}_{\ell m}[I_{p_1}\otimes\theta_{(2)}] + \lambda[I_{p_1}\otimes\theta'_{(2)}]\Omega^*_{\ell m}[I_{p_1}\otimes\delta] + 2\lambda Q_{(1)}\delta\theta'_{(2)}Q'_{(12)} + o_P(1).
\end{aligned}
$$

Putting all of the above together,

$$
\begin{aligned}
\widehat{\Omega}^*_{1\lambda} &= \lambda\Big\{\Omega^*_{kk} + [I_{p_1}\otimes\delta']\Omega^*_{mm}[I_{p_1}\otimes\delta] + [I_{p_1}\otimes\theta'_{(2)}]\Omega^*_{\ell\ell}[I_{p_1}\otimes\theta_{(2)}] - \Omega^*_{km}[I_{p_1}\otimes\delta] - [I_{p_1}\otimes\delta']\Omega^{*'}_{km}\\
&\quad + \Omega^*_{k\ell}[I_{p_1}\otimes\theta_{(2)}] + [I_{p_1}\otimes\theta'_{(2)}]\Omega^{*'}_{k\ell} - [I_{p_1}\otimes\delta']\Omega^{*'}_{\ell m}[I_{p_1}\otimes\theta_{(2)}] - [I_{p_1}\otimes\theta'_{(2)}]\Omega^*_{\ell m}[I_{p_1}\otimes\delta]\Big\} + o_P(1)\\
&= \lambda\left\{[I_{p_1}\otimes\xi']\ \Omega^*\ [I_{p_1}\otimes\xi]\right\} + o_P(1)\\
\widehat{\mathbf{V}}_{1\lambda} &= \tfrac{1}{\lambda}Q^{-1}_{(1)}\left\{[I_{p_1}\otimes\xi']\ \Omega^*\ [I_{p_1}\otimes\xi]\right\}Q^{-1}_{(1)} + o_P(1) = \tfrac{1}{\lambda}\Big\{[Q^{-1}_{(1)}\otimes\xi']\ \Omega^*\ [Q^{-1}_{(1)}\otimes\xi]\Big\} + o_P(1)\\
\widehat{\mathbf{V}}_{2\lambda} &= \tfrac{1}{1-\lambda}\Big\{[Q^{-1}_{(1)}\otimes\xi']\ \Omega^*\ [Q^{-1}_{(1)}\otimes\xi]\Big\} + o_P(1)\\
\widehat{\mathbf{V}}_{\lambda} &= \tfrac{1}{\lambda(1-\lambda)}\Big\{[Q^{-1}_{(1)}\otimes\xi']\ \Omega^*\ [Q^{-1}_{(1)}\otimes\xi]\Big\} + o_P(1).
\end{aligned}
$$

Hence,

$$
CM^*_T \Rightarrow \sup_\lambda\left\{\tfrac{1}{\lambda(1-\lambda)}\mathcal{B}^{*'}_{p_1(p+1)}(\lambda)\ A\ \mathcal{B}^*_{p_1(p+1)}(\lambda)\right\},
$$

with $A = \mathbf{H}^{*1/2'}[Q^{-1}_{(1)}\otimes\xi]\Big\{[Q^{-1}_{(1)}\otimes\xi']\ \Omega^*\ [Q^{-1}_{(1)}\otimes\xi]\Big\}^{-1}[Q^{-1}_{(1)}\otimes\xi']\mathbf{H}^{*1/2}$.

Part (ii). In this case,

$$
\begin{aligned}
\widehat{\mathbf{v}}_{\epsilon,1\lambda} &= T^{-1}_{1\lambda}\textstyle\sum_{1\lambda}\epsilon_t^2 + (\widehat{\theta}_{1\lambda}-\theta_{(1)})'\ T^{-1}_{1\lambda}\sum_{1\lambda}x_{t(1)}x'_{t(1)}(\widehat{\theta}_{1\lambda}-\theta_{(1)}) + \theta'_{(2)}\ T^{-1}_{1\lambda}\sum_{1\lambda}x_{t(2)}x'_{t(2)}\theta_{(2)}\\
&\quad - 2(\widehat{\theta}_{1\lambda}-\theta_{(1)})'\ T^{-1}_{1\lambda}\textstyle\sum_{1\lambda}\epsilon_t x_{t(1)} + 2\theta'_{(2)}\ T^{-1}_{1\lambda}\sum_{1\lambda}\epsilon_t x_{t(2)} - 2(\widehat{\theta}_{1\lambda}-\theta_{(1)})'\ T^{-1}_{1\lambda}\sum_{1\lambda}x_{t(1)}x'_{t(2)}\theta_{(2)}\\
&= \sigma^2_\epsilon + \delta' Q_{(1)}\delta + \theta'_{(2)}Q_{(2)}\theta_{(2)} - 2\delta' Q_{(12)}\theta_{(2)} + o_P(1)\\
&= \sigma^2_\epsilon + \theta'_{(2)}Q'_{(12)}Q^{-1}_{(1)}Q_{(12)}\theta_{(2)} + \theta'_{(2)}Q_{(2)}\theta_{(2)} - 2\theta'_{(2)}Q'_{(12)}Q^{-1}_{(1)}Q_{(12)}\theta_{(2)} + o_P(1)\\
&= \sigma^2_\epsilon - \theta'_{(2)}Q'_{(12)}Q^{-1}_{(1)}Q_{(12)}\theta_{(2)} + \theta'_{(2)}Q_{(2)}\theta_{(2)} + o_P(1)\\
&= \lambda\sigma^2_\epsilon + \theta'_{(2)}[Q_{(2)} - Q'_{(12)}Q^{-1}_{(1)}Q_{(12)}]\theta_{(2)} + o_P(1)\\
\widehat{\mathbf{V}}_{1\lambda} &= \Big\{\sigma^2_\epsilon + \theta'_{(2)}[Q_{(2)} - Q'_{(12)}Q^{-1}_{(1)}Q_{(12)}]\theta_{(2)}\Big\}Q^{-1}_{(1)}/\lambda + o_P(1)\\
\widehat{\mathbf{V}}_{2\lambda} &= \Big\{\sigma^2_\epsilon + \theta'_{(2)}[Q_{(2)} - Q'_{(12)}Q^{-1}_{(1)}Q_{(12)}]\theta_{(2)}\Big\}Q^{-1}_{(1)}/(1-\lambda) + o_P(1)\\
\widehat{\mathbf{V}}_{\lambda} &= \tfrac{1}{\lambda(1-\lambda)}\Big\{\sigma^2_\epsilon + \theta'_{(2)}[Q_{(2)} - Q'_{(12)}Q^{-1}_{(1)}Q_{(12)}]\theta_{(2)}\Big\}Q^{-1}_{(1)} + o_P(1) = \tfrac{\nu}{\lambda(1-\lambda)}Q^{-1}_{(1)} + o_P(1).
\end{aligned}
$$

So, CM^*_T weakly converges to:

$$
= \sup_\lambda\left\{\tfrac{1}{\nu\lambda(1-\lambda)}\mathcal{B}^{*'}_{p_1(p+1)}(\lambda)\mathbf{H}^{*1/2'}\{Q^{-1}_{(1)}\otimes\ (\xi\xi')\}\mathbf{H}^{*1/2}\mathcal{B}^*_{p_1(p+1)}(\lambda)\right\}.
$$

□