Further Results on Size and Power of Heteroskedasticity and Autocorrelation Robust Tests, with an Application to Trend Testing

Pötscher, Benedikt M. and Preinerstorfer, David

2017

Further Results on Size and Power of Heteroskedasticity and Autocorrelation Robust Tests, with an Application to Trend Testing*

Benedikt M. Pötscher and David Preinerstorfer
Department of Statistics, University of Vienna
ECARES, Université libre de Bruxelles

August 2017

Abstract

We complement the theory developed in Preinerstorfer and Pötscher (2016) with further finite sample results on size and power of heteroskedasticity and autocorrelation robust tests. These allows us, in particular, to show that the sufficient conditions for the existence of size-controlling critical values recently obtained in Pötscher and Preinerstorfer (2016) are often also necessary. We furthermore apply the results obtained to tests for hypotheses on deterministic trends in stationary time series regressions, and find that many tests currently used are strongly size-distorted.

1 Introduction

Heteroskedasticity and autocorrelation robust tests in regression models suggested in the literature (e.g., tests based on the covariance estimators in Newey and West (1987, 1994), Andrews (1991), and Andrews and Monahan (1992), or tests in Kiefer et al. (2000), Kiefer and Vogelsang (2002a,b, 2005)) often suffer from substantial size distortions or power deficiencies. This has been repeatedly documented in simulation studies, and has been explained analytically by the theory developed in Preinerstorfer and Pötscher (2016). Given a test for an affine restriction on the regression coefficient vector, the results in Preinerstorfer and Pötscher (2016) provide several sufficient conditions that imply size equal to one, or severe biasedness of the test (resulting in low power in certain regions of the alternative). The central object in that theory is the set of possible covariance matrices of the regression errors, i.e., the covariance model, and, in particular, its set

*Financial support of the second author by the Danish National Research Foundation (Grant DNRF 78, CREATES) is gratefully acknowledged. Address correspondence to Benedikt Pötscher, Department of Statistics, University of Vienna, A-1090 Oskar-Morgenstern Platz 1. E-Mail: benedikt.poetscher@univie.ac.at.

of concentration spaces. Concentration spaces are defined as the column spaces of all singular matrices belonging to the boundary of the covariance model (cf. Definition 1 in Preinerstorfer and Pötscher (2016)). In Preinerstorfer and Pötscher (2016) it was shown that the position of the concentration spaces relative to the rejection region of the test often lets one deduce whether size distortions or power problems occur. Loosely speaking, if a concentration space lies in the "interior" of the rejection region, the test has size equal to one, whereas if a concentration space lies in the "exterior" (the "interior" of the complement) of the rejection region, the test is biased and has nuisance-minimal power equal to zero.[1] These interiority (exteriority) conditions can be formulated in terms of test statistics and critical values, can be easily checked in practice, and have been made explicit in Preinerstorfer and Pötscher (2016) at different levels of generality concerning the test statistic and the covariance model (cf. their Corollary 5.17, Theorem 3.3, Theorem 3.12, Theorem 3.15, and Theorem 4.2 for more details).

Given a test statistic, the results of Preinerstorfer and Pötscher (2016) just mentioned – if applicable – all lead to implications of the following type: (i) size equals one for any choice of critical value (e.g., testing a zero restriction on the mean of a stationary AR(1) time series falls under this case); or (ii) all critical values smaller than a certain real number (depending on observable quantities only) lead to a test with size one. While implication (i) certainly rules out the existence of a size-controlling critical value, implication (ii) does not, because it only makes a statement about a certain range of critical values. Hence, the question when a size-controlling critical value actually exists has not sufficiently been answered in Preinerstorfer and Pötscher (2016). Focusing exclusively on size control, Pötscher and Preinerstorfer (2016) recently developed conditions under which size can be controlled at any level.[2] It turns out that these conditions can, in general, not be formulated in terms of concentration spaces of the covariance model alone. Rather, they are conditions involving a different, but related, set $\mathbb{J}$, say, of linear spaces obtained from the covariance model. This set $\mathbb{J}$ consists of nontrivial projections of concentration spaces as well as of spaces which might be regarded as "higher-order" concentration spaces (cf. Section 5.3.2 of Pötscher and Preinerstorfer (2016) for a detailed discussion). Again, the conditions in Pötscher and Preinerstorfer (2016) do not depend on unobservable quantities, and hence can be checked by the practitioner. Pötscher and Preinerstorfer (2016) also provide algorithms for the computation of size-controlling critical values, which are implemented in the R-package **acrt** (Preinerstorfer (2016)).

Summarizing we arrive at the following situation: Preinerstorfer and Pötscher (2016) provide sufficient conditions for non-existence of size-controlling critical values in terms of the set of concentration spaces of a covariance model, whereas Pötscher and Preinerstorfer (2016) provide sufficient conditions for the existence of size-controlling critical values formulated in terms of

[1] The situation is a bit more complex. For example, sometimes a modification of the rejection region, which leaves the rejection probabilities unchanged, is required in order to enforce the interiority (exteriority) condition; see Theorem 5.7 in Preinerstorfer and Pötscher (2016).

[2] We note that, apart from the results mentioned before, Preinerstorfer and Pötscher (2016) also contains results that ensure size control (and positive infimal power). The scope of these results is, however, substantially more narrow than the scope of the results in Pötscher and Preinerstorfer (2016).

a different set of linear spaces derived from the covariance model. Combining the results in Preinerstorfer and Pötscher (2016) and Pötscher and Preinerstorfer (2016) does in general *not* result in necessary and sufficient conditions for the existence of size-controlling critical values. [This is partly due to the fact that different sets of linear spaces associated with the covariance model are used in these two papers.] Rather, there remains a range of problems for which the existence of size-controlling critical values can be neither disproved by the results in Preinerstorfer and Pötscher (2016) nor proved by the results in Pötscher and Preinerstorfer (2016).

In the present paper we close the "gap" between the negative results in Preinerstorfer and Pötscher (2016) on the one hand, and the positive results in Pötscher and Preinerstorfer (2016) on the other hand. We achieve this by obtaining new negative results that are typically more general than the ones in Preinerstorfer and Pötscher (2016). Instead of directly working with concentration spaces of a given covariance model (as in Preinerstorfer and Pötscher (2016)) our main strategy is essentially as follows: We first show that size properties of (invariant) tests are preserved when passing from the given covariance model to a suitably constructed auxiliary covariance model which has the property that the concentration spaces of this auxiliary covariance model coincide with the set $\mathbb{J}$ of linear spaces derived from the initial covariance model (as used in the results of Pötscher and Preinerstorfer (2016)). Then we apply results in Preinerstorfer and Pötscher (2016) to the concentration spaces of the auxiliary covariance model to obtain a necessary condition for the existence of size-controlling critical values. [This result is first formulated for arbitrary covariance models, and is then further specialized to the case of stationary autocorrelated errors.] The so-obtained new result now allows us to prove that the conditions developed in Pötscher and Preinerstorfer (2016) for the possibility of size control are not only sufficient, but are – under certain (weak) conditions on the test statistic – also necessary. Additionally, we also study power properties and obtain results showing that frequently (but not always) a critical value leading to size control will lead to low power in certain regions of the alternative.

Obtaining results for the class of problems inaccessible by the results of Preinerstorfer and Pötscher (2016) and Pötscher and Preinerstorfer (2016) is not only theoretically satisfying. It is also practically important as this class contains empirically relevant testing problems: As a further contribution we thus apply our results to the important problem of testing hypotheses on polynomial or cyclical trends in stationary time series, the former being our main focus. Testing for trends certainly is an important problem (not only) in economics, and has received a great amount of attention in the literature. Using our new results we can prove that many tests currently in use (e.g., conventional tests based on long-run-variance estimators, or more specialized tests as suggested in Vogelsang (1998) and Bunzel and Vogelsang (2005)) suffer from severe size problems if the covariance model is not too small (that is, contains all covariance matrices of stationary autoregressive processes of order two or a slight enlargement of that set). Furthermore, our results show that this problem can not be resolved by increasing the critical values used.

The structure of the article is as follows: Section 2 introduces the framework and some notation. In Section 3 we present results concerning size properties of nonsphericity-corrected F-type tests. This is done on two levels of generality: In Subsection 3.1 we present results for general covariance models, whereas in Subsection 3.2 we present results for covariance models obtained from stationary autocorrelated errors. In these two sections it is also shown that the conditions for size control obtained in Theorems 3.2, 6.2, 6.5, and Corollary 5.6 of Pötscher and Preinerstorfer (2016) are not only sufficient but are also necessary in important scenarios. In Section 4 we present negative results concerning the power of tests based on size-controlling critical values. Finally, in Section 5 we discuss consequences of our results for testing restrictions on coefficients of polynomial and cyclical regressors. All proofs as well as some auxiliary results are given in the appendices.

2 Framework

2.1 The model and basic notation

Consider the linear regression model

$$\mathbf{Y} = X\beta + \mathbf{U}, \tag{1}$$

where X is a (real) nonstochastic regressor (design) matrix of dimension $n \times k$ and where $\beta \in \mathbb{R}^k$ denotes the unknown regression parameter vector. We always assume $\operatorname{rank}(X) = k$ and $1 \leq k < n$. We furthermore assume that the $n \times 1$ disturbance vector $\mathbf{U} = (\mathbf{u}_1, \ldots, \mathbf{u}_n)'$ is normally distributed with mean zero and unknown covariance matrix $\sigma^2\Sigma$, where Σ varies in a prescribed (nonempty) set $\mathfrak{C}$ of symmetric and positive definite $n \times n$ matrices and where $0 < \sigma^2 < \infty$ holds (σ always denoting the positive square root).[3] The set $\mathfrak{C}$ will be referred to as the covariance model. We shall always assume that $\mathfrak{C}$ allows σ^2 and Σ to be uniquely determined from $\sigma^2\Sigma$.[4] [This entails virtually no loss of generality and can always be achieved, e.g., by imposing some normalization assumption on the elements of $\mathfrak{C}$ such as normalizing the first diagonal element of Σ or the norm of Σ to one, etc.] The leading case will concern the situation where $\mathfrak{C}$ results from the assumption that the elements $\mathbf{u}_1, \ldots, \mathbf{u}_n$ of the $n \times 1$ disturbance vector $\mathbf{U}$ are distributed like consecutive elements of a zero mean weakly stationary Gaussian process with an unknown spectral density, but allowing for more general covariance models is useful.

The linear model described in (1) together with the Gaussianity assumption on $\mathbf{U}$ induces a collection of distributions on the Borel-sets of $\mathbb{R}^n$, the sample space of $\mathbf{Y}$. Denoting a Gaussian probability measure with mean $\mu \in \mathbb{R}^n$ and (possibly singular) covariance matrix A by $P_{\mu,A}$, the

[3] Since we are concerned with finite-sample results only, the elements of $\mathbf{Y}$, X, and $\mathbf{U}$ (and even the probability space supporting $\mathbf{Y}$ and $\mathbf{U}$) may depend on sample size n, but this will not be expressed in the notation. Furthermore, the obvious dependence of $\mathfrak{C}$ on n will also not be shown in the notation.

[4] That is, $\mathfrak{C}$ has the property that $\Sigma \in \mathfrak{C}$ implies $\delta\Sigma \notin \mathfrak{C}$ for every $\delta \neq 1$.

induced collection of distributions is then given by

$$\left\{P_{\mu,\sigma^2\Sigma} : \mu \in \operatorname{span}(X), 0 < \sigma^2 < \infty, \Sigma \in \mathfrak{C}\right\}. \tag{2}$$

Since every $\Sigma \in \mathfrak{C}$ is positive definite by assumption, each element of the set in the previous display is absolutely continuous with respect to (w.r.t.) Lebesgue measure on $\mathbb{R}^n$.

We shall consider the problem of testing a linear (better: affine) hypothesis on the parameter vector $\beta \in \mathbb{R}^k$, i.e., the problem of testing the null $R\beta = r$ against the alternative $R\beta \neq r$, where R is a $q \times k$ matrix always of rank $q \geq 1$ and $r \in \mathbb{R}^q$. Set $\mathfrak{M} = \operatorname{span}(X)$. Define the affine space

$$\mathfrak{M}_0 = \{\mu \in \mathfrak{M} : \mu = X\beta \text{ and } R\beta = r\}$$

and let

$$\mathfrak{M}_1 = \{\mu \in \mathfrak{M} : \mu = X\beta \text{ and } R\beta \neq r\}.$$

Adopting these definitions, the above testing problem can then be written more precisely as

$$H_0 : \mu \in \mathfrak{M}_0,\ 0 < \sigma^2 < \infty,\ \Sigma \in \mathfrak{C} \quad \text{vs.} \quad H_1 : \mu \in \mathfrak{M}_1,\ 0 < \sigma^2 < \infty,\ \Sigma \in \mathfrak{C}. \tag{3}$$

We also define $\mathfrak{M}_0^{lin}$ as the linear space parallel to $\mathfrak{M}_0$, i.e., $\mathfrak{M}_0^{lin} = \mathfrak{M}_0 - \mu_0$ for some $\mu_0 \in \mathfrak{M}_0$. Obviously, $\mathfrak{M}_0^{lin}$ does not depend on the choice of $\mu_0 \in \mathfrak{M}_0$. The previously introduced concepts and notation will be used throughout the paper.

The assumption of Gaussianity is made mainly in order not to obscure the structure of the problem by technicalities. Substantial generalizations away from Gaussianity are possible exactly in the same way as the extensions discussed in Section 5.5 of Preinerstorfer and Pötscher (2016); see also Section 7 of Pötscher and Preinerstorfer (2016). The assumption of nonstochastic regressors can be relaxed somewhat: If X is random and, e.g., independent of $\mathbf{U}$, the results of the paper apply after one conditions on X. For arguments supporting conditional inference see, e.g., Robinson (1979).

We next collect some further terminology and notation used throughout the paper. A (non-randomized) *test* is the indicator function of a Borel-set W in $\mathbb{R}^n$, with W called the corresponding *rejection region*. The *size* of such a test (rejection region) is the supremum over all rejection probabilities under the null hypothesis H_0, i.e.,

$$\sup_{\mu \in \mathfrak{M}_0} \sup_{0 < \sigma^2 < \infty} \sup_{\Sigma \in \mathfrak{C}} P_{\mu,\sigma^2\Sigma}(W).$$

Throughout the paper we let $\hat{\beta}_X(y) = (X'X)^{-1} X'y$, where X is the design matrix appearing in (1) and $y \in \mathbb{R}^n$. The corresponding ordinary least squares (OLS) residual vector is denoted by $\hat{u}_X(y) = y - X\hat{\beta}_X(y)$. If it is clear from the context which design matrix is being used, we shall drop the subscript X from $\hat{\beta}_X(y)$ and $\hat{u}_X(y)$ and shall simply write $\hat{\beta}(y)$ and $\hat{u}(y)$. We use Pr as a generic symbol for a probability measure. Lebesgue measure on the Borel-sets of $\mathbb{R}^n$

will be denoted by $\lambda_{\mathbb{R}^n}$, whereas Lebesgue measure on an affine subspace $\mathcal{A}$ of $\mathbb{R}^n$ (but viewed as a measure on the Borel-sets of $\mathbb{R}^n$) will be denoted by $\lambda_{\mathcal{A}}$, with zero-dimensional Lebesgue measure being interpreted as point mass. The set of real matrices of dimension $l \times m$ is denoted by $\mathbb{R}^{l \times m}$ (all matrices in the paper will be real matrices). Let B' denote the transpose of a matrix $B \in \mathbb{R}^{l \times m}$ and let $\operatorname{span}(B)$ denote the subspace in $\mathbb{R}^l$ spanned by its columns. For a symmetric and nonnegative definite matrix B we denote the unique symmetric and nonnegative definite square root by $B^{1/2}$. For a linear subspace $\mathcal{L}$ of $\mathbb{R}^n$ we let $\mathcal{L}^{\perp}$ denote its orthogonal complement and we let $\Pi_{\mathcal{L}}$ denote the orthogonal projection onto $\mathcal{L}$. For an affine subspace $\mathcal{A}$ of $\mathbb{R}^n$ we denote by $G(\mathcal{A})$ the group of all affine transformations on $\mathbb{R}^n$ of the form $y \mapsto \delta(y-a)+a^*$ where $\delta \neq 0$ and a as well as a^* belong to $\mathcal{A}$. The j-th standard basis vector in $\mathbb{R}^n$ is written as $e_j(n)$. Furthermore, we let $\mathbb{N}$ denote the set of all positive integers. A sum (product, respectively) over an empty index set is to be interpreted as 0 (1, respectively). Finally, for a subset A of a topological space we denote by $\operatorname{cl}(A)$ the closure of A (w.r.t. the ambient space).

2.2 Classes of test statistics

The rejection regions we consider will be of the form $W = \{y \in \mathbb{R}^n : T(y) \geq C\}$, where the critical value C satisfies $-\infty < C < \infty$ and the test statistic T is a Borel-measurable function from $\mathbb{R}^n$ to $\mathbb{R}$. With the exception of Section 4, the results in the present paper will concern the class of nonsphericity-corrected F-type test statistics as defined in (28) of Section 5.4 in Preinerstorfer and Pötscher (2016) that satisfy Assumption 5 in that reference. For the convenience of the reader we recall the definition of this class of test statistics. We start with the following assumption, which is Assumption 5 in Preinerstorfer and Pötscher (2016):

Assumption 1. (i) Suppose we have estimators $\check{\beta} : \mathbb{R}^n \backslash N \to \mathbb{R}^k$ and $\check{\Omega} : \mathbb{R}^n \backslash N \to \mathbb{R}^{q \times q}$ that are well-defined and continuous on $\mathbb{R}^n \backslash N$, where N is a closed $\lambda_{\mathbb{R}^n}$-null set. Furthermore, $\check{\Omega}(y)$ is symmetric for every $y \in \mathbb{R}^n \backslash N$. (ii) The set $\mathbb{R}^n \backslash N$ is assumed to be invariant under the group $G(\mathfrak{M})$, i.e., $y \in \mathbb{R}^n \backslash N$ implies $\delta y + X\eta \in \mathbb{R}^n \backslash N$ for every $\delta \neq 0$ and every $\eta \in \mathbb{R}^k$. (iii) The estimators satisfy the equivariance properties $\check{\beta}(\delta y + X\eta) = \delta\check{\beta}(y) + \eta$ and $\check{\Omega}(\delta y + X\eta) = \delta^2\check{\Omega}(y)$ for every $y \in \mathbb{R}^n \backslash N$, for every $\delta \neq 0$, and for every $\eta \in \mathbb{R}^k$. (iv) $\check{\Omega}$ is $\lambda_{\mathbb{R}^n}$-almost everywhere nonsingular on $\mathbb{R}^n \backslash N$.

Nonsphericity-corrected F-type test statistics are now of the form

$$T(y) = \begin{cases} (R\check{\beta}(y) - r)'\check{\Omega}^{-1}(y)(R\check{\beta}(y) - r), & y \in \mathbb{R}^n \backslash N^*, \\ 0, & y \in N^*, \end{cases} \tag{4}$$

where $\check{\beta}$, $\check{\Omega}$, and N satisfy Assumption 1 and where $N^* = N \cup \left\{y \in \mathbb{R}^n \backslash N : \det \check{\Omega}(y) = 0\right\}$. We recall from Lemmata 5.15 and F.1 in Preinerstorfer and Pötscher (2016) that N^* is then a closed $\lambda_{\mathbb{R}^n}$-null set that is invariant under $G(\mathfrak{M})$, and that T is continuous on $\mathbb{R}^n \backslash N^*$ (and is obviously Borel-measurable on $\mathbb{R}^n$). Furthermore, T is $G(\mathfrak{M}_0)$-invariant, i.e., $T(\delta(y - \mu_0) + \mu_0') = T(y)$ holds for every $y \in \mathbb{R}^n$, every $\delta \neq 0$, every $\mu_0 \in \mathfrak{M}_0$, and for every $\mu_0' \in \mathfrak{M}_0$.

Remark 2.1. (*Important subclasses*) (i) Classical autocorrelation robust test statistics (e.g., those considered in Newey and West (1987), Andrews (1991) Sections 3-5, or in Kiefer et al. (2000), Kiefer and Vogelsang (2002a,b, 2005)) fall into this class: More precisely, denoting such a test statistic by T_w as in Pötscher and Preinerstorfer (2016), it follows that T_w is a nonsphericity-corrected F-type test statistics with Assumption 1 above being satisfied, provided only Assumptions 1 and 2 of Pötscher and Preinerstorfer (2016) hold. Here $\check{\beta}$ is given by the ordinary least squares estimator $\hat{\beta}$, $\check{\Omega}$ is given by $\hat{\Omega}_w$ defined in Section 3 of Pötscher and Preinerstorfer (2016), and $N = \emptyset$ holds (see Remark 5.16 in Pötscher and Preinerstorfer (2016)). Furthermore, $\check{\Omega} = \hat{\Omega}_w$ is then nonnegative definite on all of $\mathbb{R}^n$ (see Section 3.2 of Preinerstorfer and Pötscher (2016) or Section 3 of Pötscher and Preinerstorfer (2016)). We also recall from Section 5.3.1 of Pötscher and Preinerstorfer (2016) that in this case the set N^* can be shown to be a finite union of proper linear subspaces of $\mathbb{R}^n$.

(ii) Classical autocorrelation robust test statistics like T_w, but where the weights are now allowed to depend on the data (e.g., through data-driven bandwidth choice or through prewithening, etc.) as considered, e.g., in Newey and West (1994), Andrews (1991), and Andrews and Monahan (1992), also fall into the class of nonsphericity-corrected F-type tests under appropriate conditions (with the set N now typically being nonempty), see Preinerstorfer (2017) for details. The same is typically true for test statistics based on parametric long-run variance estimators or test statistics based on feasible generalized least squares (cf. Section 3.3 of Preinerstorfer and Pötscher (2016)).

(iii) A statement completely analogous to (i) above applies to the more general class of test statistics T_{GQ} discussed in Section 3.4B of Pötscher and Preinerstorfer (2016), provided Assumption 1 of Pötscher and Preinerstorfer (2016) is traded for the assumption that the weighting matrix $\mathcal{W}_n^*$ appearing in the definition of T_{GQ} is positive definite (and $\check{\Omega}$ is of course now as discussed in Section 3.4B of Pötscher and Preinerstorfer (2016)); see Remark 5.16 in Pötscher and Preinerstorfer (2016). Again, $\check{\Omega}$ is then nonnegative definite on all of $\mathbb{R}^n$ (see Section 3.2.1 of Preinerstorfer and Pötscher (2016)), $N = \emptyset$ holds, and N^* is a finite union of proper linear subspaces of $\mathbb{R}^n$ (see Section 5.3.1 of Pötscher and Preinerstorfer (2016)).

(iv) The (weighted) Eicker-test statistic $T_{E,\mathsf{W}}$ (cf. Eicker (1967)) as defined in Section 3.4C of Pötscher and Preinerstorfer (2016) is also a nonsphericity-corrected F-type test statistic with Assumption 1 above being satisfied, where $\check{\beta} = \hat{\beta}$, $\check{\Omega}$ is as in Section 3.4C of Pötscher and Preinerstorfer (2016), and $N = \emptyset$ holds. Again, $\check{\Omega}$ is nonnegative definite on all of $\mathbb{R}^n$, and $N^* = \operatorname{span}(X)$ holds (see Sections 3.4 and 5.3.1 of Pötscher and Preinerstorfer (2016)).

(v) Under the assumptions of Section 4 of Preinerstorfer and Pötscher (2016) (including Assumption 3 in that reference), usual heteroskedasticity-robust test statistics considered in the literature (see Long and Ervin (2000) for an overview) also fall into the class of nonsphericity-corrected F-type test statistics with Assumption 1 being satisfied. Again, the matrix $\check{\Omega}$ is then nonnegative definite everywhere, $N = \emptyset$ holds, and N^* is a finite union of proper linear subspaces of $\mathbb{R}^n$ (the latter following from Lemma 4.1 in Preinerstorfer and Pötscher (2016) combined with

Lemma 5.17 of Pötscher and Preinerstorfer (2016)).

We shall also encounter cases where $\check{\Omega}(y)$ may not be nonnegative definite for some values of $y \in \mathbb{R}^n \backslash N$. For these cases the following assumption, which is Assumption 7 in Preinerstorfer and Pötscher (2016), will turn out to be useful. For a discussion of this assumption see p. 314 of that reference.

Assumption 2. For every $v \in \mathbb{R}^q$ with $v \neq 0$ we have $\lambda_{\mathbb{R}^n}\left(\left\{y \in \mathbb{R}^n \backslash N^* : v'\check{\Omega}^{-1}(y)v = 0\right\}\right) = 0$.

3 Negative results on the size of nonsphericity-corrected F-type test statistics

3.1 A result for general covariance models

In this subsection we provide a negative result concerning the size of a class of nonsphericity-corrected F-type test statistics that is central to many of the results in the present paper. In particular, it allows us to show that the sufficient conditions for size control obtained in Pötscher and Preinerstorfer (2016) are often also necessary. The result complements negative results in Preinerstorfer and Pötscher (2016) and is obtained by combining Lemmata A.1 and A.3 in Appendix A with Corollary 5.17 of Preinerstorfer and Pötscher (2016). Its relationship to negative results in Preinerstorfer and Pötscher (2016) is further discussed in Appendix A.1. We recall the following definition from Pötscher and Preinerstorfer (2016).

Definition 3.1. Given a linear subspace $\mathcal{L}$ of $\mathbb{R}^n$ with $\dim(\mathcal{L}) < n$ and a covariance model $\mathfrak{C}$, we let $\mathcal{L}(\mathfrak{C}) = \{\mathcal{L}(\Sigma) : \Sigma \in \mathfrak{C}\}$, where $\mathcal{L}(\Sigma) = \Pi_{\mathcal{L}^\perp}\Sigma\Pi_{\mathcal{L}^\perp}/\|\Pi_{\mathcal{L}^\perp}\Sigma\Pi_{\mathcal{L}^\perp}\|$. Furthermore, we define

$$\mathbb{J}(\mathcal{L}, \mathfrak{C}) = \left\{\operatorname{span}(\bar{\Sigma}) : \bar{\Sigma} \in \operatorname{cl}(\mathcal{L}(\mathfrak{C})),\ \operatorname{rank}(\bar{\Sigma}) < n - \dim(\mathcal{L})\right\},$$

where the closure is here understood w.r.t. $\mathbb{R}^{n\times n}$. [The symbol $\|\cdot\|$ here denotes a norm on $\mathbb{R}^{n\times n}$. Note that $\mathbb{J}(\mathcal{L}, \mathfrak{C})$ does not depend on which norm is chosen.]

The space $\mathcal{L}$ figuring in this definition will always be an appropriately chosen subspace related to invariance properties of the tests under consideration. A leading case is when $\mathcal{L} = \mathfrak{M}_0^{lin}$. Loosely speaking, the linear spaces belonging to $\mathbb{J}(\mathcal{L}, \mathfrak{C})$ are either (nontrivial) projections of concentration spaces of the covariance model $\mathfrak{C}$ (in the sense of Preinerstorfer and Pötscher (2016)) on $\mathcal{L}^\perp$, or are what one could call "higher-order" concentration spaces. For a more detailed discussion see Section 5.3.2 of Pötscher and Preinerstorfer (2016).

Theorem 3.1. *Let $\mathfrak{C}$ be a covariance model. Let T be a nonsphericity-corrected F-type test statistic of the form (4) based on $\check{\beta}$ and $\check{\Omega}$ satisfying Assumption 1 with $N = \emptyset$. Furthermore, assume that $\check{\Omega}(y)$ is nonnegative definite for every $y \in \mathbb{R}^n$. If an $\mathcal{S} \in \mathbb{J}(\mathfrak{M}_0^{lin}, \mathfrak{C})$ satisfying $\mathcal{S} \subseteq \operatorname{span}(X)$ exists, then*

$$\sup_{\Sigma\in\mathfrak{C}} P_{\mu_0,\sigma^2\Sigma}(T \geq C) = 1 \tag{5}$$

holds for every critical value C, $-\infty < C < \infty$, for every $\mu_0 \in \mathfrak{M}_0$, and for every $\sigma^2 \in (0, \infty)$.

Remark 3.2. *(On the necessity of the sufficient conditions for size control in Pötscher and Preinerstorfer (2016))* (i) The preceding theorem can be used to show that the conditions for size control obtained in Corollary 5.6 of Pötscher and Preinerstorfer (2016) are not only sufficient, but are actually necessary in some important scenarios: Suppose T is as in Theorem 3.1, additionally satisfying $N^* = \operatorname{span}(X)$. Then Corollary 5.6 of Pötscher and Preinerstorfer (2016) (with $\mathcal{V} = \{0\}$, i.e., $\mathcal{L} = \mathfrak{M}_0^{lin}$) is applicable (since Lemma 5.15 in that reference shows that T satisfies the assumptions in that corollary with $N^\dagger = N^*$) and yields "$\mathcal{S} \nsubseteq \operatorname{span}(X)$ for all $\mathcal{S} \in \mathbb{J}(\mathfrak{M}_0^{lin}, \mathfrak{C})$" as a sufficient condition for the possibility of size control. [Application of this corollary actually delivers the just given condition but with $\mathcal{S} \nsubseteq \operatorname{span}(X)$ replaced by $\mu_0 + \mathcal{S} \nsubseteq \operatorname{span}(X)$ where $\mu_0 \in \mathfrak{M}_0$. Because $\mu_0 \in \mathfrak{M}_0 \subseteq \operatorname{span}(X)$ and the latter set is a linear space, both formulations are seen to be equivalent.] Theorem 3.1 above then implies that this sufficient condition is also necessary.

(ii) The discussion in (i) showing necessity of the sufficient condition for size control applies, in particular, to the (weighted) Eicker-test statistic $T_{E,\mathsf{W}}$ in view of Remark 2.1(iv) above. Note that $N^* = \operatorname{span}(X)$ is here always satisfied.

(iii) Next consider the classical autocorrelation robust test statistic T_w with Assumptions 1 and 2 of Pötscher and Preinerstorfer (2016) being satisfied. Then the discussion in (i) showing necessity of the sufficient condition for size control also applies to T_w in view of Remark 2.1(i) above, provided $N^* = \operatorname{span}(X)$ holds. While the relation $N^* = \operatorname{span}(X)$ need not always hold for T_w (see the discussion in Section 5.3.1 of Pötscher and Preinerstorfer (2016)), it holds for many combinations of restriction matrix R and design matrix X (in fact, it holds generically in many universes of design matrices as a consequence of Lemma A.3 of Pötscher and Preinerstorfer (2016)). Hence, for such combinations of R and X, the above mentioned sufficient conditions for size control are in fact necessary.

(iv) For test statistics T_{GQ} with positive definite weighting matrix $\mathcal{W}_n^*$ a statement completely analogous to (iii) above holds in view of Remark 2.1(iii). The same is true for heteroskedasticity-robust test statistics as discussed in Remark 2.1(v).

Remark 3.3. While Theorem 3.1 applies to any combination of test statistic T and covariance model $\mathfrak{C}$ as long as they satisfy the assumptions of the theorem, in a typical application the choice of the test statistic used will certainly be dictated by properties of the covariance model $\mathfrak{C}$ one maintains. For example, in case $\mathfrak{C}$ models stationary autocorrelated errors different test statistics will be employed than in the case where $\mathfrak{C}$ models heteroskedasticity.

3.2 Results for covariance models obtained from stationary autocorrelated errors

We next specialize the results of the preceding section to the case of stationary autocorrelated errors. i.e., to the case where the elements $\mathbf{u}_1, \ldots, \mathbf{u}_n$ of the $n \times 1$ disturbance vector $\mathbf{U}$ in model

(1) are distributed like consecutive elements of a zero mean weakly stationary Gaussian process with an unknown spectral density, which is not almost everywhere equal to zero. Consequently, the covariance matrix of the disturbance vector is positive definite and can be written as $\sigma^2\Sigma(f)$ where

$$\Sigma(f) = \left[\int_{-\pi}^{\pi} e^{-\iota(j-l)\omega} f(\omega) d\omega\right]_{j,l=1}^{n},$$

with f varying in $\mathfrak{F}$, a prescribed (nonempty) family of *normalized* (i.e., $\int_{-\pi}^{\pi} f(\omega)d\omega = 1$) spectral densities, and where $0 < \sigma^2 < \infty$ holds. Here ι denotes the imaginary unit. We define the associated covariance model via $\mathfrak{C}(\mathfrak{F}) = \{\Sigma(f) : f \in \mathfrak{F}\}$. Examples for the set $\mathfrak{F}$ are (i) $\mathfrak{F}_{\text{all}}$, the set of *all* normalized spectral densities, or (ii) $\mathfrak{F}_{\text{ARMA}(p,q)}$, the set of all normalized spectral densities corresponding to stationary autoregressive moving average models of order at most (p, q), or (iii) the set of normalized spectral densities corresponding to fractional autoregressive moving average models, etc. We shall write $\mathfrak{F}_{\text{AR}(p)}$ for $\mathfrak{F}_{\text{ARMA}(p,0)}$.

We need to recall some more concepts and notation from Pötscher and Preinerstorfer (2016); for background see this reference. Let $\omega \in [0, \pi]$ and let $s \geq 0$ be an integer. Define $E_{n,s}(\omega)$ as the $n \times 2$-dimensional matrix with j-th row equal to $(j^s \cos(j\omega), j^s \sin(j\omega))$. For $\mathcal{L}$, a linear subspace of $\mathbb{R}^n$ with $\dim(\mathcal{L}) < n$, let $\underline{\omega}(\mathcal{L})$ and $\underline{d}(\mathcal{L})$ be as in Definition 3.1 of Pötscher and Preinerstorfer (2016). Recall that the coordinates $\omega_j(\mathcal{L})$ of $\underline{\omega}(\mathcal{L})$ are precisely those angular frequencies $\omega \in [0, \pi]$ for which $\operatorname{span}(E_{n,0}(\omega)) \subseteq \mathcal{L}$ holds. And, for each j, the coordinate $d_j(\mathcal{L})$ of $\underline{d}(\mathcal{L})$ is the smallest natural number d such that $\operatorname{span}(E_{n,d}(\omega_j(\mathcal{L}))) \nsubseteq \mathcal{L}$. As shown in Pötscher and Preinerstorfer (2016), the vector $\underline{\omega}(\mathcal{L})$ (and hence $\underline{d}(\mathcal{L})$) has a finite number, $p(\mathcal{L})$ say, of components; it may be the case that $p(\mathcal{L}) = 0$, in which case $\underline{\omega}(\mathcal{L})$ and $\underline{d}(\mathcal{L})$ are 0-tupels. As in Pötscher and Preinerstorfer (2016), for d a natural number we define $\kappa(\omega, d) = 2d$ for $\omega \in (0, \pi)$ and $\kappa(\omega, d) = d$ for $\omega \in \{0, \pi\}$. Furthermore, we set $\kappa(\underline{\omega}(\mathcal{L}), \underline{d}(\mathcal{L})) = \sum \kappa(\omega_j(\mathcal{L}), d_j(\mathcal{L}))$ where the sum extends over $j = 1, \ldots, p(\mathcal{L})$, with the convention that this sum is zero if $p(\mathcal{L}) = 0$. We next define $\rho(\gamma, \mathcal{L})$ for $\gamma \in [0, \pi]$ as follows: $\rho(\gamma, \mathcal{L}) = d_j(\mathcal{L})$ in case $\gamma = \omega_j(\mathcal{L})$ for some $j = 1, \ldots, p(\mathcal{L})$, and $\rho(\gamma, \mathcal{L}) = 0$ else. For ease of notation we shall simply write $\rho(\gamma)$ for $\rho(\gamma, \mathfrak{M}_0^{lin})$.

The subsequent theorem specializes Theorem 3.1 to the case where $\mathfrak{C} = \mathfrak{C}(\mathfrak{F})$. For a definition of the collection $\mathbb{S}(\mathfrak{F}, \mathcal{L})$ of certain subsets of $[0, \pi]$ figuring in this theorem see Definition 6.3 of Pötscher and Preinerstorfer (2016).

Theorem 3.4. *Let $\mathfrak{F}$ be a nonempty set of normalized spectral densities, i.e., $\emptyset \neq \mathfrak{F} \subseteq \mathfrak{F}_{\text{all}}$. Let T be a nonsphericity-corrected F-type test statistic of the form (4) based on $\check{\beta}$ and $\check{\Omega}$ satisfying Assumption 1 with $N = \emptyset$. Furthermore, assume that $\check{\Omega}(y)$ is nonnegative definite for every $y \in \mathbb{R}^n$. Suppose there exists a linear subspace $\mathcal{S}$ of $\mathbb{R}^n$ that can be written as*

$$\mathcal{S} = \operatorname{span}\left(\Pi_{(\mathfrak{M}_0^{lin})^\perp}\left(E_{n,\rho(\gamma_1)}(\gamma_1), \ldots, E_{n,\rho(\gamma_p)}(\gamma_p)\right)\right) \quad \textit{for some} \quad \Gamma \in \mathbb{S}(\mathfrak{F}, \mathfrak{M}_0^{lin}), \tag{6}$$

where the γ_i's denote the elements of Γ and $p = \operatorname{card}(\Gamma)$, such that $\mathcal{S}$ satisfies $\mathcal{S} \subseteq \operatorname{span}(X)$ (or, *equivalently,* $\mathrm{span}(E_{n,\rho(\gamma_1)}(\gamma_1), \ldots, E_{n,\rho(\gamma_p)}(\gamma_p)) \subseteq \mathrm{span}(X)$*). Then* $\dim(\mathcal{S}) < n - \dim(\mathfrak{M}_0^{lin})$ *holds. Furthermore,*

$$\sup_{f \in \mathfrak{F}} P_{\mu_0,\sigma^2\Sigma(f)}(T \geq C) = 1$$

holds for every critical value C*,* $-\infty < C < \infty$*, for every* $\mu_0 \in \mathfrak{M}_0$*, and for every* $\sigma^2 \in (0, \infty)$*.*

This theorem is applicable to any nonempty set $\mathfrak{F}$ of normalized spectral densities. In case more is known about the richness of $\mathfrak{F}$, the sufficient condition in the preceding result can sometimes be simplified substantially. Below we present such a result making use of the subsequent lemma.

Lemma 3.5. *Let* $\mathfrak{F} \subseteq \mathfrak{F}_{\mathrm{all}}$ *satisfy* $\mathfrak{F} \supseteq \mathfrak{F}_{\mathrm{AR}(2)}$ *and let* $\mathcal{L}$ *be a linear subspace of* $\mathbb{R}^n$ *with* $\dim(\mathcal{L}) < n$*. Let* $\gamma \in [0, \pi]$*. Then* $\{\gamma\} \in \mathbb{S}(\mathfrak{F}, \mathcal{L})$ *if and only if* $\kappa(\underline{\omega}(\mathcal{L}), \underline{d}(\mathcal{L})) + \kappa(\gamma, 1) < n$*. And* $\{\gamma\} \in \mathbb{S}(\mathfrak{F}, \mathcal{L})$ *holds for every* $\gamma \in [0, \pi]$ *if and only if* $\kappa(\underline{\omega}(\mathcal{L}), \underline{d}(\mathcal{L})) + 2 < n$*.*

Remark 3.6. (i) A sufficient condition for $\kappa(\underline{\omega}(\mathcal{L}), \underline{d}(\mathcal{L})) + \kappa(\gamma, 1) < n$ ($\kappa(\underline{\omega}(\mathcal{L}), \underline{d}(\mathcal{L})) + 2 < n$, respectively) is given by $\dim(\mathcal{L}) + \kappa(\gamma, 1) < n$ ($\dim(\mathcal{L}) + 2 < n$, respectively). This follows from $\kappa(\underline{\omega}(\mathcal{L}), \underline{d}(\mathcal{L})) \leq \dim(\mathcal{L})$ established in Lemma D.1 in Pötscher and Preinerstorfer (2016).

(ii) In the case $\mathcal{L} = \mathfrak{M}_0^{lin}$ the latter two conditions become $k - q + \kappa(\gamma, 1) < n$ and $k - q + 2 < n$, respectively. Note that the condition $k - q + \kappa(\gamma, 1) < n$ is always satisfied for $\gamma = 0$ or $\gamma = \pi$ (as then $\kappa(\gamma, 1) = 1$). For $\gamma \in (0, \pi)$ this condition coincides with $k - q + 2 < n$, and is always satisfied except if $k = n - 1$ and $q = 1$.

Armed with the preceding lemma we can now establish the following consequence of Theorem 3.4 provided $\mathfrak{F}$ is rich enough to encompass $\mathfrak{F}_{\mathrm{AR}(2)}$, which clearly is a very weak condition.

Theorem 3.7. *Let* $\mathfrak{F} \subseteq \mathfrak{F}_{\mathrm{all}}$ *satisfy* $\mathfrak{F} \supseteq \mathfrak{F}_{\mathrm{AR}(2)}$*. Let* T *be a nonsphericity-corrected F-type test statistic of the form (4) based on* $\check{\beta}$ *and* $\check{\Omega}$ *satisfying Assumption 1 with* $N = \emptyset$*. Furthermore, assume that* $\check{\Omega}(y)$ *is nonnegative definite for every* $y \in \mathbb{R}^n$*. Suppose there exists a* $\gamma \in [0, \pi]$ *such that* $\mathrm{span}(E_{n,\rho(\gamma)}(\gamma)) \subseteq \mathrm{span}(X)$*. Then* $\kappa(\underline{\omega}(\mathfrak{M}_0^{lin}), \underline{d}(\mathfrak{M}_0^{lin})) + \kappa(\gamma, 1) < n$ *holds, and we have*

$$\sup_{f \in \mathfrak{F}} P_{\mu_0,\sigma^2\Sigma(f)}(T \geq C) = 1 \tag{7}$$

for every critical value C*,* $-\infty < C < \infty$*, for every* $\mu_0 \in \mathfrak{M}_0$*, and for every* $\sigma^2 \in (0, \infty)$*.*

Remark 3.8. *(On the necessity of the sufficient conditions for size control in Pötscher and Preinerstorfer (2016))* (i) Theorem 3.4 can be used to show that the conditions for size control obtained in Part 1 of Theorem 6.2 of Pötscher and Preinerstorfer (2016) are not only sufficient, but are actually also necessary in some important scenarios: Suppose T is as in Theorem 3.4, additionally satisfying $N^* = \mathrm{span}(X)$. Then Part 1 of Theorem 6.2 of Pötscher and Preinerstorfer (2016) (with $\mathcal{L} = \mathfrak{M}_0^{lin}$) is applicable and yields "$\mathcal{S} \nsubseteq \mathrm{span}(X)$ for all $\mathcal{S}$ satisfying (6) and $\dim(\mathcal{S}) < n - \dim(\mathfrak{M}_0^{lin})$" as a sufficient condition for the possibility of size control (cf. Theorem 6.4 in the same reference and note that $N^\dagger = N^* = \mathrm{span}(X)$). [Application of this theorem actually delivers the just given condition but with $\mathcal{S} \nsubseteq \operatorname{span}(X)$ replaced by $\mu_0 + \mathcal{S} \nsubseteq \operatorname{span}(X)$ where $\mu_0 \in \mathfrak{M}_0$. Because $\mu_0 \in \mathfrak{M}_0 \subseteq \operatorname{span}(X)$ and the latter set is a linear space, both formulations are seen to be equivalent.] Theorem 3.4 above now implies that the before mentioned sufficient condition is also necessary.

(ii) In a similar vein Theorem 3.7 can be used to show necessity of conditions for size control obtained from Part 2 of Theorem 6.2 of Pötscher and Preinerstorfer (2016) (with $\mathcal{L} = \mathfrak{M}_0^{lin}$) in important cases: Suppose T is as in (i). Then Part 2 of Theorem 6.2 of Pötscher and Preinerstorfer (2016) delivers "$\operatorname{span}(E_{n,\rho(\gamma)}(\gamma)) \nsubseteq \operatorname{span}(X)$ for all $\gamma \in \bigcup \mathbb{S}(\mathfrak{F}, \mathfrak{M}_0^{lin})$" as a sufficient condition for possibility of size control (cf. Theorem 6.4 in Pötscher and Preinerstorfer (2016)). Consider now the case where $\mathfrak{F}$ satisfies $\mathfrak{F} \supseteq \mathfrak{F}_{\mathrm{AR}(2)}$. We can then conclude from Theorem 3.7 that the just mentioned sufficient condition is also necessary. Similarly, the even stricter sufficient condition "$\operatorname{span}(E_{n,\rho(\gamma)}(\gamma)) \nsubseteq \operatorname{span}(X)$ for all $\gamma \in [0, \pi]$" also mentioned in Part 2 of Theorem 6.2 of Pötscher and Preinerstorfer (2016) is seen to be necessary.

(iii) Similarly as in Remark 3.2, the discussion in (i) and (ii) above covers (weighted) Eicker-test statistics $T_{E,\mathsf{W}}$ as well as classical autocorrelation robust test statistics T_w (the latter under Assumptions 1 and 2 of Pötscher and Preinerstorfer (2016) and if $N^* = \operatorname{span}(X)$ holds), and thus applies to the sufficient conditions in Theorems 3.2 and Theorem 6.5 in Pötscher and Preinerstorfer (2016) as well. Furthermore, the discussion in (i) and (ii) above also covers the test statistics T_{GQ} (provided the weighting matrix $\mathcal{W}_n^*$ is positive definite and $N^* = \operatorname{span}(X)$ holds).

The results so far have only concerned the size of nonsphericity-corrected F-type test statistics for which the exceptional set N is empty and $\check{\Omega}$ is nonnegative definite everywhere. We now provide a result also for the case where this condition is not met.

Definition 3.2. Let $\mathfrak{F}_{\mathrm{AR}(2)}^{ext}$ denote the set of all normalized spectral densities of the form $c_1 f + (2\pi)^{-1} c_2$ with $f \in \mathfrak{F}_{\mathrm{AR}(2)}$ and $c_1 + c_2 = 1$, $c_1 \geq 0$, $c_2 \geq 0$.

Obviously, $\mathfrak{F}_{\mathrm{AR}(2)} \subseteq \mathfrak{F}_{\mathrm{AR}(2)}^{ext} \subseteq \mathfrak{F}_{\mathrm{ARMA}(2,1)}$ holds. While the preceding result maintained that $\mathfrak{F}$ contains $\mathfrak{F}_{\mathrm{AR}(2)}$, the next result maintains the slightly stronger condition that $\mathfrak{F} \supseteq \mathfrak{F}_{\mathrm{AR}(2)}^{ext}$.

Theorem 3.9. *Let $\mathfrak{F} \subseteq \mathfrak{F}_{\mathrm{all}}$ satisfy $\mathfrak{F} \supseteq \mathfrak{F}_{\mathrm{AR}(2)}^{ext}$. Let T be a nonsphericity-corrected F-type test statistic of the form (4) based on $\check{\beta}$ and $\check{\Omega}$ satisfying Assumption 1. Furthermore, assume that $\check{\Omega}$ also satisfies Assumption 2. Suppose there exists a $\gamma \in [0, \pi]$ such that* $\operatorname{span}(E_{n,\rho(\gamma)}(\gamma)) \subseteq \operatorname{span}(X)$. *Then for every critical value C, $-\infty < C < \infty$, for every $\mu_0 \in \mathfrak{M}_0$, and for every $\sigma^2 \in (0, \infty)$ it holds that*

$$P_{0,I_n}(\check{\Omega} \text{ is nonnegative definite}) \leq K(\gamma) \leq \sup_{f \in \mathfrak{F}} P_{\mu_0, \sigma^2 \Sigma(f)} \left(T \geq C\right), \tag{8}$$

where $K(\gamma)$ is defined by

$$K(\gamma) = \int \Pr\left(\bar{\xi}_\gamma(x) \geq 0\right) dP_{0, I_{\kappa(\gamma,1)}}(x)$$

with the random variable $\bar{\xi}_\gamma(x)$ *given by*

$$\bar{\xi}_\gamma(x) = (R\hat{\beta}_X(\bar{E}_{n,\rho(\gamma)}(\gamma)x))'\check{\Omega}^{-1}(\mathbf{G})\, R\hat{\beta}_X(\bar{E}_{n,\rho(\gamma)}(\gamma)x)$$

on the event where $\{\mathbf{G} \in \mathbb{R}^n \backslash N^*\}$ *and by* $\bar{\xi}_\gamma(x) = 0$ *otherwise. Here* $\mathbf{G}$ *is a standard normal* n*-vector,* $\bar{E}_{n,\rho(\gamma)}(\gamma) = E_{n,\rho(\gamma)}(\gamma)$ *if* $\gamma \in (0, \pi)$ *and* $\bar{E}_{n,\rho(\gamma)}(\gamma)$ *denotes the first column of* $E_{n,\rho(\gamma)}(\gamma)$ *otherwise. [Recall that* $\hat{\beta}_X(y) = (X'X)^{-1}X'y$*.]*

The significance of the preceding theorem is that it provides a lower bound for the size of a large class of nonsphericity-corrected F-type tests, including those with $N \neq \emptyset$ or with $\check{\Omega}$ not necessarily nonnegative definite. In particular, it shows that size can not be controlled at a given desired significance level α, if α is below the threshold given by the lower bound in (8). Observe that this threshold will typically be close to 1, at least if n is sufficiently large, since (possibly after rescaling) $\check{\Omega}$ will often approach a positive definite matrix as $n \to \infty$.

Remark 3.10. (i) There are at most finitely many γ satisfying the assumption $\operatorname{span}(E_{n,\rho(\gamma)}(\gamma)) \subseteq \operatorname{span}(X)$ in the preceding theorem. To see this note that any such γ must coincide with a coordinate of $\underline{\omega}(\operatorname{span}(X))$ (since trivially $\operatorname{span}(E_{n,0}(\gamma)) \subseteq \operatorname{span}(X)$ in case $\rho(\gamma) = 0$ by this assumption, and since $\operatorname{span}(E_{n,0}(\gamma)) \subseteq \mathfrak{M}_0^{lin} \subseteq \operatorname{span}(X)$ in case $\rho(\gamma) > 0$), and that the dimension of the vector $\underline{\omega}(\operatorname{span}(X))$ is finite as discussed subsequent to Definition 3.1 in Pötscher and Preinerstorfer (2016).

(ii) If $\mathbb{G}$ denotes the (finite) set of γ's satisfying the assumption $\operatorname{span}(E_{n,\rho(\gamma)}(\gamma)) \subseteq \operatorname{span}(X)$ in the theorem, relation (8) in fact implies

$$P_{0,I_n}(\check{\Omega} \text{ is nonnegative definite}) \leq \min_{\gamma \in \mathbb{G}} K(\gamma) \leq \max_{\gamma \in \mathbb{G}} K(\gamma) \leq \sup_{f \in \mathfrak{F}} P_{\mu_0,\sigma^2\Sigma(f)}(T \geq C)\,.$$

(iii) Similar to Theorem 3.7, Theorem 3.9 also delivers (7) in case $\check{\Omega}$ is nonnegative definite $\lambda_{\mathbb{R}^n}$-almost everywhere. However, note that the latter theorem imposes a stronger condition on the set $\mathfrak{F}$.

Remark 3.11. Some results in this section are formulated for sets of spectral densities $\mathfrak{F}$ satisfying $\mathfrak{F} \supseteq \mathfrak{F}_{\mathrm{AR}(2)}$ or $\mathfrak{F} \supseteq \mathfrak{F}_{\mathrm{AR}(2)}^{ext}$, and thus for covariance models $\mathfrak{C}(\mathfrak{F})$ satisfying $\mathfrak{C}(\mathfrak{F}) \supseteq \mathfrak{C}(\mathfrak{F}_{\mathrm{AR}(2)})$ or $\mathfrak{C}(\mathfrak{F}) \supseteq \mathfrak{C}(\mathfrak{F}_{\mathrm{AR}(2)}^{ext})$, respectively. Trivially, these results also hold for any covariance model $\mathfrak{C}$ (not necessarily of the form $\mathfrak{C}(\mathfrak{F})$) that satisfies $\mathfrak{C} \supseteq \mathfrak{C}(\mathfrak{F}_{\mathrm{AR}(2)})$ or $\mathfrak{C} \supseteq \mathfrak{C}(\mathfrak{F}_{\mathrm{AR}(2)}^{ext})$, respectively. This observation also applies to other results in this paper further below and will not be repeated.

4 Negative results concerning power

We now show for a large class of test statistics, even larger than the class of nonsphericity-corrected F-type test statistics, that – under certain conditions – a choice of critical value leading to size less than one necessarily implies that the test is severely biased and thus has bad power properties in certain regions of the alternative hypothesis (cf. Part 3 of Theorem 5.7 and Remark 5.5(iii) in Preinerstorfer and Pötscher (2016)). The relevant conditions essentially say that a collection $\mathbb{K}$ as in the subsequent lemma can be found that is nonempty. It should be noted, however, that there are important instances where (i) the relevant conditions are not satisfied (that is, a nonempty $\mathbb{K}$ satisfying the properties required in the lemma does not exist) and (ii) small size and good power properties coexist. For results in that direction see Theorems 3.7, 5.10, 5.12, and 5.21 in Preinerstorfer and Pötscher (2016) as well as Proposition 5.2 and Theorem 5.4 in Preinerstorfer (2017).

The subsequent lemma is a variant of Lemma 5.10 in Pötscher and Preinerstorfer (2016). Recall that $\mathbb{H}$, defined in that lemma, certainly contains all one-dimensional $\mathcal{S} \in \mathbb{J}(\mathcal{L}, \mathfrak{C})$.

Lemma 4.1. *Let $\mathfrak{C}$ be a covariance model. Assume that the test statistic $T : \mathbb{R}^n \to \mathbb{R}$ is Borel-measurable and is continuous on the complement of a closed set $N^\dagger$. Assume that T and $N^\dagger$ are $G(\mathfrak{M}_0)$-invariant, and are also invariant w.r.t. addition of elements of a linear subspace $\mathcal{V}$ of $\mathbb{R}^n$. Define $\mathcal{L} = \operatorname{span}(\mathfrak{M}_0^{lin} \cup \mathcal{V})$ and assume that $\dim \mathcal{L} < n$. Let $\mathbb{H}$ and $C(\mathcal{S})$ be defined as in Lemma 5.10 of Pötscher and Preinerstorfer (2016). Let $\mathbb{K}$ be a subset of $\mathbb{H}$ and define $C_*(\mathbb{K}) = \inf_{\mathcal{S}\in\mathbb{K}} C(\mathcal{S})$ and $C^*(\mathbb{K}) = \sup_{\mathcal{S}\in\mathbb{K}} C(\mathcal{S})$, with the convention that $C_*(\mathbb{K}) = \infty$ and $C^*(\mathbb{K}) = -\infty$ if $\mathbb{K}$ is empty. Suppose that $\mathbb{K}$ has the property that for every $\mathcal{S} \in \mathbb{K}$ the set $N^\dagger$ is a $\lambda_{\mu_0+\mathcal{S}}$-null set for some $\mu_0 \in \mathfrak{M}_0$ (and hence for all $\mu_0 \in \mathfrak{M}_0$). Then the following holds:*

1. *For every $C \in (-\infty, C^*(\mathbb{K}))$, every $\mu_0 \in \mathfrak{M}_0$, and every $\sigma^2 \in (0,\infty)$ we have*

$$\sup_{\Sigma\in\mathfrak{C}} P_{\mu_0,\sigma^2\Sigma}(T \geq C) = 1.$$

2. *For every $C \in (C_*(\mathbb{K}), \infty)$, every $\mu_0 \in \mathfrak{M}_0$, and every $\sigma^2 \in (0,\infty)$ we have*

$$\inf_{\Sigma\in\mathfrak{C}} P_{\mu_0,\sigma^2\Sigma}(T \geq C) = 0.$$

Part 1 of the lemma implies that the size of the test equals 1 if $C < C^*(\mathbb{K})$. Part 2 shows that the test is severely biased for $C > C_*(\mathbb{K})$, which – in view of the invariance properties of T (cf. Theorem 5.7 and Remark 5.5(iii) in Preinerstorfer and Pötscher (2016)) – implies bad power properties such as (11) and (12) below. In particular, Part 2 implies that infimal power is zero for such choices of C. [Needless to say, the lemma neither implies that $\sup_{\Sigma\in\mathfrak{C}} P_{\mu_0,\sigma^2\Sigma}(T \geq C)$ is less than 1 for $C > C^*(\mathbb{K})$ nor that $\inf_{\Sigma\in\mathfrak{C}} P_{\mu_0,\sigma^2\Sigma}(T \geq C)$ is positive for $C < C_*(\mathbb{K})$. For conditions implying that size is less than 1 for appropriate choices of C see Pötscher and Preinerstorfer (2016).] Before proceeding we want to note that the preceding lemma also provides a negative size result, namely that the test based on T has size equal to 1 for *every* C, if $C^*(\mathbb{K}) = \infty$ holds for a collection $\mathbb{K}$ satisfying the assumptions of that lemma.

The announced theorem is now as follows and builds on the preceding lemma.

Theorem 4.2. *Let $\mathfrak{C}$ be a covariance model. Assume that the test statistic $T : \mathbb{R}^n \to \mathbb{R}$ is Borel-measurable and is continuous on the complement of a closed set $N^\dagger$. Assume that T and $N^\dagger$ are $G(\mathfrak{M}_0)$-invariant, and are also invariant w.r.t. addition of elements of a linear subspace $\mathcal{V}$ of $\mathbb{R}^n$. Define $\mathcal{L} = \operatorname{span}(\mathfrak{M}_0^{lin} \cup \mathcal{V})$ and assume that $\dim \mathcal{L} < n$. Then the following hold:*

1. *Suppose there exist two elements $\mathcal{S}_1$ and $\mathcal{S}_2$ of $\mathbb{H}$ such that $C(\mathcal{S}_1) \neq C(\mathcal{S}_2)$. Suppose further that for $i = 1, 2$ the set $N^\dagger$ is a $\lambda_{\mu_0+\mathcal{S}_i}$-null set for some $\mu_0 \in \mathfrak{M}_0$ (and hence for all $\mu_0 \in \mathfrak{M}_0$). Then for any critical value C, $-\infty < C < \infty$, satisfying*[5]
$$\sup_{\mu_0 \in \mathfrak{M}_0} \sup_{0<\sigma^2<\infty} \sup_{\Sigma \in \mathfrak{C}} P_{\mu_0,\sigma^2\Sigma}(T \geq C) < 1, \tag{9}$$
we have
$$\inf_{\mu_0 \in \mathfrak{M}_0} \inf_{0<\sigma^2<\infty} \inf_{\Sigma \in \mathfrak{C}} P_{\mu_0,\sigma^2\Sigma}(T \geq C) = 0. \tag{10}$$

2. *Suppose there exists an element $\mathcal{S}$ of $\mathbb{H}$ such that $N^\dagger$ is a $\lambda_{\mu_0+\mathcal{S}}$-null set for some $\mu_0 \in \mathfrak{M}_0$ (and hence for all $\mu_0 \in \mathfrak{M}_0$). Then (9) implies that $C \geq C(\mathcal{S})$ must hold; furthermore, (9) implies (10), except possibly if $C = C(\mathcal{S})$ holds.*

3. *Whenever (10) holds for some C, $-\infty < C < \infty$, it follows that*
$$\inf_{0<\sigma^2<\infty} \inf_{\Sigma \in \mathfrak{C}} P_{\mu_1,\sigma^2\Sigma}(T \geq C) = 0 \tag{11}$$
for every $\mu_1 \in \mathfrak{M}_1$. Also
$$\inf_{\mu_1 \in \mathfrak{M}_1} \inf_{\Sigma \in \mathfrak{C}} P_{\mu_1,\sigma^2\Sigma}(T \geq C) = 0 \tag{12}$$
for every $\sigma^2 \in (0, \infty)$.

In the important special case where $\mathcal{V} = \{0\}$, the assumptions on T and the associated set $N^\dagger$ in the second and third sentence of the preceding theorem are satisfied, e.g., for nonsphericity-corrected F-type test statistics (under Assumption 1), including the test statistics T_w, T_{GQ}, and $T_{E,\mathsf{W}}$; see Remark 2.1 above as well as Section 5.3.1 in Pötscher and Preinerstorfer (2016). Furthermore, for the class of test statistics T such that Theorem 3.1 applies (and for which $N^\dagger = N^* = \operatorname{span}(X)$ holds), it can be shown that $N^\dagger$ is a $\lambda_{\mu_0+\mathcal{S}}$-null set for any $\mathcal{S} \in \mathbb{H}$ (in fact, for any $\mathcal{S} \in \mathbb{J}(\mathcal{L}, \mathfrak{C})$) provided (9) holds. These observations lead to the following corollary.

Corollary 4.3. *Let $\mathfrak{C}$ be a covariance model and let T be a nonsphericity-corrected F-type test statistic of the form (4) based on $\check{\beta}$ and $\check{\Omega}$ satisfying Assumption 1 with $N = \emptyset$. Furthermore, assume that $\check{\Omega}(y)$ is nonnegative definite for every $y \in \mathbb{R}^n$ and that $N^* = \operatorname{span}(X)$.*

[5] Because of $G(\mathfrak{M}_0)$-invariance (cf. Remark 5.5(iii) in Preinerstorfer and Pötscher (2016)), the left-hand side of (9) coincides with $\sup_{\Sigma \in \mathfrak{C}} P_{\mu_0,\sigma^2\Sigma}(T \geq C)$ for any $\mu_0 \in \mathfrak{M}_0$ and any $\sigma^2 \in (0, \infty)$. Similarly, the left-hand side of (10) coincides with $\inf_{\Sigma \in \mathfrak{C}} P_{\mu_0,\sigma^2\Sigma}(T \geq C)$ for any $\mu_0 \in \mathfrak{M}_0$ and any $\sigma^2 \in (0, \infty)$.

1. *Suppose there exist two elements $\mathcal{S}_1$ and $\mathcal{S}_2$ of $\mathbb{H}$ (where $\mathbb{H}$ is as in Theorem 4.2 with $\mathcal{V} = \{0\}$) such that $C(\mathcal{S}_1) \neq C(\mathcal{S}_2)$. If a critical value C, $-\infty < C < \infty$, satisfies (9), then it also satisfies (10); and thus it also satisfies (11) and (12).*

2. *Suppose that $\mathbb{H}$ is nonempty (where $\mathbb{H}$ is as in Theorem 4.2 with $\mathcal{V} = \{0\}$) but $C(\mathcal{S})$ is the same for all $\mathcal{S} \in \mathbb{H}$. Then (9) implies that $C \geq C(\mathcal{S})$ must hold; furthermore, (9) implies (10) (and thus (11) and (12)), except possibly if $C = C(\mathcal{S})$ holds.*

Theorem 4.2 as well as the preceding corollary maintain conditions that, in particular, require $\mathbb{H}$ to be nonempty. In view of Lemma 5.10 in Pötscher and Preinerstorfer (2016), $\mathbb{H}$ is certainly nonempty if a one-dimensional $\mathcal{S} \in \mathbb{J}(\mathcal{L}, \mathfrak{C})$ exists. The following lemma shows that for $\mathfrak{C} = \mathfrak{C}(\mathfrak{F})$ with $\mathfrak{F} \supseteq \mathfrak{F}_{AR(2)}$ this is indeed the case; in fact, for such $\mathfrak{C}$ typically at least two such spaces exist.[6]

Lemma 4.4. *Let $\mathfrak{F} \subseteq \mathfrak{F}_{\mathrm{all}}$ satisfy $\mathfrak{F} \supseteq \mathfrak{F}_{\mathrm{AR}(2)}$. Define $\mathcal{L} = \mathrm{span}((\mathfrak{M}_0^{lin}) \cup \mathcal{V})$, where $\mathcal{V}$ is a linear subspace of $\mathbb{R}^n$, and assume that $\dim(\mathcal{L})+1 < n$. Then, for $\gamma \in \{0, \pi\}$, $\mathrm{span}\left(\Pi_{\mathcal{L}^\perp}\left(E_{n,\rho(\gamma,\mathcal{L})}(\gamma)\right)\right)$ belongs to $\mathbb{J}(\mathcal{L}, \mathfrak{C}(\mathfrak{F}))$ and is one-dimensional.*

The preceding lemma continues to hold for any covariance model $\mathfrak{C} \supseteq \mathfrak{C}(\mathfrak{F}_{\mathrm{AR}(2)})$ in a trivial way, since $\mathbb{J}(\mathcal{L}, \mathfrak{C}) \supseteq \mathbb{J}(\mathcal{L}, \mathfrak{C}(\mathfrak{F}_{\mathrm{AR}(2)}))$ then certainly holds.

5 Consequences for testing hypotheses on deterministic trends

In this section we discuss important consequences of the results obtained so far for testing restrictions on coefficients of polynomial and cyclical regressors. Such testing problems have, for obvious reasons, received a great deal of attention in econometrics, and are relevant in many other fields besides economics, e.g., climate research. In particular, we show that a large class of nonsphericity-corrected F-type test statistics leads to unsatisfactory test procedures in this context. In Subsection 5.1 we present results concerning hypotheses on the coefficients of polynomial regressors. Results concerning tests for hypotheses on the coefficients of cyclical regressors are briefly discussed in Subsection 5.2.

5.1 Polynomial regressors

We consider here the case where one tests hypotheses that involve the coefficient of a polynomial regressor as expressed in the subsequent assumption:

Assumption 3. Suppose that $X = (F, \tilde{X})$, where F is an $n \times k_F$-dimensional matrix ($1 \leq k_F \leq k$), the j-th column being given by $(1^{j-1}, \ldots, n^{j-1})'$, and where $\tilde{X}$ is an $n \times (k - k_F)$-dimensional

[6] While the one-dimensional spaces given in the lemma typically will be different, it is *not* established in the lemma that this is necessarily always the case.

matrix such that X has rank k (here $\tilde{X}$ is the empty matrix if $k_F = k$). Furthermore, suppose that the restriction matrix R has a nonzero column $R_{\cdot i}$ for some $i = 1, \ldots, k_F$, i.e., the hypothesis involves coefficients of the polynomial trend.

Under this assumption one obtains the subsequent theorem as a consequence of Theorem 3.7.

Theorem 5.1. *Let $\mathfrak{F} \subseteq \mathfrak{F}_{\mathrm{all}}$ satisfy $\mathfrak{F} \supseteq \mathfrak{F}_{\mathrm{AR}(2)}$. Suppose that Assumption 3 holds. Let T be a nonsphericity-corrected F-type test statistic of the form (4) based on $\check{\beta}$ and $\check{\Omega}$ satisfying Assumption 1 with $N = \emptyset$. Furthermore, assume that $\check{\Omega}(y)$ is nonnegative definite for every $y \in \mathbb{R}^n$. Then*

$$\sup_{f \in \mathfrak{F}} P_{\mu_0, \sigma^2 \Sigma(f)}(T \geq C) = 1$$

holds for every critical value C, $-\infty < C < \infty$, for every $\mu_0 \in \mathfrak{M}_0$, and for every $\sigma^2 \in (0, \infty)$.

The previous theorem relies in particular on the assumption that $N = \emptyset$ and that $\check{\Omega}$ is nonnegative definite everywhere. While these two assumptions may appear fairly natural and are widely satisfied, e.g., for the test statistics T_w, T_{GQ}, and $T_{E,\mathsf{W}}$ as discussed in Remark 2.1, we shall see in Subsections 5.1.1 and 5.1.2 below that they are not satisfied by some tests suggested in the literature. To obtain results also for tests that are not covered by the previous theorem we can apply Theorem 3.9. The following result is then obtained.

Theorem 5.2. *Let $\mathfrak{F} \subseteq \mathfrak{F}_{\mathrm{all}}$ satisfy $\mathfrak{F} \supseteq \mathfrak{F}^{ext}_{\mathrm{AR}(2)}$. Suppose that Assumption 3 holds. Let T be a nonsphericity-corrected F-type test statistic of the form (4) based on $\check{\beta}$ and $\check{\Omega}$ satisfying Assumption 1. Furthermore, assume that $\check{\Omega}$ also satisfies Assumption 2. Then for every critical value C, $-\infty < C < \infty$, for every $\mu_0 \in \mathfrak{M}_0$, and for every $\sigma^2 \in (0, \infty)$ it holds that*

$$P_{0,I_n}(\check{\Omega} \text{ is nonnegative definite}) \leq P_{0,I_n}(R_{\cdot i_0}' \check{\Omega}^{-1} R_{\cdot i_0} \geq 0) \leq \sup_{f \in \mathfrak{F}} P_{\mu_0, \sigma^2 \Sigma(f)}\left(T \geq C\right), \qquad (13)$$

where $R_{\cdot i_0}$ denotes the first nonzero column of R. [Note that $\check{\Omega}$ is P_{0,I_n}-almost everywhere nonsingular in view of Assumption 1.]

Theorem 5.2 shows that under Assumption 3 a large class of nonsphericity-corrected F-type tests, including cases with $N \neq \emptyset$ or with $N = \emptyset$ but where $\check{\Omega}$ is not necessarily nonnegative definite everywhere, typically have large size. In particular, size can not be controlled at a given desired significance level α, if α is below the lower bound in (13). Observe that this lower bound will typically be close to 1, at least if n is sufficiently large.

To illustrate the scope and applicability of Theorems 5.1 and 5.2 above (beyond the test statistics such as T_w, T_{GQ}, and $T_{E,\mathsf{W}}$ mentioned before), we shall now apply them to some commonly used test statistics that have been designed for testing polynomial trends. First, in Subsection 5.1.1, we shall derive properties of conventional tests for polynomial trends. Such tests are based on long-run-variance estimators and classical results due to Grenander (1954). In Subsection 5.1.2 we shall discuss properties of tests that have been introduced more recently by

Vogelsang (1998) and Bunzel and Vogelsang (2005). While our discussion of methods is certainly not exhaustive (for example, we do not discuss tests in Harvey et al. (2007) or Perron and Yabu (2009), which have been suggested only for the special case of testing a restriction on the slope in a "linear trend plus noise model"), it should also serve the purpose of presenting a general pattern how one can check the reliability of polynomial trend tests. It might also help to avoid pitfalls in the construction of novel tests for polynomial trends.

5.1.1 Properties of conventional tests for hypotheses on polynomial trends

The structure of tests that have traditionally been used for testing restrictions on coefficients of polynomial trends (i.e., when the design matrix X satisfies Assumption 3, and in particular if $k_F = k$) is motivated by results concerning the asymptotic covariance matrix of the OLS estimator (and its efficiency) in regression models with stationary error processes and deterministic polynomial time trends by Grenander (1954) (cf. also the discussion in Bunzel and Vogelsang (2005) on p. 383). The corresponding test statistics are nonsphericity-corrected F-type test statistics as in (4). They are based on the OLS estimator $\hat{\beta}$ and a covariance matrix estimator

$$\check{\Omega}_{\mathcal{W}}(y) = \hat{\omega}_{\mathcal{W}}(y) R(X'X)^{-1}R'. \tag{14}$$

Here the "long-run-variance estimator" $\hat{\omega}_{\mathcal{W}}$ is of the form

$$\hat{\omega}_{\mathcal{W}}(y) = n^{-1}\hat{u}'(y)\mathcal{W}(y)\hat{u}(y), \tag{15}$$

where $\mathcal{W}(y)$ is a symmetric, possibly data-dependent, $n \times n$-dimensional matrix that might not be well-defined on all of $\mathbb{R}^n$.[7] In many cases, however, $\mathcal{W}$ is constant, i.e., does not depend on y, and is also positive definite. For example, this is so in the leading case where the (i, j)-th element of $\mathcal{W}$ is of the form $\kappa(|i-j|/M)$ for some (deterministic) $M > 0$ (typically depending on n) and a kernel function κ such as the Bartlett, Parzen, Quadratic-Spectral, or Daniell kernel (positive definiteness does not hold, e.g., for the rectangular kernel). Note that in case $\mathcal{W}$ is given by a kernel κ the estimator $\hat{\omega}_{\mathcal{W}}$ in the previous display can be written in the more familiar form

$$\hat{\omega}_{\mathcal{W}}(y) = \sum_{i=-(n-1)}^{n-1} \kappa(|i|/M)\hat{\gamma}_i(y),$$

where $\hat{\gamma}_i(y) = \hat{\gamma}_{-i}(y) = n^{-1}\sum_{j=i+1}^{n} \hat{u}_j(y)\hat{u}_{j-i}(y)$ for $i \geq 0$. For trend tests based on the OLS estimator $\hat{\beta}$ and a covariance estimator $\check{\Omega}_{\mathcal{W}}$ as in (14) we shall first obtain two corollaries from Theorems 5.1 and 5.2 that cover the case where $\mathcal{W}$ is constant.[8] Further below we shall then

[7] The matrix $\mathcal{W}$ may depend on n, a dependence not shown in the notation. Furthermore, assuming symmetry of $\mathcal{W}$ entails no loss of generality, since given a long-run-variance-estimator as in (15) and based on a non-symmetric weights matrix $\mathcal{W}_*$, one can always pass to an equivalent long-run-variance estimator by replacing $\mathcal{W}_*$ with the symmetric matrix $\mathcal{W} = (\mathcal{W}_* + \mathcal{W}_*')/2$.

[8] The slightly more general case, where $\mathcal{W}$ is not constant in y (and is defined on all of $\mathbb{R}^n$) but $\mathcal{W}^* := \Pi_{\mathrm{span}(X)^\perp}\mathcal{W}\Pi_{\mathrm{span}(X)^\perp}$ is so, can immediately be subsumed under the present discussion, if one observes that

address the case where $\mathcal{W}$ is allowed to depend on y. Note that the assumptions on $\mathcal{W}$ in the subsequent corollary are certainly met if $\mathcal{W}$ is constant, symmetric, and positive definite, and hence are satisfied in the leading case mentioned before (provided M is deterministic).

Corollary 5.3. *Let $\mathfrak{F} \subseteq \mathfrak{F}_{\mathrm{all}}$ satisfy $\mathfrak{F} \supseteq \mathfrak{F}_{\mathrm{AR}(2)}$ and suppose that Assumption 3 holds. Suppose further that $\mathcal{W}$ is constant and symmetric, and that $\Pi_{\mathrm{span}(X)^{\perp}} \mathcal{W} \Pi_{\mathrm{span}(X)^{\perp}}$ is nonzero and nonnegative definite. Then $\check{\beta} = \hat{\beta}$ and $\check{\Omega} = \hat{\Omega}_{\mathcal{W}}$ satisfy Assumption 1 with $N = \emptyset$. Let T be of the form (4) with $\check{\beta} = \hat{\beta}$, $\check{\Omega} = \hat{\Omega}_{\mathcal{W}}$, and $N = \emptyset$. Then*

$$\sup_{f \in \mathfrak{F}} P_{\mu_0, \sigma^2 \Sigma(f)}(T \geq C) = 1$$

holds for every critical value C, $-\infty < C < \infty$, for every $\mu_0 \in \mathfrak{M}_0$, and for every $\sigma^2 \in (0, \infty)$.

We next consider the case where the matrix $\Pi_{\mathrm{span}(X)^{\perp}} \mathcal{W} \Pi_{\mathrm{span}(X)^{\perp}}$ is nonzero, but not (necessarily) nonnegative definite, and thus the previous corollary is not applicable. The subsequent corollary covers this case and is obtained under the slightly stronger assumption that $\mathfrak{F} \supseteq \mathfrak{F}_{AR(2)}^{ext}$. [Note also that the case where $\mathcal{W}$ is constant but $\Pi_{\mathrm{span}(X)^{\perp}} \mathcal{W} \Pi_{\mathrm{span}(X)^{\perp}}$ is equal to zero is of no interest as it leads to a long-run-variance estimator that vanishes identically.]

Corollary 5.4. *Let $\mathfrak{F} \subseteq \mathfrak{F}_{\mathrm{all}}$ satisfy $\mathfrak{F} \supseteq \mathfrak{F}_{\mathrm{AR}(2)}^{ext}$ and suppose that Assumption 3 holds. Suppose further that $\mathcal{W}$ is constant and symmetric, and that $\Pi_{\mathrm{span}(X)^{\perp}} \mathcal{W} \Pi_{\mathrm{span}(X)^{\perp}}$ is nonzero. Then $\check{\beta} = \hat{\beta}$ and $\check{\Omega} = \hat{\Omega}_{\mathcal{W}}$ satisfy Assumption 1 with $N = \emptyset$. Let T be of the form (4) with $\check{\beta} = \hat{\beta}$, $\check{\Omega} = \hat{\Omega}_{\mathcal{W}}$, and $N = \emptyset$. Then*

$$P_{0, I_n}(\hat{\omega}_{\mathcal{W}} \geq 0) \leq \sup_{f \in \mathfrak{F}} P_{\mu_0, \sigma^2 \Sigma(f)}(T \geq C) \tag{16}$$

holds for every critical value C, $-\infty < C < \infty$, for every $\mu_0 \in \mathfrak{M}_0$, and for every $\sigma^2 \in (0, \infty)$. Furthermore, for every $0 \leq C < \infty$ the lower bound in the previous display is an upper bound for the maximal power of the test under i.i.d. errors, i.e.,

$$\sup_{\mu_1 \in \mathfrak{M}_1} \sup_{0 < \sigma^2 < \infty} P_{\mu_1, \sigma^2 I_n}(T \geq C) \leq P_{0, I_n}(\hat{\omega}_{\mathcal{W}} \geq 0). \tag{17}$$

The previous corollary shows that the size of the test is bounded from below by the probability that the long-run-variance estimator $\hat{\omega}_{\mathcal{W}}$ used in the construction of the test statistic is nonnegative, where the probability is taken under $N(0, I_n)$-distributed errors. For consistent long-run-variance estimators this probability approaches 1 as sample size increases, and hence the size of tests based on such estimators $\hat{\omega}_{\mathcal{W}}$ will exceed any prescribed nominal significance level $\alpha \in (0, 1)$ eventually. Additionally, it is shown in that corollary that for nonnegative critical values (the standard in applications) the probability $P_{0, I_n}(\hat{\omega}_{\mathcal{W}} \geq 0)$ also provides an upper bound on the *maximal power* of the test under i.i.d. errors. Thus, if the lower bound in (16)

$\hat{\omega}_{\mathcal{W}}$ coincides with $\hat{\omega}_{\mathcal{W}^*}$ and $\mathcal{W}^*$ is constant.

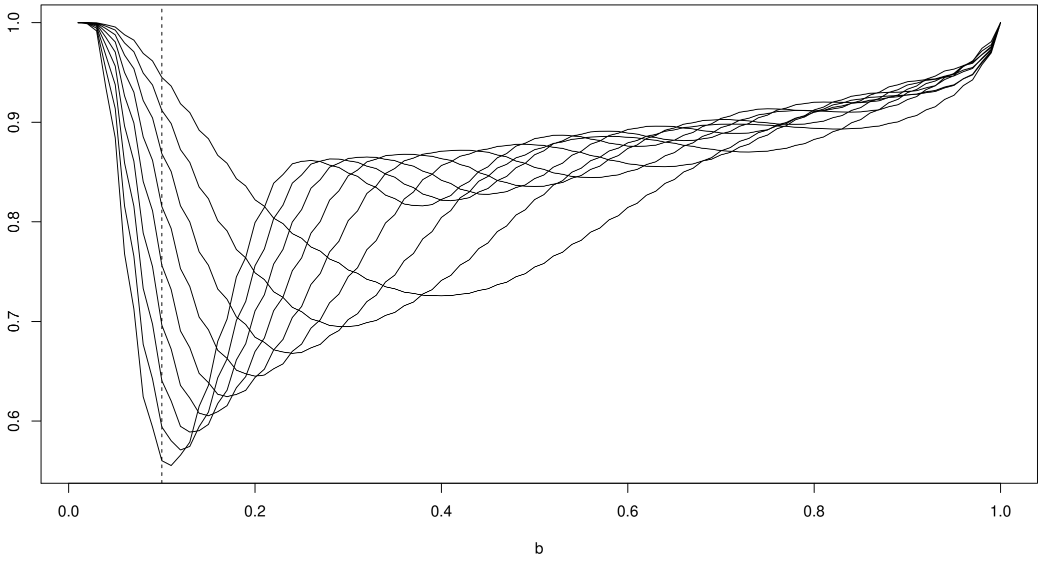

Figure 1: Numerical values of $P_{0,I_n}(\hat{\omega}_{\mathcal{W}} \geq 0)$ for $\mathcal{W}_{ij} = \mathbf{1}_{(-1,1)}((i-j)/(bn))$ as a function of $b \in (0,1)$. Sample size $n = 150$ and Assumption 3 holds with $k_F = k$ and for different values of $k \in \{2, 3, 4, \ldots, 10\}$. The probabilities for $k = 2$ correspond to the function with the largest value at the dashed vertical line, the probablities for $k = 3$ correspond to the function with the second largest value at the dashed vertical line, etc.

is small, and hence (16) does not tell us much about size, the inequality in (17) shows that power must then be small over a substantial subset of the parameter space (unless perhaps one chooses a negative critical value). To get an idea of the magnitude of the lower (upper) bound in (16) ((17)) in a special case, we computed $P_{0,I_n}(\hat{\omega}_{\mathcal{W}} \geq 0)$ numerically for the rectangular kernel, i.e., for $\mathcal{W}_{ij} = \mathbf{1}_{(-1,1)}((i-j)/M)$, for the cases when Assumption 3 is satisfied with $k_F = k \in \{2, 3, 4, \ldots, 10\}$, respectively, sample size $n = 150$, and bandwidth parameter $M = bn$ for $b \in \{0.01, 0.02, \ldots, 1\}$. The results are presented in Figure 1. For all values of b and k the probability $P_{0,I_n}(\hat{\omega}_{\mathcal{W}} \geq 0)$ is quite large, in particular is larger than 1/2, and thus exceeds commonly used significance levels. Thus, as a consequence of (16), one has strong size distortions regardless of the value of b chosen if one decides to use a test based on the rectangular kernel. Note also that the probability $P_{0,I_n}(\hat{\omega}_{\mathcal{W}} \geq 0)$ can be easily obtained numerically in any other case, as it is the probability that a quadratic form in a standard Gaussian random vector is nonnegative (for the actual computation we used the algorithm by Davies (1980)).

The assumption of $\mathcal{W}$ being data-independent, i.e., constant as a function of $y \in \mathbb{R}^n$, in the previous two corollaries is not satisfied for the important class of long-run-variance estimators that incorporate prewhitening or data-dependent bandwidth parameters (e.g., Andrews (1991), Andrews and Monahan (1992) and Newey and West (1994)). An additional complication for such estimators is that the corresponding weights matrix $\mathcal{W}(y)$, and thus also $\check{\Omega}_{\mathcal{W}}$, are in general not well-defined for every $y \in \mathbb{R}^n$. Nevertheless, after a careful structural analysis of such estimators (similar to the results obtained in Section 3.3 of Preinerstorfer (2017)), one can typically show that the resulting test statistic satisfies the assumptions of Theorem 5.2 above and thus one can obtain suitable versions of the above corollaries tailored towards test statistics based on specific classes of prewhitened long-run-variance estimators with data-dependent bandwidth parameters. To make this more compelling, we provide in the following such a result for a widely used procedure in that class. We consider a version of the AR(1)-prewhitened long-run-variance estimator based on auxiliary AR(1) models for bandwidth selection and the Quadratic-Spectral kernel as discussed in Andrews and Monahan (1992). This is a long-run-variance estimator as in (15), where the weights matrix is obtained as follows (the set where all involved quantities are well-defined is given in (19) further below): Let

$$\hat{\rho}(y) = \frac{\sum_{i=2}^{n} \hat{u}_i(y)\hat{u}_{i-1}(y)}{\sum_{i=1}^{n-1} \hat{u}_i^2(y)}, \tag{18}$$

and define $\hat{v}_i(y) = \hat{u}_{i+1}(y) - \hat{\rho}(y)\hat{u}_i(y)$ for $i = 1, \ldots, n-1$, which one can write in an obvious way as $\hat{v}(y) = \mathsf{A}(\hat{\rho}(y))\hat{u}(y)$ with $\rho \mapsto \mathsf{A}(\rho) \in \mathbb{R}^{(n-1)\times n}$ a continuous function on $\mathbb{R}$. Define the data-dependent bandwidth parameter M_{AM} via

$$M_{\mathrm{AM}}(y) = 1.3221 \left(n \frac{4\tilde{\rho}^2(y)}{(1-\tilde{\rho}(y))^4} \right)^{1/5} \qquad \text{with} \qquad \tilde{\rho}(y) = \frac{\sum_{i=2}^{n-1} \hat{v}_i(y)\hat{v}_{i-1}(y)}{\sum_{i=1}^{n-2} \hat{v}_i^2(y)}.$$

The long-run-variance estimator $\hat{\omega}_{\mathcal{W}_{\mathrm{AM}}}$ is now obtained (granted the involved expressions are well-defined) by choosing $\mathcal{W}$ in (15) equal to

$$\mathcal{W}_{\mathrm{AM}}(y) = (1-\hat{\rho}(y))^{-2} \mathsf{A}'(\hat{\rho}(y)) \left[\kappa_{\mathrm{QS}}(|i-j|/M_{\mathrm{AM}}(y))\right]_{i,j=1}^{n-1} \mathsf{A}(\hat{\rho}(y)),$$

where $[\kappa_{\mathrm{QS}}(|i-j|/M_{\mathrm{AM}}(y))]_{i,j=1}^{n-1}$ is defined as I_{n-1} in case $M_{\mathrm{AM}}(y) = 0$ holds (cf., e.g., p. 821 in Andrews (1991) for a definition of the Quadratic-Spectral kernel κ_{QS}). The corresponding covariance matrix estimator $\check{\Omega}_{\mathcal{W}_{\mathrm{AM}}}$ is then given by plugging $\hat{\omega}_{\mathcal{W}_{\mathrm{AM}}}$ into (14). The set where $\mathcal{W}_{\mathrm{AM}}$ (and hence $\check{\Omega}_{\mathcal{W}_{\mathrm{AM}}}$) is well-defined is easily seen to coincide with the set of all $y \in \mathbb{R}^n$ such that $\hat{\rho}(y)$ and $\tilde{\rho}(y)$ are both well-defined and are not equal to 1, i.e., with the set

$$\left\{ y \in \mathbb{R}^n : \sum_{i=1}^{n-1} \hat{u}_i(y)(\hat{u}_{i+1}(y) - \hat{u}_i(y)) \neq 0, \ \sum_{i=1}^{n-2} \hat{v}_i(y)(\hat{v}_{i+1}(y) - \hat{v}_i(y)) \neq 0 \right\}. \tag{19}$$

Define N_{AM} as the complement of the set (19) in $\mathbb{R}^n$. A result concerning size properties of polynomial trend tests based on the long-run-variance estimator $\hat{\omega}_{\mathcal{W}_{\mathrm{AM}}}$ is now obtained by combining Theorem 5.2 above with results obtained in Lemma D.3 in Appendix D, showing, in particular, that $\hat{\beta}$ and $\check{\Omega}_{\mathcal{W}_{\mathrm{AM}}}$ satisfy Assumptions 1 with $N = N_{\mathrm{AM}}$, provided $N_{\mathrm{AM}} \neq \mathbb{R}^n$ holds. Note that

(i) the condition $N_{\mathrm{AM}} \neq \mathbb{R}^n$ only depends on properties of the design matrix X and hence can be checked, and that (ii) in case $N_{\mathrm{AM}} = \mathbb{R}^n$, the matrix $\check{\Omega}_{\mathcal{W}_{\mathrm{AM}}}$ is *nowhere* well-defined, and tests based on this estimator hence break down in a trivial way.

Corollary 5.5. *Let $\mathfrak{F} \subseteq \mathfrak{F}_{\mathrm{all}}$ satisfy $\mathfrak{F} \supseteq \mathfrak{F}_{\mathrm{AR}(2)}^{ext}$ and suppose Assumption 3 holds. Suppose further that $N_{\mathrm{AM}} \neq \mathbb{R}^n$. Then $\check{\beta} = \hat{\beta}$ and $\check{\Omega} = \check{\Omega}_{\mathcal{W}_{\mathrm{AM}}}$ satisfy Assumption 1 with $N = N_{\mathrm{AM}}$. Let T be of the form (4) with $\check{\beta} = \hat{\beta}$, $\check{\Omega} = \check{\Omega}_{\mathcal{W}_{\mathrm{AM}}}$, and $N = N_{\mathrm{AM}}$. Then*

$$\sup_{f \in \mathfrak{F}} P_{\mu_0, \sigma^2 \Sigma(f)}(T \geq C) = 1$$

holds for every critical value C, $-\infty < C < \infty$, for every $\mu_0 \in \mathfrak{M}_0$, and for every $\sigma^2 \in (0, \infty)$.

5.1.2 Properties of some recently suggested tests for hypotheses on polynomial trends

In this subsection we discuss finite sample properties of classes of tests for polynomial trends that have been suggested in Vogelsang (1998) and Bunzel and Vogelsang (2005). We start with a discussion of the tests introduced in the former article. Vogelsang (1998) introduces two classes of tests for testing hypotheses on trends, in particular polynomial trends. From Section 3.2 of Vogelsang (1998) it is not difficult to see that these classes of test statistics (i.e., the classes referred to as PS_T^i and PSW_T^i in that reference) are (possibly up to a constant positive multiplicative factor that can be absorbed into the critical value) of the form (4). More specifically, the test statistics in Vogelsang (1998) are based on a combination of one of the two estimators

$$\check{\beta}_V(y) = \hat{\beta}_{VX}(Vy) = (X'V'VX)^{-1} X'V'Vy \quad \text{for } V \in \{A, I_n\}, \tag{20}$$

with a corresponding covariance estimator of the form

$$\check{\Omega}^{\mathrm{Vo}}_{c,U,i,V}(y) = n^{j(V)} s^2_{A,X}(y) \exp(c J^i_{n,U}(y)) R (X'V'VX)^{-1} R', \tag{21}$$

for $i \in \{1, 2\}$ and where $j(V) = 1$ if $V = A$ and $j(V) = -1$ if $V = I_n$. Here A is the $n \times n$-dimensional matrix that has 0 above the main diagonal and 1 on and below the main diagonal, c is a real number[9], U is an $n \times m$-dimensional matrix (with $m \geq 1$) such that (X, U) is of full column-rank $k + m < n$. [In Vogelsang (1998) the column vectors of U correspond to polynomial trends of an order exceeding the polynomial trends already contained in span(X).] Furthermore,

$$J^1_{n,U}(y) = n^{-1} \hat{\beta}'_{(X,U)}(y) G' \left(s^2_{I_n,(X,U)}(y) G ((X,U)'(X,U))^{-1} G' \right)^{-1} G \hat{\beta}_{(X,U)}(y), \tag{22}$$

[9] We here also allow for the value $c = 0$ in the formulation of the covariance estimators because this turns out to be convenient in the proofs.

and

$$J^2_{n,U}(y) = n^{-1}\hat{\beta}'_{A(X,U)}(Ay)G'\left(s^2_{A,(X,U)}(y)G((X,U)'A'A(X,U))^{-1}G'\right)^{-1}G\hat{\beta}_{A(X,U)}(Ay),$$

with $G = (0, I_m) \in \mathbb{R}^{m\times(k+m)}$, where we use the notation

$$s^2_{D_1,D_2}(y) = n^{-1}y'D_1'\Pi_{\operatorname{span}(D_1D_2)^\perp}D_1y$$

for nonsingular $D_1 \in \mathbb{R}^{n\times n}$ and for $D_2 \in \mathbb{R}^{n\times l}$ of rank $l \leq n$. It is obvious from the above expressions that the covariance estimator $\check{\Omega}^{\mathrm{Vo}}_{c,U,i,V}$ is not well-defined on all of $\mathbb{R}^n$. However, it is also not difficult to see that the set where such an estimator is well-defined coincides with $\mathbb{R}^n\setminus\operatorname{span}(X,U)$, see the proof of Lemma D.4 in Appendix D. We stress once more that the matrix U used in the construction above is chosen in a particular way in Vogelsang (1998). We do not impose such a restriction here, because it would unnecessarily complicate the presentation of the result below, and because this restrictions is actually not necessary for establishing the result. The following result now shows that the tests suggested in Vogelsang (1998) suffer from substantial size distortions in case $\mathfrak{F} \supseteq \mathfrak{F}^{ext}_{AR(2)}$.

Corollary 5.6. *Let $\mathfrak{F} \subseteq \mathfrak{F}_{\mathrm{all}}$ satisfy $\mathfrak{F} \supseteq \mathfrak{F}^{ext}_{\mathrm{AR}(2)}$ and suppose Assumption 3 holds. Let $V \in \{A, I_n\}$, $c \in \mathbb{R}$, $i \in \{1,2\}$, and let U be an $n\times m$-dimensional matrix with $m \geq 1$, $k+m < n$, such that (X,U) is of full column-rank. Then $\check{\beta} = \check{\beta}_V$ and $\check{\Omega} = \check{\Omega}^{\mathrm{Vo}}_{c,U,i,V}$ satisfy Assumption 1 with $N = \operatorname{span}(X,U)$. Let T be of the form (4) with $\check{\beta} = \check{\beta}_V$, $\check{\Omega} = \check{\Omega}^{\mathrm{Vo}}_{c,U,i,V}$, and $N = \operatorname{span}(X,U)$. Then*

$$\sup_{f\in\mathfrak{F}} P_{\mu_0,\sigma^2\Sigma(f)}(T \geq C) = 1$$

holds for every critical value C, $-\infty < C < \infty$, for every $\mu_0 \in \mathfrak{M}_0$, and for every $\sigma^2 \in (0,\infty)$.

Next we turn to the tests introduced in Bunzel and Vogelsang (2005). We first discuss tests introduced in that article with data-independent tuning parameters and data-independent critical values: These tests are based on the OLS estimator $\hat{\beta}$ and two classes of covariance matrix estimators, both of which incorporate a tuning parameter $c \in \mathbb{R}$, and which are defined as

$$\check{\Omega}^{\mathrm{BV},J}_{\mathcal{W},U,c}(y) = \hat{\omega}_{\mathcal{W}}(y)\exp(cJ^1_{n,U}(y))R(X'X)^{-1}R', \tag{23}$$

where U is an $n\times m$-dimensional matrix with $m \geq 1$ such that (X,U) is of full column-rank $k+m < n$ (note that $\hat{\omega}_{\mathcal{W}}$ and $J^1_{n,U}$ have been defined in (15) and (22) above), and

$$\check{\Omega}^{\mathrm{BV}}_{\mathcal{W},c}(y) = \hat{\omega}_{\mathcal{W}}(y)\exp\left(cn^{-2}\frac{\hat{u}'(y)A'A\hat{u}(y)}{\hat{u}'(y)\hat{u}(y)}\right)R(X'X)^{-1}R' \tag{24}$$

where A has been defined below (21). The subsequent result applies, in particular, if $\mathcal{W}_{ij} = \kappa(|i-j|/M)$ where $M > 0$ is a (fixed) real number and κ is a kernel function such that $\mathcal{W}$ is positive definite, including the recommendation in Bunzel and Vogelsang (2005) to use the

Daniell kernel. In that case, and more generally whenever $\Pi_{\mathrm{span}(X)^\perp}\mathcal{W}\Pi_{\mathrm{span}(X)^\perp}$ is nonzero and nonnegative definite (with $\mathcal{W}$ constant[10] and symmetric), the subsequent corollary shows that the above mentioned tests in Bunzel and Vogelsang (2005) have size equal to one if $\mathfrak{F} \supseteq \mathfrak{F}_{AR(2)}^{ext}$; in case $\Pi_{\mathrm{span}(X)^\perp}\mathcal{W}\Pi_{\mathrm{span}(X)^\perp}$ is nonzero but not nonnegative definite, a lower bound on the size is obtained, which also provides an upper bound for the power in the case of i.i.d. errors. A discussion similar to the discussion following Corollary 5.4 also applies here (cf. also Figure 1).

Corollary 5.7. *Let $\mathfrak{F} \subseteq \mathfrak{F}_{\mathrm{all}}$ satisfy $\mathfrak{F} \supseteq \mathfrak{F}_{\mathrm{AR}(2)}^{ext}$ and suppose Assumption 3 holds. Suppose that $\mathcal{W}$ is constant and symmetric, that $\Pi_{\mathrm{span}(X)^\perp}\mathcal{W}\Pi_{\mathrm{span}(X)^\perp}$ is nonzero, and that $c \in \mathbb{R}$. Furthermore, for the statements that involve U, suppose U is an $n \times m$-dimensional matrix with $m \geq 1$ such that (X, U) is of full column-rank $k + m < n$. Then, $\check{\beta} = \hat{\beta}$ and $\check{\Omega} = \check{\Omega}_{\mathcal{W},c}^{\mathrm{BV}}$ ($\check{\beta} = \hat{\beta}$ and $\check{\Omega} = \check{\Omega}_{\mathcal{W},U,c}^{\mathrm{BV},J}$, respectively) satisfy Assumption 1 with $N = \mathrm{span}(X)$ ($N = \mathrm{span}(X, U)$, respectively). Let T be of the form (4) with $\check{\beta} = \hat{\beta}$, $\check{\Omega} = \check{\Omega}_{\mathcal{W},c}^{\mathrm{BV}}$, and $N = \mathrm{span}(X)$, or with $\check{\beta} = \hat{\beta}$, $\check{\Omega} = \check{\Omega}_{\mathcal{W},U,c}^{\mathrm{BV},J}$, and $N = \mathrm{span}(X, U)$. Then*

$$P_{0,I_n}(\hat{\omega}_{\mathcal{W}} \geq 0) \leq \sup_{f \in \mathfrak{F}} P_{\mu_0,\sigma^2\Sigma(f)}(T \geq C)$$

holds for every critical value C, $-\infty < C < \infty$, for every $\mu_0 \in \mathfrak{M}_0$, and for every $\sigma^2 \in (0, \infty)$. The lower bound equals 1 in case $\Pi_{\mathrm{span}(X)^\perp}\mathcal{W}\Pi_{\mathrm{span}(X)^\perp}$ is nonnegative definite. Furthermore, for every $0 \leq C < \infty$ the lower bound in the previous display is an upper bound for the maximal power of the test under i.i.d. errors, i.e.,

$$\sup_{\mu_1 \in \mathfrak{M}_1} \sup_{0 < \sigma^2 < \infty} P_{\mu_1,\sigma^2 I_n}(T \geq C) \leq P_{0,I_n}(\hat{\omega}_{\mathcal{W}} \geq 0). \tag{25}$$

We shall now turn to the approach Bunzel and Vogelsang (2005) suggest for practical applications. This approach is based on a data-driven selection of the weights matrix $\mathcal{W}$ and of the tuning parameter c, and on a data-driven selection of the critical value C. Their approach is as follows: Bunzel and Vogelsang (2005) focus on $\hat{\omega}_{\mathcal{W}}$ based on the Daniell kernel. More specifically, they set $\mathcal{W}_{ij} = \kappa_D(|i - j| / \max(bn, 2))$ (cf. Bunzel and Vogelsang (2005), Appendix B, for a definition of the Daniell kernel). Recall that, regardless of the value of b, the matrix with elements $\mathcal{W}_{ij} = \kappa_D(|i - j| / \max(bn, 2))$ based on the Daniell kernel is positive definite. The authors recommend to choose b as a positive piecewise constant function of $\hat{\rho}$ (which has been defined in (18) above), i.e., for constants $a_i \in (0, \infty)$, $i = 0, \ldots, m'$ ($m' \in \mathbb{N}$), and $\bar{a}_i \in \mathbb{R}$, $i = 1, \ldots, m'$, they suggest to use

$$b_{\mathrm{BV}}(y, a, \bar{a}) = a_0 + \sum_{i=1}^{m'} a_i \mathbf{1}_{[\bar{a}_i, \infty)}(\hat{\rho}(y)).$$

For a recommendation concerning the choice of these constants see Bunzel and Vogelsang (2005), p. 388. Furthermore, Bunzel and Vogelsang (2005) suggest to choose their data-driven critical

[10] Cf. Footnote 8

value C and a data-driven tuning parameter c as a polynomial function of $b_{\mathrm{BV}}(y, a, \bar{a})$, respectively. More precisely, for constants $h_0, \ldots, h_{m''} \in \mathbb{R}$ ($m'' \in \mathbb{N}$, $h_{m''} \neq 0$) and $p_0, \ldots, p_{m'''} \in \mathbb{R}$ ($m''' \in \mathbb{N}$, $p_{m'''} \neq 0$) they suggest to use

$$C_{\mathrm{BV}}(y, h) = \sum_{i=0}^{m''} h_i (b_{\mathrm{BV}}(y, a, \bar{a}))^i \quad \text{and} \quad c_{\mathrm{BV}}(y, p) = \sum_{i=0}^{m'''} p_i (b_{\mathrm{BV}}(y, a, \bar{a}))^i.$$

Then they set

$$\mathcal{W}_{\mathrm{BV}}(y) = \left[\kappa_D(|i-j| / \max(b_{\mathrm{BV}}(y, a, \bar{a})n, 2))\right]_{i,j=1}^n,$$

and define, in correspondence with (23) and (24), the covariance estimators

$$\check{\Omega}^{\mathrm{BV},J}_{U,a,\bar{a},h,p}(y) = \hat{\omega}_{\mathcal{W}_{\mathrm{BV}}}(y) \exp\left(c_{\mathrm{BV}}(y,p) J^1_{n,U}(y)\right) R(X'X)^{-1}R'$$

and

$$\check{\Omega}^{\mathrm{BV}}_{a,\bar{a},h,p}(y) = \hat{\omega}_{\mathcal{W}_{\mathrm{BV}}}(y) \exp\left(c_{\mathrm{BV}}(y,p) n^{-2} \frac{\hat{u}'(y) A' A \hat{u}(y)}{\hat{u}'(y)\hat{u}(y)}\right) R(X'X)^{-1}R'.$$

The vectors of (constant) tuning parameters $a = (a_0, \ldots, a_{m'})'$, $\bar{a} = (\bar{a}_1, \ldots, \bar{a}_{m'})'$, $h = (h_0, \ldots, h_{m''})'$, and $p = (p_0, \ldots, p_{m'''})'$ this approach is based on are tabulated in Bunzel and Vogelsang (2005) for certain cases, and need to be obtained via simulations, following the rationale in Bunzel and Vogelsang (2005), for the cases not tabulated in that paper. Furthermore, the data-driven tuning parameters b_{BV} and c_{BV} as well as the data-driven critical value C_{BV} are well-defined for a given $y \in \mathbb{R}^n$ if and only if $\hat{\rho}(y)$ is well-defined, i.e., these quantities are well-defined on the complement of the closed set

$$\tilde{N} := \left\{ y \in \mathbb{R}^n : \sum_{i=1}^{n-1} \hat{u}_i^2(y) = 0 \right\}. \tag{26}$$

Clearly, $\mathrm{span}(X)$ is contained in $\tilde{N}$. Hence, it is not difficult to see that the estimator $\check{\Omega}^{\mathrm{BV}}_{a,\bar{a},h,p}$ is well-defined on $\mathbb{R}^n \backslash \tilde{N}$ and that the estimator $\check{\Omega}^{\mathrm{BV},J}_{U,a,\bar{a},h,p}$ is well-defined on $\mathbb{R}^n \backslash (\mathrm{span}(X,U) \cup \tilde{N})$. In fact, under Assumption 3 we have that $\tilde{N} = \mathrm{span}(X)$ (see the proof of the subsequent corollary). Consequently, under Assumption 3, the estimator $\check{\Omega}^{\mathrm{BV}}_{a,\bar{a},h,p}$ is well defined on $\mathbb{R}^n \backslash \mathrm{span}(X)$ and $\check{\Omega}^{\mathrm{BV},J}_{U,a,\bar{a},h,p}$ is well-defined on $\mathbb{R}^n \backslash \mathrm{span}(X,U)$. [In order that the data-driven critical value is also defined for every y, we set $C_{\mathrm{BV}}(y,h)$ equal to an arbitrary value (0, say) on the null-set $\tilde{N}$. Of course, the choice of assignment on this null-set is inconsequential for the result below.]

The following corollary shows that the tests for hypotheses concerning polynomial trends based on data-driven tuning parameters and a data-driven critical value as suggested in Bunzel and Vogelsang (2005) have size one in case $\mathfrak{F} \supseteq \mathfrak{F}^{ext}_{AR(2)}$. The proof of this is based on a similar approach as used in the proof of Corollary 5.7 above, but has to deal with the fact that the choice of the tuning parameters and the critical value is data-driven, and hence is more involved. In particular, it turns out that in order for Assumption 1 to be satisfied for the covariance estimators used here, one has to work with null-sets $N_{\mathrm{BV},U}$ and N_{BV} that are larger than $\mathrm{span}(X,U)$ and $\mathrm{span}(X)$, respectively.

Corollary 5.8. *Let $\mathfrak{F} \subseteq \mathfrak{F}_{\text{all}}$ satisfy $\mathfrak{F} \supseteq \mathfrak{F}_{\text{AR}(2)}^{ext}$ and suppose Assumption 3 holds. Let $a_i \in (0, \infty)$ for $i = 0, \ldots, m'$ ($m' \in \mathbb{N}$), $\bar{a}_i \in \mathbb{R}$ for $i = 1, \ldots, m'$, $h_i \in \mathbb{R}$ for $i = 0, \ldots, m''$ with $h_{m''} \neq 0$ and $m'' \in \mathbb{N}$, and $p_i \in \mathbb{R}$ for $i = 0, \ldots, m'''$ with $p_{m'''} \neq 0$ and $m''' \in \mathbb{N}$. Furthermore, for the statements that involve U, suppose U is an $n \times m$-dimensional matrix with $m \geq 1$ such that (X, U) is of full column-rank $k + m < n$. Then, $\check{\beta} = \hat{\beta}$ and $\check{\Omega} = \check{\Omega}^{\text{BV}}_{a,\bar{a},h,p}$ satisfies Assumption 1 with $N = N_{\text{BV}}$ (defined in Lemma D.6 in Appendix D), and $\check{\beta} = \hat{\beta}$ and $\check{\Omega} = \check{\Omega}^{\text{BV},J}_{U,a,\bar{a},h,p}$ satisfies Assumption 1 with $N = N_{\text{BV},U}$ (defined in Lemma D.6). Let T be of the form (4) with $\check{\beta} = \hat{\beta}$, $\check{\Omega} = \check{\Omega}^{\text{BV}}_{a,\bar{a},h,p}$, and $N = N_{\text{BV}}$, or with $\check{\beta} = \hat{\beta}$, $\check{\Omega} = \check{\Omega}^{\text{BV},J}_{U,a,\bar{a},h,p}$, and $N = N_{\text{BV},U}$. Then*

$$\sup_{f \in \mathfrak{F}} P_{\mu_0, \sigma^2 \Sigma(f)}(\{y \in \mathbb{R}^n : T(y) \geq C_{\text{BV}}(y, h)\}) = 1$$

holds for every $\mu_0 \in \mathfrak{M}_0$ and for every $\sigma^2 \in (0, \infty)$.

Remark 5.9. Alternatively one can consider T^*, where

$$T^*(y) = (R\hat{\beta}(y) - r)' \left(\check{\Omega}^{\text{BV}}_{a,\bar{a},h,p}(y)\right)^{-1} (R\hat{\beta}(y) - r)$$

for all $y \in \mathbb{R}^n \backslash \operatorname{span}(X)$ such that $\check{\Omega}^{\text{BV}}_{a,\bar{a},h,p}(y)$ is nonsingular, and where $T^*(y) = 0$ else, (and we can similarly define a test statistic T^{**} with $\check{\Omega}^{\text{BV},J}_{U,a,\bar{a},h,p}$ and $\operatorname{span}(X, U)$ in place of $\check{\Omega}^{\text{BV}}_{a,\bar{a},h,p}$ and $\operatorname{span}(X)$, respectively). While T^* and T^{**} are well-defined test statistics, we are not guaranteed that $\hat{\beta}$ and $\check{\Omega}^{\text{BV}}_{a,\bar{a},h,p}$ ($\hat{\beta}$ and $\check{\Omega}^{\text{BV},J}_{U,a,\bar{a},h,p}$, respectively) satisfy Assumption 1 with $N = \operatorname{span}(X)$ ($N = \operatorname{span}(X, U)$, respectively). However, T^* as well as T^{**} differ from the corresponding test statistics considered in the preceding corollary at most on a null-set, hence the conclusions of the corollary carry over to T^* and T^{**}.

5.2 Cyclical trends

We here consider briefly the case when one tests hypotheses concerning a cyclical trend, i.e., when the following assumption is satisfied:

Assumption 4. Suppose that $X = (E_{n,0}(\omega), \tilde{X})$ for some $\omega \in (0, \pi)$ where $\tilde{X}$ is an $n \times (k-2)$-dimensional matrix such that X has rank k (here $\tilde{X}$ is the empty matrix if $k = 2$). Furthermore, suppose that the restriction matrix R has a nonzero column $R_{\cdot i}$ for some $i = 1, 2$, i.e., the hypothesis involves coefficients of the cyclical component.

Under this assumption we obtain the subsequent theorem from Theorem 3.7.

Theorem 5.10. *Let $\mathfrak{F} \subseteq \mathfrak{F}_{\text{all}}$ satisfy $\mathfrak{F} \supseteq \mathfrak{F}_{\text{AR}(2)}$ and suppose Assumption 4 holds. Let T be a nonsphericity-corrected F-type test statistic of the form (4) based on $\check{\beta}$ and $\check{\Omega}$ satisfying Assumption 1 with $N = \emptyset$. Furthermore, assume that $\check{\Omega}(y)$ is nonnegative definite for every $y \in \mathbb{R}^n$. Then*

$$\sup_{f \in \mathfrak{F}} P_{\mu_0, \sigma^2 \Sigma(f)}(T \geq C) = 1$$

holds for every critical value C, $-\infty < C < \infty$, for every $\mu_0 \in \mathfrak{M}_0$, and for every $\sigma^2 \in (0,\infty)$.

Under a slightly stronger condition on $\mathfrak{F}$, the following theorem is applicable in case the assumption that $N = \emptyset$ or the nonnegative definiteness assumption on $\check{\Omega}$ in the previous theorem are violated.

Theorem 5.11. *Let $\mathfrak{F} \subseteq \mathfrak{F}_{\mathrm{all}}$ satisfy $\mathfrak{F} \supseteq \mathfrak{F}^{ext}_{\mathrm{AR}(2)}$. Suppose Assumption 4 holds. Let T be a nonsphericity-corrected F-type test statistic of the form (4) based on $\check{\beta}$ and $\check{\Omega}$ satisfying Assumption 1. Furthermore, assume that $\check{\Omega}$ also satisfies Assumption 2. Then for every critical value C, $-\infty < C < \infty$, for every $\mu_0 \in \mathfrak{M}_0$, and for every $\sigma^2 \in (0,\infty)$ it holds that*

$$P_{0,I_n}(\check{\Omega} \text{ is nonnegative definite}) \leq K(\omega) \leq \sup_{f \in \mathfrak{F}} P_{\mu_0,\sigma^2\Sigma(f)}\left(T \geq C\right),$$

where $K(\omega)$ is defined in Theorem 3.9.

Using these results, one can now obtain similar results as in Subsection 5.1.2 concerning the tests developed in Vogelsang (1998) and Bunzel and Vogelsang (2005) under Assumption 4. Due to space constraints, however, we do not spell out the details.

Remark 5.12. *(The cases $\omega = 0$ or $\omega = \pi$)* (i) In case $\omega = 0$ (or $\omega = \pi$) consider Assumption 4 with the understanding that $X = (\bar{E}_{n,0}(\omega), \tilde{X})$, that $\tilde{X}$ is now $n \times (k-1)$-dimensional, and that $R_{\cdot 1} \neq 0$, where $\bar{E}_{n,0}(\omega)$ denotes the first column of $E_{n,0}(\omega)$. Then Theorems 5.10 and 5.11 continue to hold with this interpretation of Assumption 4. Also note that the case $\omega = 0$ can be subsumed under the results of Subsection 5.1 by setting $k_F = 1$.

(ii) In case $\omega = 0$ (or $\omega = \pi$), Theorem 5.10 (with the before mentioned interpretation of Assumption 4) in fact continues to hold under the weaker assumption that $\mathfrak{F} \supseteq \mathfrak{F}_{\mathrm{AR}(1)}$. This follows from Part 3 of Corollary 5.17 in Preinerstorfer and Pötscher (2016) upon noting that $\mathcal{Z} = \mathrm{span}(\bar{E}_{n,0}(\omega))$ is a concentration space of the covariance model $\mathfrak{C}(\mathfrak{F})$, and that $\check{\Omega}$ vanishes on $\mathrm{span}(X) \supseteq \mathcal{Z}$ as a consequence of the assumption $N = \emptyset$ (see the discussion following (27) in Preinerstorfer and Pötscher (2016)).

(iii) In case $\omega = 0$ (or $\omega = \pi$), Theorem 5.11 (with the before mentioned interpretation of Assumption 4) also continues to hold under the weaker assumption that $\mathfrak{F} \supseteq \mathfrak{F}_{\mathrm{AR}(1)}$ if $\bar{\xi}_\omega(x)$ in the definition of $K(\omega)$ is now replaced by $\breve{\xi}_\omega(x)$ defined as

$$\breve{\xi}_\omega(x) = (R\hat{\beta}_X(\bar{E}_{n,0}(\omega)x))'\check{\Omega}^{-1}\left(\left(\left(\bar{E}_{n,0}(\omega)\bar{E}_{n,0}(\omega)'\right)^{1/2} + D(\omega)^{1/2}\right)\mathbf{G}\right)R\hat{\beta}_X(\bar{E}_{n,0}(\omega)x)$$

on the event where $\{((\bar{E}_{n,0}(\omega)\bar{E}_{n,0}(\omega)')^{1/2} + D(\omega)^{1/2})\mathbf{G} \in \mathbb{R}^n \backslash N^*\}$ and by $\breve{\xi}_\omega(x) = 0$ otherwise, and if the distribution P_{0,I_n} appearing in the lower bound is replaced by $P_{0,\Phi}$ where $\Phi = \bar{E}_{n,0}(\omega)\bar{E}_{n,0}(\omega)' + D(\omega)$. Here $D(0)$ is the matrix D given in Part 3 and $D(\pi)$ is the matrix D given in Part 4 of Lemma G.1 in Preinerstorfer and Pötscher (2016). This can be proved by making use of Theorem 5.19 and Lemma G.1 in Preinerstorfer and Pötscher (2016).

A Appendix: Proofs and auxiliary results for Section 3.1

Lemma A.1. *Let $\mathfrak{C}$ be a covariance model and let $\mathcal{L}$ be a linear subspace of $\mathbb{R}^n$ with $\dim(\mathcal{L}) = l < n$. Let $\mathfrak{C}^\sharp = \{\Sigma^\sharp : \Sigma \in \mathfrak{C}\}$ and $\mathfrak{C}^\natural = \{\Sigma^\natural : \Sigma \in \mathfrak{C}\}$, where $\Sigma^\sharp = \mathcal{L}(\Sigma) + \lambda_{l+1}(\mathcal{L}(\Sigma))\Pi_{\mathcal{L}}$ and where $\Sigma^\natural = \mathcal{L}(\Sigma) + \Pi_{\mathcal{L}}$; here $\lambda_{l+1}(\mathcal{L}(\Sigma))$ denotes the $(l+1)$-th eigenvalue of $\mathcal{L}(\Sigma)$ counting (with multiplicity) from smallest to largest. Then $\mathfrak{C}^\sharp$ and $\mathfrak{C}^\natural$ are covariance models. Furthermore, the collection of concentration spaces of $\mathfrak{C}^\sharp$ coincides with $\mathbb{J}(\mathcal{L}, \mathfrak{C})$, and the collection of concentration spaces of $\mathfrak{C}^\natural$ coincides with the collection $\{\mathcal{S} + \mathcal{L} : \mathcal{S} \in \mathbb{J}(\mathcal{L}, \mathfrak{C})\}$.*

Proof: 1. That $\mathfrak{C}^\sharp$ and $\mathfrak{C}^\natural$ are covariance models is obvious since the elements of these two collections are clearly symmetric and positive definite matrices (as $\lambda_{l+1}(\mathcal{L}(\Sigma)) > 0$ by construction).

2. Suppose $\mathcal{S} \in \mathbb{J}(\mathcal{L}, \mathfrak{C})$. Then $\mathcal{S} = \operatorname{span}(\bar{\Sigma})$ for some $\bar{\Sigma} \in \operatorname{cl}(\mathcal{L}(\mathfrak{C}))$ with $\operatorname{rank}(\bar{\Sigma}) < n - l$. In particular, $\bar{\Sigma}$ is the limit of $\mathcal{L}(\Sigma_m)$ for a sequence $\Sigma_m \in \mathfrak{C}$. But then $\Sigma_m^\sharp = \mathcal{L}(\Sigma_m) + \lambda_{l+1}(\mathcal{L}(\Sigma_m))\Pi_{\mathcal{L}}$ belongs to $\mathfrak{C}^\sharp$ and converges to $\bar{\Sigma}$ for $m \to \infty$, since $\lambda_{l+1}(\mathcal{L}(\Sigma_m))$ converges to $\lambda_{l+1}(\bar{\Sigma})$, which equals zero as a consequence of $\operatorname{rank}(\bar{\Sigma}) < n - l$. This shows that $\operatorname{span}(\bar{\Sigma})$, and hence $\mathcal{S}$, is a concentration space of $\mathfrak{C}^\sharp$. Conversely, suppose $\mathcal{Z}$ is a concentration space of $\mathfrak{C}^\sharp$. Then $\mathcal{Z} = \operatorname{span}(\breve{\Sigma})$ for some singular matrix that is the limit of some sequence $\Sigma_m^\sharp \in \mathfrak{C}^\sharp$. In particular, $\Sigma_m^\sharp = \mathcal{L}(\Sigma_m) + \lambda_{l+1}(\mathcal{L}(\Sigma_m))\Pi_{\mathcal{L}}$ holds for some sequence $\Sigma_m \in \mathfrak{C}$. Since the matrices $\mathcal{L}(\Sigma_m)$ reside in the unit sphere in $\mathbb{R}^{n\times n}$, we have convergence of $\mathcal{L}(\Sigma_{m_i})$ to a limit $\bar{\Sigma} \in \mathbb{R}^{n\times n}$ along an appropriate subsequence m_i; in particular, $\bar{\Sigma} \in \operatorname{cl}(\mathcal{L}(\mathfrak{C}))$ follows. Furthermore, we conclude that $\Sigma_{m_i}^\sharp$ converges to $\bar{\Sigma} + \lambda_{l+1}(\bar{\Sigma})\Pi_{\mathcal{L}}$, and hence obtain the equality $\breve{\Sigma} = \bar{\Sigma} + \lambda_{l+1}(\bar{\Sigma})\Pi_{\mathcal{L}}$. Since $\bar{\Sigma}$ is certainly symmetric and nonnegative definite, we have that $\lambda_{l+1}(\bar{\Sigma}) \geq 0$. Note that $\bar{\Sigma}x = 0$ for every $x \in \mathcal{L}$ by construction of $\bar{\Sigma}$. Hence $\operatorname{rank}(\bar{\Sigma}) \leq n - l$ must hold. If $\operatorname{rank}(\bar{\Sigma}) = n - l$ would hold we would have $\lambda_{l+1}(\bar{\Sigma}) > 0$, implying that $\bar{\Sigma} + \lambda_{l+1}(\bar{\Sigma})\Pi_{\mathcal{L}}$ is nonsingular, contradicting singularity of $\breve{\Sigma}$. Consequently, $\operatorname{rank}(\bar{\Sigma}) < n - l$ and $\lambda_{l+1}(\bar{\Sigma}) = 0$ must hold, implying that $\mathcal{S} = \operatorname{span}(\bar{\Sigma})$ belongs to $\mathbb{J}(\mathcal{L}, \mathfrak{C})$ and that $\breve{\Sigma} = \bar{\Sigma}$ holds. But this shows $\mathcal{Z} = \mathcal{S} \in \mathbb{J}(\mathcal{L}, \mathfrak{C})$.

3. Suppose $\mathcal{S} \in \mathbb{J}(\mathcal{L}, \mathfrak{C})$. Then $\mathcal{S} = \operatorname{span}(\bar{\Sigma})$ for some $\bar{\Sigma} \in \operatorname{cl}(\mathcal{L}(\mathfrak{C}))$ with $\operatorname{rank}(\bar{\Sigma}) < n - l$. In particular, $\bar{\Sigma}$ is the limit of $\mathcal{L}(\Sigma_m)$ for a sequence $\Sigma_m \in \mathfrak{C}$. But then $\Sigma_m^\natural = \mathcal{L}(\Sigma_m) + \Pi_{\mathcal{L}}$ belongs to $\mathfrak{C}^\natural$ and converges to $\bar{\Sigma} + \Pi_{\mathcal{L}}$ for $m \to \infty$. Now $\bar{\Sigma} + \Pi_{\mathcal{L}}$ is singular since $\operatorname{rank}(\bar{\Sigma}) < n - l$. Hence, $\operatorname{span}(\bar{\Sigma} + \Pi_{\mathcal{L}})$ is a concentration space of $\mathfrak{C}^\natural$ and $\operatorname{span}(\bar{\Sigma} + \Pi_{\mathcal{L}}) = \operatorname{span}(\bar{\Sigma}) + \mathcal{L} = \mathcal{S} + \mathcal{L}$ clearly holds. This proves one direction. Conversely, suppose $\mathcal{Z}$ is a concentration space of $\mathfrak{C}^\natural$. Then $\mathcal{Z} = \operatorname{span}(\breve{\Sigma})$ for some singular matrix that is the limit of some sequence $\Sigma_m^\natural \in \mathfrak{C}^\natural$, where $\Sigma_m^\natural = \mathcal{L}(\Sigma_m) + \Pi_{\mathcal{L}}$ for some $\Sigma_m \in \mathfrak{C}$. By the same compactness argument as before, we have $\mathcal{L}(\Sigma_{m_i}) \to \bar{\Sigma}$ implying that $\bar{\Sigma} \in \operatorname{cl}(\mathcal{L}(\mathfrak{C}))$. Furthermore, we immediately arrive at $\breve{\Sigma} = \bar{\Sigma} + \Pi_{\mathcal{L}}$. As before it follows that $\operatorname{rank}(\bar{\Sigma}) < n - l$ must hold and hence that $\mathcal{S} = \operatorname{span}(\bar{\Sigma}) \in \mathbb{J}(\mathcal{L}, \mathfrak{C})$. But then $\mathcal{Z} = \operatorname{span}(\breve{\Sigma}) = \operatorname{span}(\bar{\Sigma} + \Pi_{\mathcal{L}}) = \operatorname{span}(\bar{\Sigma}) + \mathcal{L}$ holds, implying the result. ■

Remark A.2. (i) By construction $\mathbb{J}(\mathcal{L}, \mathfrak{C}) = \mathbb{J}(\mathcal{L}, \mathfrak{C}^\sharp) = \mathbb{J}(\mathcal{L}, \mathfrak{C}^\natural)$. Furthermore, all three collections coincide with the collection of all concentration spaces of $\mathfrak{C}^\sharp$ (the union over which is $J(\mathfrak{C}^\sharp)$)

in the notation of Preinerstorfer and Pötscher (2016)).

(ii) The sum $\mathcal{S}+\mathcal{L}$ is an orthogonal sum and hence $\mathcal{S}$ is uniquely determined.

(iii) The map $\Sigma \mapsto \Sigma^{\sharp}$ is surjective from $\mathfrak{C}$ to $\mathfrak{C}^{\sharp}$ by definition, and the analogous statement holds for the map $\Sigma \mapsto \Sigma^{\natural}$. But these maps need not be injective.

Lemma A.3. *Let $\mathfrak{C}$ be a covariance model and let $\mathcal{L}$ be a linear subspace of $\mathbb{R}^n$ with $\dim(\mathcal{L}) < n$. Furthermore, let $W \subseteq \mathbb{R}^n$ be a rejection region of a test, which is $G(a+\mathcal{L})$-invariant for some $a \in \mathbb{R}^n$. Then for every σ, $0 < \sigma < \infty$, and every $\Sigma \in \mathfrak{C}$ we have*

$$P_{a,\sigma^2\Sigma}(W) = P_{a,\sigma^2\mathcal{L}(\Sigma)}(W) = P_{a,\sigma^2\Sigma^{\sharp}}(W) = P_{a,\sigma^2\Sigma^{\natural}}(W).$$

Furthermore, these probabilities do not depend on σ and they are unaffected if a is replaced by an arbitrary element of $a+\mathcal{L}$.

Proof: The first claim is essentially proved by the argument establishing (47) in Pötscher and Preinerstorfer (2016). The second claim is an immediate consequence of the assumed invariance (cf. also Proposition 5.4 in Preinerstorfer and Pötscher (2016)). ■

Proof of Theorem 3.1: By monotonicity w.r.t. C we may assume $C > 0$. Note that $\dim(\mathfrak{M}_0^{lin}) = k - q < n$ by our general model assumptions. Since T is $G(\mathfrak{M}_0)$-invariant by Lemma 5.15 in Preinerstorfer and Pötscher (2016), the preceding Lemma A.3, applied with $\mathcal{L} = \mathfrak{M}_0^{lin}$ and $a = \mu_0$, hence shows that it suffices to prove the theorem with $\mathfrak{C}$ replaced by $\mathfrak{C}^{\sharp}$. By Lemma A.1, also applied with $\mathcal{L} = \mathfrak{M}_0^{lin}$, the space $\mathcal{S}$ appearing in the formulation of the theorem is a concentration space of $\mathfrak{C}^{\sharp}$. We now apply Part 3 of Corollary 5.17 of Preinerstorfer and Pötscher (2016) to the linear model (1) considered in the present paper, but with $\mathfrak{C}$ replaced by $\mathfrak{C}^{\sharp}$. All assumptions of that result, except for the assumption that $\check{\Omega}(z) = 0$ and $R\check{\beta}(z) \neq 0$ simultaneously hold $\lambda_{\mathcal{S}}$-almost everywhere, are easily seen to be satisfied. We verify the remaining assumption now as follows: The discussion following (27) in Section 5.4 of Preinerstorfer and Pötscher (2016) shows that in case $N = \emptyset$ (which is assumed here) $\check{\Omega}(z) = 0$ holds for every $z \in \operatorname{span}(X)$, and thus for every $z \in \mathcal{S}$ (since $\mathcal{S} \subseteq \operatorname{span}(X)$ has been assumed). Hence, $\check{\Omega}(z) = 0$ $\lambda_{\mathcal{S}}$-almost everywhere follows. Furthermore, Assumption 1 together with $N = \emptyset$ imply that $\check{\beta}(X\gamma) = \check{\beta}(\varepsilon \cdot 0 + X\gamma) = \varepsilon\check{\beta}(0) + \gamma$ for every $\gamma \in \mathbb{R}^k$ and every $\varepsilon \neq 0$, which of course implies $\check{\beta}(X\gamma) = \gamma$ for every $\gamma \in \mathbb{R}^k$. Since we have assumed $\mathcal{S} \subseteq \operatorname{span}(X)$, it follows on the one hand that for every $z \in \mathcal{S}$ we have $R\check{\beta}(z) = 0$ if and only if $z \in \mathfrak{M}_0^{lin}$. On the other hand, by construction $\mathcal{S} \subseteq (\mathfrak{M}_0^{lin})^{\perp}$ holds, showing that $R\check{\beta}(z) \neq 0$ must hold for all nonzero $z \in \mathcal{S}$ in view of the fact that $\mathcal{S} \subseteq \operatorname{span}(X)$ has been assumed. Since $\mathcal{S}$ can not be zero-dimensional in view of its definition (cf. the discussion in Pötscher and Preinerstorfer (2016) following Definition 5.1), $\lambda_{\mathcal{S}}(\{0\}) = 0$ follows, which completes the proof. ■

A.1 Some comments on Lemmata A.1 and A.3

Lemmata A.1 and A.3 allow one to derive results regarding the rejection probabilities under a covariance model $\mathfrak{C}$ by working with a different, though related, covariance model $\mathfrak{C}^{\sharp}$. [Note that this covariance model has the property that its concentration spaces in the sense of Preinerstorfer and Pötscher (2016) are precisely given by the elements $\mathcal{S}$ of $\mathbb{J}(\mathcal{L}, \mathfrak{C})$.] A point in case is Theorem 3.1 in Section 3.1, which provides a "size one" result for the covariance model $\mathfrak{C}$, and which has been derived by applying Part 3 of Corollary 5.17 in Preinerstorfer and Pötscher (2016) to the covariance model $\mathfrak{C}^{\sharp}$, followed by an appeal to the aforementioned lemmata. In a similar vein one can combine other results of Preinerstorfer and Pötscher (2016) with these lemmata, but we do not spell this out here. Often this will lead to improvements over what one obtains from a direct application of the respective result of Preinerstorfer and Pötscher (2016) to the covariance model $\mathfrak{C}$. We illustrate this in the following by comparing the result in Theorem 3.1 with what one gets if instead one works with the originally given $\mathfrak{C}$ and directly applies Part 3 of Corollary 5.17 in Preinerstorfer and Pötscher (2016) to $\mathfrak{C}$.

Suppose $\mathfrak{C}$ and T are as in Theorem 3.1 (again with $N = \emptyset$ and nonnegative definiteness of $\check{\Omega}(y)$ for every $y \in \mathbb{R}^n$). Applying Part 3 of Corollary 5.17 in Preinerstorfer and Pötscher (2016) to the originally given covariance model $\mathfrak{C}$ allows one to obtain the following result: If a concentration space $\mathcal{Z}$ of $\mathfrak{C}$ exists that satisfies $\mathcal{Z} \subseteq \operatorname{span}(X)$ and $\mathcal{Z} \nsubseteq \mathfrak{M}_0^{lin}$, then (5) holds (for every C, every $\mu_0 \in \mathfrak{M}_0$, and every $\sigma^2 \in (0, \infty)$). [To see this note that by Corollary 5.17 in Preinerstorfer and Pötscher (2016) one only has to verify that $\check{\Omega}(z) = 0$ and $R\check{\beta}(z) \neq 0$ hold $\lambda_{\mathcal{Z}}$-almost everywhere. The argument for $\check{\Omega}(z) = 0$ $\lambda_{\mathcal{Z}}$-a.e. is identical to the corresponding argument given in the proof of Theorem 3.1. For the second claim a similar argument as in the proof of Theorem 3.1 shows that for $z \in \mathcal{Z}$ we have $R\check{\beta}(z) = 0$ if and only if $z \in \mathfrak{M}_0^{lin}$. In other words, $R\check{\beta}(z) = 0$ for $z \in \mathcal{Z}$ only occurs when $z \in \mathcal{Z} \cap \mathfrak{M}_0^{lin}$, which is a $\lambda_{\mathcal{Z}}$-null set, since $\mathcal{Z} \nsubseteq \mathfrak{M}_0^{lin}$.]

We now show that Theorem 3.1 is indeed at least as good a result as the result obtained in the preceding paragraph. For this it suffices to show that a concentration space $\mathcal{Z}$ of $\mathfrak{C}$ satisfying $\mathcal{Z} \subseteq \operatorname{span}(X)$ and $\mathcal{Z} \nsubseteq \mathfrak{M}_0^{lin}$ gives rise to an element $\mathcal{S} \in \mathbb{J}(\mathfrak{M}_0^{lin}, \mathfrak{C})$ satisfying the assumptions of Theorem 3.1: To see this, set $\mathcal{S} = \Pi_{(\mathfrak{M}_0^{lin})^{\perp}} \mathcal{Z}$ and observe that $\mathcal{S} \in \mathbb{J}(\mathfrak{M}_0^{lin}, \mathfrak{C})$ by Lemma 5.19(a) in Pötscher and Preinerstorfer (2016) (since $\Pi_{(\mathfrak{M}_0^{lin})^{\perp}} \mathcal{Z} \neq \{0\}$ in view of $\mathcal{Z} \nsubseteq \mathfrak{M}_0^{lin}$, and since $\Pi_{(\mathfrak{M}_0^{lin})^{\perp}} \mathcal{Z} \neq (\mathfrak{M}_0^{lin})^{\perp}$ in view of $\mathcal{Z} \subseteq \operatorname{span}(X)$, $\mathfrak{M}_0^{lin} \subseteq \operatorname{span}(X)$, and $\operatorname{rank}(X) < n$).[11] Furthermore, observe that $\mathcal{S} \subseteq \operatorname{span}(X)$ must also hold, since $\mathcal{Z} \subseteq \operatorname{span}(X)$ and $\mathfrak{M}_0^{lin} \subseteq \operatorname{span}(X)$.

Theorem 3.1 will sometimes actually give a strictly better result for the following reason (at least for covariance models $\mathfrak{C}$ that are bounded and bounded away from zero, an essentially costfree assumption in view of Remark 5.1(ii) in Pötscher and Preinerstorfer (2016)): Concentration spaces $\mathcal{Z}$ of $\mathfrak{C}$, that satisfy $\mathcal{Z} \subseteq \operatorname{span}(X)$ but also $\mathcal{Z} \subseteq \mathfrak{M}_0^{lin}$, can not be used in a direct application of Part 3 of Corollary 5.17 in Preinerstorfer and Pötscher (2016) since such spaces do not satisfy the relevant assumptions (note that $R\check{\beta}(z) = 0$ for all $z \in \mathcal{Z}$ holds for such spaces $\mathcal{Z}$); hence they do not help in establishing a result of the form (5) via a direct application of Part 3 of Corollary 5.17 in Preinerstorfer and Pötscher (2016). Nevertheless, such concentration spaces can have associated with them spaces $\mathcal{S} \in \mathbb{J}(\mathfrak{M}_0^{lin}, \mathfrak{C})$ in the way as described in Lemma

[11]Note that the boundedness assumptions on $\mathfrak{C}$ in Lemma 5.19 in Pötscher and Preinerstorfer (2016) have not been used in the proof of Part (a) of that lemma.

5.19(b) in Pötscher and Preinerstorfer (2016), that then will allow one to establish (5) via an application of Theorem 3.1 (provided the condition $\mathcal{S} \subseteq \operatorname{span}(X)$ can be shown to hold).

B Appendix: Proofs and auxiliary results for Section 3.2

Proof of Theorem 3.4: First, that $\mathcal{S} \subseteq \operatorname{span}(X)$ is equivalent to $\mathcal{A} \subseteq \operatorname{span}(X)$ where $\mathcal{A} := \operatorname{span}(E_{n,\rho(\gamma_1)}(\gamma_1), \ldots, E_{n,\rho(\gamma_p)}(\gamma_p))$ is obvious since any element of $\mathcal{A}$ is the sum of an element of $\mathcal{S}$ and an element of $\mathfrak{M}_0^{lin} \subseteq \operatorname{span}(X)$. Second, $\mathcal{S} \subseteq \operatorname{span}(X)$, $\mathfrak{M}_0^{lin} \subseteq \operatorname{span}(X)$, and the fact that $\mathcal{S}$ is certainly orthogonal to $\mathfrak{M}_0^{lin}$ imply $\dim(\mathcal{S}) + \dim(\mathfrak{M}_0^{lin}) \leq \dim(\operatorname{span}(X)) = k$. Since we always maintain $k < n$ we can conclude that $\dim(\mathcal{S}) < n - \dim(\mathfrak{M}_0^{lin})$ must hold. This together with Proposition 6.1 of Pötscher and Preinerstorfer (2016) now shows that the linear subspace figuring in the theorem belongs to $\mathbb{J}(\mathfrak{M}_0^{lin}, \mathfrak{C}(\mathfrak{F}))$ as clearly $\dim(\mathfrak{M}_0^{lin}) = k - q < n$ holds. An application of Theorem 3.1 then completes the proof. ■

Proof of Lemma 3.5: If $\{\gamma\} \in \mathbb{S}(\mathfrak{F}, \mathcal{L})$ holds, the definition of $\mathbb{S}(\mathfrak{F}, \mathcal{L})$ (Definition 6.3 in Pötscher and Preinerstorfer (2016)) immediately implies that $\kappa(\underline{\omega}(\mathcal{L}), \underline{d}(\mathcal{L})) + \kappa(\gamma, 1) < n$ must hold. To prove the converse, we first claim that there exists a sequence of spectral densities f_m in $\mathfrak{F}$ so that the sequence of spectral measures m_{g_m} defined by their spectral densities

$$g_m(\nu) = |\Delta_{\underline{\omega}(\mathcal{L}),\underline{d}(\mathcal{L})}(e^{\iota\nu})|^2 f_m(\nu) / \int_{-\pi}^{\pi} |\Delta_{\underline{\omega}(\mathcal{L}),\underline{d}(\mathcal{L})}(e^{\iota\nu})|^2 f_m(\nu) d\nu$$

converges weakly to a spectral measure m that satisfies $\operatorname{supp}(\mathsf{m}) \cap [0, \pi] = \{\gamma\}$. Here $\Delta_{\underline{\omega}(\mathcal{L}),\underline{d}(\mathcal{L})}$ is a certain differencing operator given in Definition 6.2 of Pötscher and Preinerstorfer (2016) and $\operatorname{supp}(\mathsf{m})$ denotes the support of m. To prove this claim, let $\rho_m \in (0, 1)$ converge to 1 as $m \to \infty$, and let ξ_j for $j \in \mathbb{N}$ be a sequence in $[0, \pi] \backslash \{0, \omega_1(\mathcal{L}), \ldots, \omega_p(\mathcal{L}), \pi\}$, where $\underline{\omega}(\mathcal{L}) = (\omega_1(\mathcal{L}), \ldots, \omega_p(\mathcal{L}))$, that converges to γ as $j \to \infty$. Now for every fixed $j \in \mathbb{N}$ the sequence of spectral measures $\mathsf{m}_{h_{m,j}}$ with spectral density

$$h_{m,j}(\nu) = (2\pi)^{-1} \frac{(1 - \rho_m^2)((1 + \rho_m^2)^2 - 4\rho_m^2 \cos^2(\xi_j))}{1 + \rho_m^2} \left|1 - \rho_m e^{-\iota\xi_j} e^{-\iota\nu}\right|^{-2} \left|1 - \rho_m e^{\iota\xi_j} e^{-\iota\nu}\right|^{-2}$$

converges weakly to $(\delta_{-\xi_j} + \delta_{\xi_j})/2$ as $m \to \infty$ (cf., e.g., the argument given in the proof of Lemma G.2 in Preinerstorfer and Pötscher (2016)). Note that $h_{m,j} \in \mathfrak{F}_{\mathrm{AR}(2)}$ and thus $h_{m,j} \in \mathfrak{F}$. Since $\xi_j \notin \{\omega_1(\mathcal{L}), \ldots, \omega_p(\mathcal{L})\}$, we can conclude that the map $\nu \mapsto \Delta_{\underline{\omega}(\mathcal{L}),\underline{d}(\mathcal{L})}(e^{\iota\nu})$ does not vanish on $\{-\xi_j, \xi_j\}$. It follows that the spectral measures $\mathsf{m}_{g_{m,j}}$ with spectral densities

$$g_{m,j}(\nu) = |\Delta_{\underline{\omega}(\mathcal{L}),\underline{d}(\mathcal{L})}(e^{\iota\nu})|^2 h_{m,j}(\nu) / \int_{-\pi}^{\pi} |\Delta_{\underline{\omega}(\mathcal{L}),\underline{d}(\mathcal{L})}(e^{\iota\nu})|^2 h_{m,j}(\nu) d\nu$$

also converge weakly to $(\delta_{-\xi_j} + \delta_{\xi_j})/2$, for fixed j and for $m \to \infty$. Since $(\delta_{-\xi_j} + \delta_{\xi_j})/2$ certainly converges weakly to $(\delta_{-\gamma} + \delta_{\gamma})/2$ as $j \to \infty$, a standard diagonal argument now delivers a sequence $f_m = h_{m,j(m)}$ as required above, for $j(m)$ a suitable subsequence of j. Together with the condition $\kappa(\underline{\omega}(\mathcal{L}), \underline{d}(\mathcal{L})) + \kappa(\gamma, 1) < n$ we see that $\{\gamma\} \in \mathbb{S}(\mathfrak{F}, \mathcal{L})$ follows. This proves the first claim. The second claim is a trivial consequence of the first claim, since $\kappa(\gamma, 1) = 1$ for $\gamma = 0, \pi$ and $\kappa(\gamma, 1) = 2$ for $\gamma \in (0, \pi)$. ■

Proof of Theorem 3.7: Since $\mathrm{span}(E_{n,\rho(\gamma)}(\gamma)) \subseteq \mathrm{span}(X)$ but $\mathrm{span}(E_{n,\rho(\gamma)}(\gamma)) \nsubseteq \mathfrak{M}_0^{lin} \subseteq \mathrm{span}(X)$ in view of the definition of $\rho(\gamma)$, it easily follows that

$$\kappa(\underline{\omega}(\mathfrak{M}_0^{lin}), \underline{d}(\mathfrak{M}_0^{lin})) + \kappa(\gamma, 1) \leq \kappa(\underline{\omega}(\mathrm{span}(X)), \underline{d}(\mathrm{span}(X)))$$

must hold. The r.h.s. of the above inequality is now not larger than k in view of Lemma D.1 in Pötscher and Preinerstorfer (2016). As we always maintain $k < n$, the first claim follows. Because of the claim just established and since $\mathfrak{F} \supseteq \mathfrak{F}_{\mathrm{AR}(2)}$, we conclude from Lemma 3.5 that $\{\gamma\} \in \mathbb{S}(\mathfrak{F}, \mathfrak{M}_0^{lin})$ (note that $\dim(\mathfrak{M}_0^{lin}) = k - q < n$ always holds). Set $\mathcal{S} = \mathrm{span}(\Pi_{(\mathfrak{M}_0^{lin})^\perp} E_{n,\rho(\gamma)}(\gamma))$ and observe that $\mathcal{S}$ satisfies all the conditions of Theorem 3.4 (recall that $\mathcal{S} \subseteq \mathrm{span}(X)$ if and only if $\mathrm{span}(E_{n,\rho(\gamma)}(\gamma)) \subseteq \mathrm{span}(X)$ holds as noted in that theorem). An application of Theorem 3.4 then establishes (7). ■

Lemma B.1. *For every $\gamma \in [0, \pi]$ and every $c > 0$ there exists a sequence $h_m \in \mathfrak{F}_{\mathrm{AR}(2)}^{ext}$ and a sequence σ_m^2 of positive real numbers such that*

$$\sigma_m^2 \Pi_{(\mathfrak{M}_0^{lin})^\perp} \Sigma(h_m) \Pi_{(\mathfrak{M}_0^{lin})^\perp} \to \Pi_{(\mathfrak{M}_0^{lin})^\perp} \left(E_{n,\rho(\gamma)}(\gamma) E'_{n,\rho(\gamma)}(\gamma) + cI_n \right) \Pi_{(\mathfrak{M}_0^{lin})^\perp} \text{ as } m \to \infty. \tag{27}$$

Proof: Let $\gamma \in [0, \pi]$ and $c > 0$ be given. For ease of notation we set $\mathcal{L} = \mathfrak{M}_0^{lin}$ in the remainder of the proof. We can use the argument in the proof of Lemma 3.5 to obtain a sequence of spectral densities f_m in $\mathfrak{F}_{\mathrm{AR}(2)}$ so that the sequence m_{g_m} with spectral density given by

$$g_m(\nu) = |\Delta_{\underline{\omega}(\mathcal{L}),\underline{d}(\mathcal{L})}(e^{\iota\nu})|^2 f_m(\nu) / \int_{-\pi}^{\pi} |\Delta_{\underline{\omega}(\mathcal{L}),\underline{d}(\mathcal{L})}(e^{\iota\nu})|^2 f_m(\nu) d\nu$$

converges weakly to the spectral measure $(\delta_{-\gamma} + \delta_\gamma)/2$. Now, set $e_m := \int_{-\pi}^{\pi} |\Delta_{\underline{\omega}(\mathcal{L}),\underline{d}(\mathcal{L})}(e^{\iota\nu})|^2 f_m(\nu) d\nu$, which is a sequence of positive real numbers (since $\Delta_{\underline{\omega}(\mathcal{L}),\underline{d}(\mathcal{L})}$ is a polynomial and f_m is nonzero a.e.). By Lemma D.2 in Pötscher and Preinerstorfer (2016) we have

$$\begin{aligned}
&e_m^{-1} \Pi_{\mathcal{L}^\perp} \Sigma(f_m) \Pi_{\mathcal{L}^\perp} \\
=&e_m^{-1} \Pi_{\mathcal{L}^\perp} H_n(\underline{\omega}(\mathcal{L}), \underline{d}(\mathcal{L})) \Sigma(\Delta_{\underline{\omega}(\mathcal{L}),\underline{d}(\mathcal{L})} \odot \mathsf{m}_{f_m}, n - \kappa(\underline{\omega}(\mathcal{L}), \underline{d}(\mathcal{L}))) H'_n(\underline{\omega}(\mathcal{L}), \underline{d}(\mathcal{L})) \Pi_{\mathcal{L}^\perp} \\
=&\Pi_{\mathcal{L}^\perp} H_n(\underline{\omega}(\mathcal{L}), \underline{d}(\mathcal{L})) \Sigma(\mathsf{m}_{g_m}, n - \kappa(\underline{\omega}(\mathcal{L}), \underline{d}(\mathcal{L}))) H'_n(\underline{\omega}(\mathcal{L}), \underline{d}(\mathcal{L})) \Pi_{\mathcal{L}^\perp} \\
\to&\Pi_{\mathcal{L}^\perp} H_n(\underline{\omega}(\mathcal{L}), \underline{d}(\mathcal{L})) E_{n-\kappa(\underline{\omega}(\mathcal{L}),\underline{d}(\mathcal{L})),0}(\gamma) E'_{n-\kappa(\underline{\omega}(\mathcal{L}),\underline{d}(\mathcal{L})),0}(\gamma) H'_n(\underline{\omega}(\mathcal{L}), \underline{d}(\mathcal{L})) \Pi_{\mathcal{L}^\perp}
\end{aligned}$$

as $m \to \infty$, where the convergence is due to weak convergence of m_{g_m} to $(\delta_{-\omega} + \delta_\omega)/2$; see Pötscher and Preinerstorfer (2016) for a definition of H_n, $\Sigma(\cdot, \cdot)$, as well as $\odot$. Lemma D.3 in the same reference now shows that the limit in the preceding display can be written as

$$a\Pi_{\mathcal{L}^\perp} E_{n,\rho(\gamma)}(\gamma) E'_{n,\rho(\gamma)}(\gamma)\Pi_{\mathcal{L}^\perp}$$

for some positive real number $a = a(\gamma)$. Now set $\sigma_m^2 = e_m^{-1}\left(a^{-1} + ce_m\right)$ and set

$$h_m = \left(a^{-1} f_m + (2\pi)^{-1} ce_m\right) / \left(a^{-1} + ce_m\right).$$

Observe that $h_m \in \mathfrak{F}_{\mathrm{AR}(2)}^{ext}$ holds. But then

$$\sigma_m^2 \Pi_{\mathcal{L}^\perp} \Sigma(h_m)\Pi_{\mathcal{L}^\perp} = a^{-1} e_m^{-1} \Pi_{\mathcal{L}^\perp} \Sigma(f_m)\Pi_{\mathcal{L}^\perp} + c\Pi_{\mathcal{L}^\perp}$$

obtains, implying (27). ■

Proof of Theorem 3.9: It suffices to prove the result for $C > 0$, which we henceforth assume. For ease of notation we set $\mathcal{L} = \mathfrak{M}_0^{lin}$ in the remainder of the proof. Let $\gamma \in [0, \pi]$ satisfy $\mathrm{span}(E_{n,\rho(\gamma)}(\gamma)) \subseteq \mathrm{span}(X)$. Observe that for $\mu_0 \in \mathfrak{M}_0$, $0 < \tau^2 < \infty$, and $h \in \mathfrak{F}_{\mathrm{AR}(2)}^{ext}$ it holds that

$$P_{\mu_0,\tau^2\Sigma(h)}(T \geq C) = P_{\mu_0,\tau^2\Pi_{\mathcal{L}^\perp}\Sigma(h)\Pi_{\mathcal{L}^\perp}}(T \geq C) = P_{\mu_0,\tau^2\left[\Pi_{\mathcal{L}^\perp}\Sigma(h)\Pi_{\mathcal{L}^\perp}+\Pi_{\mathcal{L}}\right]}(T \geq C). \tag{28}$$

This follows from $G(\mathfrak{M}_0)$-invariance of T and is proved in the same way as is relation (47) in Pötscher and Preinerstorfer (2016). Let now $c > 0$ and fix $\mu_0 \in \mathfrak{M}_0$, $0 < \sigma^2 < \infty$. By Lemma B.1 there exists a sequence $h_m \in \mathfrak{F}_{\mathrm{AR}(2)}^{ext}$ and a sequence σ_m^2 of positive real numbers such that

$$\sigma_m^2 \Pi_{\mathcal{L}^\perp}\Sigma(h_m)\Pi_{\mathcal{L}^\perp} + \Pi_{\mathcal{L}} \to \Pi_{\mathcal{L}^\perp}\left(E_{n,\rho(\gamma)}(\gamma)E'_{n,\rho(\gamma)}(\gamma) + cI_n\right)\Pi_{\mathcal{L}^\perp} + \Pi_{\mathcal{L}},$$

where the limit matrix is obviously nonsingular. Consequently,

$$P_{\mu_0,\sigma_m^2\left[\Pi_{\mathcal{L}^\perp}\Sigma(h_m)\Pi_{\mathcal{L}^\perp}+\Pi_{\mathcal{L}}\right]} \to P_{\mu_0,\Pi_{\mathcal{L}^\perp}E_{n,\rho(\gamma)}(\gamma)E'_{n,\rho(\gamma)}(\gamma)\Pi_{\mathcal{L}^\perp}+c\Pi_{\mathcal{L}^\perp}+\Pi_{\mathcal{L}}}$$

for $m \to \infty$ in total variation norm (by an application of Scheffé's Lemma). By $G(\mathfrak{M}_0)$-invariance of T we also have

$$P_{\mu_0,\sigma^2\Sigma(h_m)}(T \geq C) = P_{\mu_0,\sigma_m^2\Sigma(h_m)}(T \geq C),$$

cf. Remark 5.5(iii) in Preinerstorfer and Pötscher (2016). Using (28), the preceding displays now imply that

$$\begin{aligned} P_{\mu_0,\sigma^2\Sigma(h_m)}(T \geq C) &= P_{\mu_0,\sigma_m^2\left[\Pi_{\mathcal{L}^\perp}\Sigma(h_m)\Pi_{\mathcal{L}^\perp}+\Pi_{\mathcal{L}}\right]}(T \geq C) \\ &\to P_{\mu_0,\Pi_{\mathcal{L}^\perp}E_{n,\rho(\gamma)}(\gamma)E'_{n,\rho(\gamma)}(\gamma)\Pi_{\mathcal{L}^\perp}+c\Pi_{\mathcal{L}^\perp}+\Pi_{\mathcal{L}}}(T \geq C). \end{aligned}$$

The limit in the preceding display coincides – using again $G(\mathfrak{M}_0)$-invariance of T similarly as in

(28) – with

$$P_{\mu_0,\sigma^2\left[E_{n,\rho(\gamma)}(\gamma)E'_{n,\rho(\gamma)}(\gamma)+cI_n\right]}(T \geq C).$$

Since $\mathfrak{F} \supseteq \mathfrak{F}_{\mathrm{AR}(2)}^{ext}$ has been assumed and since $c > 0$ was arbitrary in the above discussion, it follows that $\sup_{f\in\mathfrak{F}} P_{\mu_0,\sigma^2\Sigma(f)}(T \geq C)$ is not smaller than $\sup_{\Sigma\in\mathfrak{C}(\gamma)} P_{\mu_0,\sigma^2\Sigma}(T \geq C)$ where $\mathfrak{C}(\gamma)$ denotes the auxiliary covariance model

$$\mathfrak{C}(\gamma) = \{E_{n,\rho(\gamma)}(\gamma)E'_{n,\rho(\gamma)}(\gamma) + cI_n : c > 0\}.$$

To prove the right-most inequality in (8) it hence suffices to verify that for every $\mu_0 \in \mathfrak{M}_0$ and every $0 < \sigma^2 < \infty$ it holds that

$$K(\gamma) \leq \sup_{\Sigma\in\mathfrak{C}(\gamma)} P_{\mu_0,\sigma^2\Sigma}(T \geq C). \tag{29}$$

To this end, we shall use Theorem 5.19 of Preinerstorfer and Pötscher (2016) applied to the linear model (1) together with the covariance model $\mathfrak{C}(\gamma)$. Let c_m be a sequence of positive real numbers satisfying $c_m \to 0$, and consider the corresponding sequence $\Sigma_m = E_{n,\rho(\gamma)}(\gamma)E'_{n,\rho(\gamma)}(\gamma) + c_mI_n$ in $\mathfrak{C}(\gamma)$. Obviously $\Sigma_m \to E_{n,\rho(\gamma)}(\gamma)E'_{n,\rho(\gamma)}(\gamma) =: \bar{\Sigma}$ and $\mathrm{span}(\bar{\Sigma}) = \mathrm{span}(E_{n,\rho(\gamma)}(\gamma))$ is $\kappa(\gamma,1)$-dimensional. Note that $\kappa(\gamma,1)$ is positive and that the $n \times n$-matrix $\bar{\Sigma}$ is singular because the assumption $\mathrm{span}(E_{n,\rho(\gamma)}(\gamma)) \subseteq \mathrm{span}(X)$ implies $\kappa(\gamma,1) \leq k < n$. Next, observe that

$$\Pi_{\mathrm{span}(\bar{\Sigma})^\perp}\Sigma_m\Pi_{\mathrm{span}(\bar{\Sigma})^\perp} = \Pi_{\mathrm{span}(E_{n,\rho(\gamma)}(\gamma))^\perp}\Sigma_m\Pi_{\mathrm{span}(E_{n,\rho(\gamma)}(\gamma))^\perp} = c_m\Pi_{\mathrm{span}(\bar{\Sigma})^\perp},$$

and that

$$\Pi_{\mathrm{span}(\bar{\Sigma})^\perp}\Sigma_m\Pi_{\mathrm{span}(\bar{\Sigma})} = 0.$$

Hence the additional assumption on Σ_m appearing in Theorem 5.19 of Preinerstorfer and Pötscher (2016) is satisfied with $s_m = c_m$ and $D = \Pi_{\mathrm{span}(E_{n,\rho(\gamma)}(\gamma))^\perp}$. Note also that $\mathrm{span}(\bar{\Sigma}) \subseteq \mathfrak{M} = \mathrm{span}(X)$ holds by our assumption on γ. Furthermore, since $\mathrm{span}(\bar{\Sigma}) = \mathrm{span}(E_{n,\rho(\gamma)}(\gamma))$ is not contained in $\mathcal{L} = \mathfrak{M}_0^{lin}$ in view of the definition of $\rho(\gamma)$, it follows that there exists a $z \in \mathrm{span}(\bar{\Sigma})$ so that $z \notin \mathcal{L}$. As both spaces are linear it even follows that $z \notin \mathcal{L}$ is true for $\lambda_{\mathrm{span}(\bar{\Sigma})}$-almost all $z \in \mathrm{span}(\bar{\Sigma})$. In view of the $\mathrm{span}(\bar{\Sigma}) \subseteq \mathrm{span}(X)$, this implies that $R\hat{\beta}(z) \neq 0$ holds $\lambda_{\mathrm{span}(\bar{\Sigma})}$-almost everywhere. Thus Theorem 5.19 of Preinerstorfer and Pötscher (2016) is applicable, and delivers (setting $Z = \bar{E}_{n,\rho(\gamma)}(\gamma)$ in that theorem) the claim (29), upon observing that in the definition of $\bar{\xi}(\gamma)$ in Theorem 5.19 of Preinerstorfer and Pötscher (2016) and in the event following that definition given in Theorem 5.19 of Preinerstorfer and Pötscher (2016) one can replace $\bar{\Sigma}^{1/2}$ by $\Pi_{\mathrm{span}(\bar{\Sigma})}$ due to $\mathrm{span}(\bar{\Sigma}) \subseteq \mathfrak{M}$, due to the equivariance property of $\check{\Omega}$ expressed in Assumption 1, and due to $G(\mathfrak{M})$-invariance of N^* (and noting that in the case considered here $\Pi_{\mathrm{span}(\bar{\Sigma})} + D^{1/2}$ translates into I_n). It remains to show the left-most inequality in (8). But this is obvious upon noting that the event where $\check{\Omega}(\mathbf{G})$ is nonnegative definite is contained in the event $\left\{\bar{\xi}_\gamma(x) \geq 0\right\}$ for every x. ■

C Appendix: Proofs for Section 4

Proof of Lemma 4.1: In view of $G(\mathfrak{M}_0)$-invariance of T we may set $\sigma^2 = 1$. In case $\mathbb{K}$ is empty there is nothing to prove. Hence assume $\mathbb{K} \neq \varnothing$. To prove Part 1, observe that then $C^*(\mathbb{K}) > -\infty$. Choose $C \in (-\infty, C^*(\mathbb{K}))$. Since $C < C^*(\mathbb{K})$, there exists an $\mathcal{S} \in \mathbb{K}$ with $C < C(\mathcal{S}) \leq C^*(\mathbb{K})$. Now repeat, with obvious modifications, the arguments in the proof of Part 2 of Lemma 5.10 of Pötscher and Preinerstorfer (2016) that establish (23) in that reference. To prove Part 2, observe that $C_*(\mathbb{K}) < \infty$, and choose $C \in (C_*(\mathbb{K}), \infty)$. Then there exists an $\mathcal{S} \in \mathbb{K}$ with $C_*(\mathbb{K}) \leq C(\mathcal{S}) < C$. Now repeat, with obvious modifications, the arguments in the proof of Part 3 of Lemma 5.10 of Pötscher and Preinerstorfer (2016). ■

Proof of Theorem 4.2: 1. Applying Part 1 of Lemma 4.1 with $\mathbb{K} = \{\mathcal{S}_1, \mathcal{S}_2\}$ shows that C satisfying (9) must also satisfy $C \in [C^*(\mathbb{K}), \infty)$. Since $C(\mathcal{S}_1) \neq C(\mathcal{S}_2)$ by assumption, it follows that $C_*(\mathbb{K}) < C^*(\mathbb{K})$. Hence, we arrive at $C > C_*(\mathbb{K})$, which in view of Part 2 of Lemma 4.1 implies (10).

2. The same reasoning, but now with $\mathbb{K} = \{\mathcal{S}\}$, where $\mathcal{S}$ is as in the theorem, yields $C \geq C(\mathcal{S})(= C_*(\mathbb{K}) = C^*(\mathbb{K}))$. Furthermore, note that $C > C(\mathcal{S})$ obviously implies $C > C_*(\mathbb{K})$ and thus (10) follows from Part 2 of Lemma 4.1.

3. From $G(\mathfrak{M}_0)$-invariance of T (cf. Footnote 5) we know that (10) implies

$$\inf_{\Sigma \in \mathfrak{C}} P_{\mu_0, \sigma^2\Sigma}(T \geq C) = 0$$

for every $\mu_0 \in \mathfrak{M}_0$ and every σ^2, $0 < \sigma^2 < \infty$. Since $G(\mathfrak{M}_0)$-invariance of T implies $G(\{\mu_0\})$-almost invariance of T for every $\mu_0 \in \mathfrak{M}_0$, (11) now follows from the preceding display together with Part 3 of Theorem 5.7 in Preinerstorfer and Pötscher (2016). Finally, (12) follows immediately from the preceding display by noting that for every $\Sigma \in \mathfrak{C}$ and every $\sigma^2 \in (0, \infty)$ the measures $P_{\mu_1, \sigma^2\Sigma}$ converge to $P_{\mu_0, \sigma^2\Sigma}$ in the total variation distance when μ_1 converges to μ_0 (cf. the proof of Theorem 5.7, Part 2, in Preinerstorfer and Pötscher (2016)). ■

Proof of Corollary 4.3: Set $\mathcal{V} = \{0\}$. The assumptions on T and on $N^\dagger = N^*$ in the second and third sentence of Theorem 4.2 are obviously satisfied in view of Lemma 5.15 in Pötscher and Preinerstorfer (2016). The assumption on the dimension of $\mathcal{L} := \mathfrak{M}_0^{lin}$ is also satisfied since we always maintain $k < n$. If (9) holds for a given C, Theorem 3.1 implies that any $\mathcal{S} \in \mathbb{J}(\mathfrak{M}_0^{lin}, \mathfrak{C})$ must satisfy $\mathcal{S} \subsetneqq \operatorname{span}(X)$; and thus $\mathcal{S} \subsetneqq N^*$, since $N^* = \operatorname{span}(X)$ is assumed in the corollary. Since N^* is $G(\mathfrak{M}_0)$-invariant, we also have $\mu_0 + \mathcal{S} \subsetneqq N^*$ for every $\mu_0 \in \mathfrak{M}_0$. As $\mu_0 + \mathcal{S}$ and N^* are affine subspaces of $\mathbb{R}^n$, this implies $\lambda_{\mu_0 + \mathcal{S}}(N^*) = 0$ for every $\mu_0 \in \mathfrak{M}_0$. Since $N^\dagger$ coincides with N^* for the class of test statistics considered, we obtain that $N^\dagger$ is a $\lambda_{\mu_0 + \mathcal{S}}$-null set for every $\mu_0 \in \mathfrak{M}_0$ and for every $\mathcal{S} \in \mathbb{J}(\mathfrak{M}_0^{lin}, \mathfrak{C})$, and thus a fortiori for every $\mathcal{S} \in \mathbb{H}$. We now see that Part 1 (Part 2, respectively) follows from the corresponding parts of Theorem 4.2 together with Part 3 of that theorem. ■

Proof of Lemma 4.4: Because of the assumption that $\mathfrak{F}$ contains $\mathfrak{F}_{\mathrm{AR}(2)}$ and that $\dim(\mathcal{L}) + 1 < n$, Lemma 3.5 implies (cf. Remark 3.6(i)) that $\{\gamma\} \in \mathbb{S}(\mathfrak{F}, \mathcal{L})$ for every $\gamma \in \{0, \pi\}$ (recall that $\kappa(\gamma,1)=1$ for these γ's). Furthermore, the dimension of

$$\mathcal{S} := \operatorname{span}\left(\Pi_{\mathcal{L}^{\perp}}\left(E_{n,\rho(\gamma,\mathcal{L})}(\gamma)\right)\right)$$

is 1 (since the dimension of $\operatorname{span}(E_{n,\rho(\gamma,\mathcal{L})}(\gamma))$ is 1 for $\gamma \in \{0,\pi\}$ and since $E_{n,\rho(\gamma,\mathcal{L})}(\gamma) \subsetneqq \mathcal{L}$ in view of the definition of $\rho(\gamma,\mathcal{L})$). Therefore the dimension of $\mathcal{S}$ is smaller than $n-\dim(\mathcal{L})$, and it follows from Proposition 6.1 in Pötscher and Preinerstorfer (2016) that $\mathcal{S} \in \mathbb{J}(\mathcal{L},\mathfrak{C}(\mathfrak{F}))$. ■

D Appendix: Auxiliary results and proofs for Section 5

Proof of Theorem 5.1: We first show that $\operatorname{span}(E_{n,\rho(0)}(0)) \subseteq \operatorname{span}(X)$ is satisfied: For any $i=1,\ldots,k_F$ with $R_{\cdot i} \neq 0$, the i-th column of F does not belong to $\mathfrak{M}_0^{lin}$. Observe that the i-th column of F spans $\operatorname{span}(E_{n,i-1}(0))$. Hence $\rho(0)$ must satisfy $0 \leq \rho(0) \leq k_F - 1$. But clearly $\operatorname{span}(E_{n,\rho(0)}(0)) \subseteq \operatorname{span}(F) \subseteq \operatorname{span}(X)$. All the other assumptions being obviously satisfied, Theorem 3.7 completes the proof. ■

Proof of Theorem 5.2: We apply Theorem 3.9. It suffices to verify that $\gamma = 0$ satisfies the assumption $\operatorname{span}(E_{n,\rho(\gamma)}(\gamma)) \subseteq \operatorname{span}(X)$ in that theorem. But this can be established exactly in the same way as in the proof of Theorem 5.1. It remains to verify that $K(0) = P_{0,I_n}(R'_{\cdot i_0}\check{\Omega}^{-1}R_{\cdot i_0} \geq 0)$: Recall that $\kappa(0,1)=1$, and note that

$$\bar{\xi}_0(x) = x^2\bar{\xi}_0(1) \quad \text{for every } x \in \mathbb{R}.$$

This is trivial on the event $\{\mathbf{G} \in N^*\}$. On the complement of this event, it follows from $\bar{E}_{n,\rho(0)}(0)$ being $n \times 1$-dimensional, and by using that $\hat{\beta}_X(\bar{E}_{n,\rho(0)}(0)x) = x\hat{\beta}_X(\bar{E}_{n,\rho(0)}(0))$ holds for every $x \in \mathbb{R}$. From the equation in the previous display, we now obtain $K(0) = \Pr(\bar{\xi}_0(1) \geq 0)$. To prove the statement, we thus need to show that $R\hat{\beta}_X(\bar{E}_{n,\rho(0)}(0))$ coincides with $R_{\cdot i_0}$, the first nonzero column of R. From a similar reasoning as in the proof of Theorem 5.1, we see that $\bar{E}_{n,i}(0) = F_{\cdot(i+1)}$ holds for $i = 0,\ldots,\rho(0)$. Hence, $\hat{\beta}_X(\bar{E}_{n,\rho(0)}(0)) = e_{\rho(0)+1}(k)$ holds. Furthermore, from the definition of $\rho(0)$, it follows that the first $\rho(0)$ columns of R are zero, and that the $(\rho(0)+1)$-th column of R is nonzero. The statement follows. ■

Lemma D.1. *Let $H \in \mathbb{R}^{n\times n}$ be nonsingular and define $\check{\beta}(y) = \hat{\beta}_{HX}(y) = (X'H'HX)^{-1}X'H'Hy$. Let $\nu : \mathbb{R}^n \backslash N' \to \mathbb{R}$, for N' a subset of $\mathbb{R}^n$, and set*

$$\check{\Omega}(y) = \nu(y)R(X'H'HX)^{-1}R' \quad \text{for every } y \notin N'.$$

Suppose that the following holds:

(a) N' is closed and $\lambda_{\mathbb{R}^n}(N') = 0$,

(b) $\delta y + X\eta \in \mathbb{R}^n\backslash N'$ and $\nu(\delta y + X\eta) = \delta^2\nu(y)$ holds for every $y \in \mathbb{R}^n\backslash N'$, every $\delta \neq 0$, and every $\eta \in \mathbb{R}^k$,

(c) ν is continuous on $\mathbb{R}^n\backslash N'$

(d) ν is $\lambda_{\mathbb{R}^n}$-almost everywhere nonzero on $\mathbb{R}^n \backslash N'$.

Then $\check{\beta}$ and $\check{\Omega}$ satisfy Assumption 1 with $N = N'$, and $\check{\Omega}$ satisfies Assumption 2. Furthermore, if ν is nonnegative (positive) everywhere on $\mathbb{R}^n \backslash N'$, then $\check{\Omega}(z)$ is nonnegative (positive) definite everywhere on $\mathbb{R}^n \backslash N'$.

Proof: Obviously $\check{\beta}$ is well-defined and continuous on all of $\mathbb{R}^n$, and thus also when restricted to $\mathbb{R}^n \backslash N'$. Furthermore, $\check{\Omega}$ is clearly well-defined and symmetric on $\mathbb{R}^n \backslash N'$, and is continuous on $\mathbb{R}^n \backslash N'$ in view of (c). Since N' is a closed $\lambda_{\mathbb{R}^n}$-null set by (a), we have verified Part (i) of Assumption 1 with $N = N'$. Part (ii) of this assumption is contained in (b). That $\check{\beta}$ satisfies the required equivariance property in Part (iii) of Assumption 1 is obvious. That $\check{\Omega}$ satisfies the required equivariance property in that assumption follows immediately from (b), completing the verification of Part (iii) of Assumption 1. Part (iv) in that assumption follows from (d) together with $R(X'H'HX)^{-1}R'$ being positive definite. The same argument also shows that $\check{\Omega}$ satisfies Assumption 2. The final statement is trivial. ■

Lemma D.2. *Suppose $\mathcal{W}$ is constant and symmetric, and that $\Pi_{\mathrm{span}(X)^\perp} \mathcal{W} \Pi_{\mathrm{span}(X)^\perp}$ is nonzero. Then the estimators $\hat{\beta}$ and $\check{\Omega}_{\mathcal{W}}$ satisfy Assumption 1 with $N = \emptyset$, and $\check{\Omega}_{\mathcal{W}}$ satisfies Assumption 2. If, additionally, $\Pi_{\mathrm{span}(X)^\perp} \mathcal{W} \Pi_{\mathrm{span}(X)^\perp}$ is nonnegative definite, then $\check{\Omega}_{\mathcal{W}}(y)$ is nonnegative definite for every $y \in \mathbb{R}^n$.*

Proof: We verify (a)-(d) in Lemma D.1 for $H = I_n$, $\nu = \hat{\omega}_{\mathcal{W}}$, and $N' = \emptyset$. Obviously (a) is satisfied, and (c) follows immediately from the constancy assumption on $\mathcal{W}$, since $\nu = \hat{\omega}_{\mathcal{W}}$ can clearly be written as a quadratic form in y. Concerning (d), note that $\hat{\omega}_{\mathcal{W}}(y) = 0$ is equivalent to $y'\Pi_{\mathrm{span}(X)^\perp} \mathcal{W} \Pi_{\mathrm{span}(X)^\perp} y = 0$. In view of the constancy assumption on $\mathcal{W}$, the subset of $\mathbb{R}^n$ on which $\hat{\omega}_{\mathcal{W}}$ vanishes is the zero set of a multivariate polynomial, in fact of a quadratic form, on $\mathbb{R}^n$. Since the (constant) matrix $\Pi_{\mathrm{span}(X)^\perp} \mathcal{W} \Pi_{\mathrm{span}(X)^\perp}$ is symmetric and nonzero, the polynomial under consideration does not vanish everywhere on $\mathbb{R}^n$, implying that the zero set is a $\lambda_{\mathbb{R}^n}$-null set. This completes the verification of (d). That (b) is satisfied follows immediately from $\nu(y) = \hat{\omega}_{\mathcal{W}}(y) = n^{-1} y'\Pi_{\mathrm{span}(X)^\perp} \mathcal{W} \Pi_{\mathrm{span}(X)^\perp} y$, the constancy of $\mathcal{W}$, and from $\Pi_{\mathrm{span}(X)^\perp}(\delta y + X\eta) = \delta \Pi_{\mathrm{span}(X)^\perp}(y)$ for every $\delta \in \mathbb{R}$, every $y \in \mathbb{R}^n$ and every $\eta \in \mathbb{R}^k$. Now apply Lemma D.1. Note that the final statement concerning nonnegative definiteness follows from the last part of Lemma D.1, since nonnegative definiteness of $\Pi_{\mathrm{span}(X)^\perp} \mathcal{W} \Pi_{\mathrm{span}(X)^\perp}$ obviously implies nonnegativity of $\hat{\omega}_{\mathcal{W}}$ on $\mathbb{R}^n$. ■

Proof of Corollary 5.3: The statement follows upon combining Lemma D.2 with Theorem 5.1. ■

Proof of Corollary 5.4: The first part of the corollary follows upon combining Lemma D.2 with Theorem 5.2 noting that $\check{\Omega}_{\mathcal{W}}(z)$ is nonnegative definite if and only if $\hat{\omega}_{\mathcal{W}}(z) \geq 0$. For the second statement, note that $R(X'X)^{-1}R'$ is positive definite, and hence

$$\{T \geq 0\} = \{\hat{\omega}_{\mathcal{W}} \geq 0\} \cup \{R\hat{\beta} = r\},$$

from which it follows (note that $\{y : R\hat{\beta}(y) = r\}$ is an affine subspace of $\mathbb{R}^n$ that does not coincide with $\mathbb{R}^n$, and is hence a $\lambda_{\mathbb{R}^n}$-null set) that $P_{\mu,\sigma^2 I_n}(T \geq 0)$ coincides with $P_{\mu,\sigma^2 I_n}(\hat{\omega}_{\mathcal{W}} \geq 0)$. For $C \geq 0$ we then have (using monotonicity w.r.t. C)

$$\sup_{\mu \in \mathfrak{M}_1} \sup_{0<\sigma^2<\infty} P_{\mu,\sigma^2 I_n}(T \geq C) \leq \sup_{\mu \in \mathfrak{M}_1} \sup_{0<\sigma^2<\infty} P_{\mu,\sigma^2 I_n}(\hat{\omega}_{\mathcal{W}} \geq 0). \tag{30}$$

But from the equivariance property $\hat{\omega}_{\mathcal{W}}(\delta y + X\eta) = \delta^2 \hat{\omega}_{\mathcal{W}}(y)$ for $\delta \neq 0$, $y \in \mathbb{R}^n$ and $\eta \in \mathbb{R}^k$, which was established in the proof of Lemma D.2, it follows straightforwardly that $P_{\mu,\sigma^2 I_n}(\hat{\omega}_{\mathcal{W}} \geq 0) = P_{0,I_n}(\hat{\omega}_{\mathcal{W}} \geq 0)$ holds for every $\mu \in \mathfrak{M}$ and every $0 < \sigma < \infty$. This completes the proof. ■

Lemma D.3. *If $N_{\mathrm{AM}} \neq \mathbb{R}^n$, then the estimators $\hat{\beta}$ and $\check{\Omega}_{\mathcal{W}_{\mathrm{AM}}}$ satisfy Assumption 1 with $N = N_{\mathrm{AM}}$, and $\check{\Omega}_{\mathcal{W}_{\mathrm{AM}}}$ satisfies Assumption 2; furthermore $\check{\Omega}_{\mathcal{W}_{\mathrm{AM}}}(z)$ is positive definite for every $z \in \mathbb{R}^n \backslash N_{\mathrm{AM}}$.*

Proof: Observe that $\hat{\rho}$, $\tilde{\rho}$, M_{AM}, $\mathcal{W}_{\mathrm{AM}}$, and $\hat{\omega}_{\mathcal{W}_{\mathrm{AM}}}$ are well-defined on $\mathbb{R}^n \backslash N_{\mathrm{AM}}$. We next verify (a)-(d) in Lemma D.1 for $H = I_n$, $\nu = \hat{\omega}_{\mathcal{W}_{\mathrm{AM}}}$, and $N' = N_{\mathrm{AM}}$. We start with (a): Using arguments as in the proof of Lemma 3.9 in Preinerstorfer (2017), or in the proof of Lemma B.1 in Preinerstorfer and Pötscher (2016), it is not difficult to verify that N_{AM} is an algebraic set. We leave the details to the reader. This, and the assumption $N_{\mathrm{AM}} \neq \mathbb{R}^n$, implies that N_{AM} is a closed $\lambda_{\mathbb{R}^n}$-null set. To verify (c) in Lemma D.1 it suffices to establish continuity of $\mathcal{W}_{\mathrm{AM}}$ on $\mathbb{R}^n \backslash N_{\mathrm{AM}}$, since $\hat{u}(y)$ is certainly continuous on $\mathbb{R}^n$. To achieve this note that, since $\hat{\rho}$ is obviously continuous on $\mathbb{R}^n \backslash N_{\mathrm{AM}}$, since $\hat{\rho}(y) \neq 1$ for $y \in \mathbb{R}^n \backslash N_{\mathrm{AM}}$, and since $\mathsf{A}(\cdot)$ is continuous on $\mathbb{R}$, it suffices to verify that $[\kappa_{\mathrm{QS}}(|i-j|/M_{\mathrm{AM}})]_{i,j=1}^{n-1}$ is continuous on $\mathbb{R}^n \backslash N_{\mathrm{AM}}$. Now, M_{AM} is certainly continuous on $\mathbb{R}^n \backslash N_{\mathrm{AM}}$ and κ_{QS} is continuous on $\mathbb{R}$. Hence, $[\kappa_{\mathrm{QS}}(|i-j|/M_{\mathrm{AM}})]_{i,j=1}^{n-1}$ is easily seen to be continuous at every $y \in \mathbb{R}^n \backslash N_{\mathrm{AM}}$ that satisfies $M_{\mathrm{AM}}(y) \neq 0$. For $y \in \mathbb{R}^n \backslash N_{\mathrm{AM}}$ satisfying $M_{\mathrm{AM}}(y) = 0$ continuity of $[\kappa_{\mathrm{QS}}(|i-j|/M_{\mathrm{AM}})]_{i,j=1}^{n-1}$ follows from continuity of M_{AM} on $\mathbb{R}^n \backslash N_{\mathrm{AM}}$ together with $\kappa_{\mathrm{QS}}(x) \to 0$ as $|x| \to \infty$, $\kappa_{\mathrm{QS}}(0) = 1$, and the convention $[\kappa_{\mathrm{QS}}(|i-j|/M_{\mathrm{AM}}(y))]_{i,j=1}^{n-1} = I_{n-1}$ for y so that $M_{\mathrm{AM}}(y) = 0$. That (b) in Lemma D.1 holds is easily seen to follow from $\hat{u}(\delta y + X\eta) = \delta \hat{u}(y)$ for every $\delta \in \mathbb{R}$, every $y \in \mathbb{R}^n$ and every $\eta \in \mathbb{R}^k$, which in particular implies $\hat{\rho}(\delta y + X\eta) = \hat{\rho}(y)$ and $\tilde{\rho}(\delta y + X\eta) = \tilde{\rho}(y)$ for every $\delta \neq 0$, every $y \in \mathbb{R}^n \backslash N_{\mathrm{AM}}$ and every $\eta \in \mathbb{R}^k$. Finally, note that (d) in Lemma D.1 is satisfied, because $\hat{\omega}_{\mathcal{W}_{\mathrm{AM}}}(y) > 0$ holds if $y \in \mathbb{R}^n \backslash N_{\mathrm{AM}}$. The latter follows from the well-known fact that $[\kappa_{\mathrm{QS}}(|i-j|/M_{\mathrm{AM}}(y))]_{i,j=1}^{n-1}$ is positive definite in case $M_{\mathrm{AM}}(y)$ is well-defined (recall that this matrix is defined as I_{n-1} in case $M_{\mathrm{AM}}(y) = 0$), together with the observation that $y \in \mathbb{R}^n \backslash N_{\mathrm{AM}}$ implies $\mathsf{A}(\hat{\rho}(y))\hat{u}(y) = \hat{v}(y) \neq 0$. Now apply Lemma D.1. Note that the just established fact, that $\hat{\omega}_{\mathcal{W}_{\mathrm{AM}}}(y) > 0$ holds if $y \in \mathbb{R}^n \backslash N_{\mathrm{AM}}$, also shows that the last part of Lemma D.1 applies, and hence shows that $\check{\Omega}_{\mathcal{W}_{\mathrm{AM}}}(y)$ is positive definite for every $y \in \mathbb{R}^n \backslash N_{\mathrm{AM}}$. ■

Proof of Corollary 5.5: This follows upon combining Lemma D.3 and Theorem 5.2, noting that the lower bound obtained via Theorem 5.2 equals 1 due to nonnegative definiteness of $\check{\Omega}_{\mathcal{W}_{\mathrm{AM}}}(y)$ for every $y \in \mathbb{R}^n \backslash N_{\mathrm{AM}}$, which is the complement of a $\lambda_{\mathbb{R}^n}$-null set. ■

Lemma D.4. *Let $V \in \{A, I_n\}$, $c \in \mathbb{R}$, let $i \in \{1, 2\}$, and let U be an $n \times m$-dimensional matrix* *with $m \geq 1$ such that (X, U) is of full column-rank $k + m < n$. Then the estimators $\check{\beta}_V$ and $\check{\Omega}^{\mathrm{Vo}}_{c,U,i,V}$ satisfy Assumption 1 with $N = \operatorname{span}(X, U)$, and $\check{\Omega}^{\mathrm{Vo}}_{c,U,i,V}$ also satisfies Assumption 2; furthermore $\check{\Omega}^{\mathrm{Vo}}_{c,U,i,V}$ is positive definite on $\mathbb{R}^n \backslash \operatorname{span}(X, U)$.*

Proof: We verify (a)-(d) in Lemma D.1 for $H = V$ (which is invertible), $\nu = n^{j(V)} s^2_{A,X} \exp(cJ^i_{n,U})$, and $N' = \operatorname{span}(X, U)$. By assumption, $k + m < n$, hence $\operatorname{span}(X, U)$ is a closed $\lambda_{\mathbb{R}^n}$-null set, showing that (a) in Lemma D.1 is satisfied. Next, note that $s^2_{A,X}$, $s^2_{I_n,(X,U)}$, and $s^2_{A,(X,U)}$ are well-defined and continuous on $\mathbb{R}^n$; and that $J^1_{n,U}$ and $J^2_{n,U}$ are well-defined and continuous on the set where $s^2_{I_n,(X,U)}$ and $s^2_{A,(X,U)}$ are nonzero, respectively. Obviously, $s^2_{I_n,(X,U)}(y) = 0$ if and only if $y \in \operatorname{span}(X, U)$. Similarly, $s^2_{A,(X,U)}(y) = 0$ if and only if $Ay \in \operatorname{span}(A(X, U))$, or equivalently, $y \in \operatorname{span}(X, U)$. Hence (c) in Lemma D.1 follows. For (b) note first that obviously $\delta y + X\eta \notin \operatorname{span}(X, U)$ holds for every $y \notin \operatorname{span}(X, U)$, every $\delta \neq 0$ and every $\eta \in \mathbb{R}^k$. Second, concerning the equivariance property of ν, we note that for every $y \in \mathbb{R}^n$, every $\delta \in \mathbb{R}$, and every $\eta \in \mathbb{R}^k$

$$s^2_{A,X}(\delta y + X\eta) = \delta^2 s^2_{A,X}(y) \tag{31}$$

$$s^2_{I_n,(X,U)}(\delta y + X\eta) = \delta^2 s^2_{I_n,(X,U)}(y) \tag{32}$$

$$s^2_{A,(X,U)}(\delta y + X\eta) = \delta^2 s^2_{A,(X,U)}(y). \tag{33}$$

From Equations (31)-(33) we hence see that the required equivariance property follows if we can show that

$$J^i_{n,U}(\delta y + X\eta) = J^i_{n,U}(y) \text{ for every } y \in \mathbb{R}^n \backslash \operatorname{span}(X, U), \text{ every } \delta \neq 0, \text{ and every } \eta \in \mathbb{R}^k. \tag{34}$$

To see this let $y \in \mathbb{R}^n \backslash \operatorname{span}(X, U)$, $\delta \neq 0$, and $\eta \in \mathbb{R}^k$. We consider first the case where $i = 1$. Note that $G\hat{\beta}_{(X,U)}(\delta y + X\eta) = \delta G\hat{\beta}_{(X,U)}(y)$, and recall from (32) that $s^2_{I_n,(X,U)}(\delta y + X\eta) = \delta^2 s^2_{I_n,(X,U)}(y) > 0$ (positivity following from $y \notin \operatorname{span}(X, U)$), showing that $J^1_{n,U}(\delta y + X\eta) = J^1_{n,U}(y)$. For $i = 2$, note that $G\hat{\beta}'_{A(X,U)}(A(\delta y + X\eta)) = \delta G\hat{\beta}'_{A(X,U)}(Ay)$, and recall from (33) that $s^2_{A,(X,U)}(\delta y + X\eta) = \delta^2 s^2_{A,(X,U)}(y) > 0$ (positivity following from $y \notin \operatorname{span}(X, U)$), showing that $J^2_{n,U}(\delta y + X\eta) = J^2_{n,U}(y)$. This verifies the statement in (34) and thus (b). Concerning (d) (and the final claim in the lemma) note that for $y \notin \operatorname{span}(X, U)$ it holds that $s^2_{A,X}(y) \exp(cJ^i_{n,U}(y)) > 0$. ■

Proof of Corollary 5.6: This follows upon combining Lemma D.4 and Theorem 5.2, noting that the lower bound obtained via Theorem 5.2 equals 1 due to nonnegative definiteness of $\check{\Omega}^{\mathrm{Vo}}_{c,U,i,V}$ on the complement of the $\lambda_{\mathbb{R}^n}$-null set $\operatorname{span}(X, U)$. ■

Lemma D.5. *Suppose that $\mathcal{W}$ is constant and symmetric, that $\Pi_{\operatorname{span}(X)^\perp} \mathcal{W} \Pi_{\operatorname{span}(X)^\perp}$ is nonzero, and that $c \in \mathbb{R}$. Then the following holds:*

1. *If U is an $n \times m$-dimensional matrix with $m \geq 1$ such that (X, U) is of full column-rank $k + m < n$, then the estimators $\hat{\beta}$ and $\check{\Omega}^{\mathrm{BV},J}_{\mathcal{W},U,c}$ satisfy Assumption 1 with $N = \operatorname{span}(X, U)$,*

and $\check{\Omega}^{\mathrm{BV},J}_{\mathcal{W},U,c}$ *satisfies Assumption 2. If, additionally,* $\Pi_{\mathrm{span}(X)^\perp}\mathcal{W}\Pi_{\mathrm{span}(X)^\perp}$ *is nonnegative definite, then* $\check{\Omega}^{\mathrm{BV},J}_{\mathcal{W},U,c}$ *is nonnegative definite on* $\mathbb{R}^n\backslash\operatorname{span}(X,U)$.

2. *The estimators* $\hat{\beta}$ *and* $\check{\Omega}^{\mathrm{BV}}_{\mathcal{W},c}$ *satisfy Assumption 1 with* $N = \operatorname{span}(X)$, *and* $\check{\Omega}^{BV}_{\mathcal{W},c}$ *satisfies Assumption 2. If, additionally,* $\Pi_{\mathrm{span}(X)^\perp}\mathcal{W}\Pi_{\mathrm{span}(X)^\perp}$ *is nonnegative definite, then* $\check{\Omega}^{\mathrm{BV}}_{\mathcal{W},c}$ *is nonnegative definite on* $\mathbb{R}^n\backslash\operatorname{span}(X)$.

Proof: 1. We verify (a)-(d) in Lemma D.1 for $H = I_n$, $\nu = \hat{\omega}_{\mathcal{W}}\exp(cJ^1_{n,U})$, and $N' = \operatorname{span}(X,U)$. That (a) holds follows from the same argument as in the proof of Lemma D.4. That (c) holds, follows from continuity of $\hat{\omega}_{\mathcal{W}}$ on $\mathbb{R}^n$ (cf. the proof of Lemma D.2), together with continuity of $\exp(cJ^1_{n,U})$ on the complement of $\operatorname{span}(X,U)$ (cf. the proof of Lemma D.4). The first Part of (b) was established in the proof of Lemma D.4. The second Part of (b) follows from the corresponding equivariance property of $\hat{\omega}_{\mathcal{W}}$, which was verified in the proof of Lemma D.2, together with the invariance property in Equation (34) established in the proof of Lemma D.4. Part (d) follows from $\hat{\omega}_{\mathcal{W}}(y) \neq 0$ for $\lambda_{\mathbb{R}^n}$-almost every $y \in \mathbb{R}^n$ (cf. the proof of Lemma D.2) together with $\exp(cJ^1_{n,U}(y)) > 0$ for every $y \notin \operatorname{span}(X,U)$. The final claim follows from the final statement in Lemma D.1 since (cf. the proof of Lemma D.2) $\hat{\omega}_{\mathcal{W}}(y) \geq 0$ holds for every $y \in \mathbb{R}^n$ in case $\Pi_{\mathrm{span}(X)^\perp}\mathcal{W}\Pi_{\mathrm{span}(X)^\perp}$ is nonnegative definite.

2. The proof is very similar to the proof of the first part. It follows along the same lines observing that the function defined via

$$y \mapsto \frac{\hat{u}'(y)A'A\hat{u}'(y)}{\hat{u}'(y)\hat{u}'(y)}$$

is well-defined and continuous on $\mathbb{R}^n\backslash\operatorname{span}(X)$, and is $G(\mathfrak{M})$-invariant. We skip the details. ■

Proof of Corollary 5.7: Noting that

$$P_{0,I_n}(\check{\Omega}\text{ is nonnegative definite}) = P_{0,I_n}(\hat{\omega}_{\mathcal{W}} \geq 0)$$

in our present context, the first part follows upon combining Lemma D.5 and Theorem 5.2 (the statement concerning the lower bound being 1 if $\Pi_{\mathrm{span}(X)^\perp}\mathcal{W}\Pi_{\mathrm{span}(X)^\perp}$ is nonnegative definite follows from nonnegative definiteness of $\check{\Omega}^{\mathrm{BV},J}_{\mathcal{W},U,c}$, or of $\check{\Omega}^{\mathrm{BV}}_{\mathcal{W},c}$, respectively, on the complement of $\lambda_{\mathbb{R}^n}$-null sets in that case). For the last part of the corollary, we can apply a similar argument as the one that was given to verify the analogous statement in Corollary 5.4: Note that now $\{T \geq 0\} = \{\hat{\omega}_{\mathcal{W}} \geq 0\} \cup \{R\hat{\beta} = r\} \cup N$, where $N = \operatorname{span}(X,U)$ if T is based on $\check{\Omega}^{\mathrm{BV},J}_{\mathcal{W},U,c}$, and $N = \operatorname{span}(X)$ if T is based on $\check{\Omega}^{\mathrm{BV}}_{\mathcal{W},c}$. In both cases $N \cup \{R\hat{\beta} = r\}$ is a $\lambda_{\mathbb{R}^n}$-null set, and we see that (30) holds also in the situation of the present lemma. The remainder of the proof is now analogous to the argument given at the end of the proof of Corollary 5.4. ■

Lemma D.6. *Let* $a_i \in (0,\infty)$ *for* $i = 0,\ldots,m'$ $(m' \in \mathbb{N})$, $\bar{a}_i \in \mathbb{R}$ *for* $i = 1,\ldots,m'$, $h_i \in \mathbb{R}$ *for* $i = 0,\ldots,m''$ $(m'' \in \mathbb{N})$ *with* $h_{m''} \neq 0$, *and* $p_i \in \mathbb{R}$ *for* $i = 0,\ldots,m'''$ $(m''' \in \mathbb{N})$ *with* $p_{m'''} \neq 0$. *Suppose further that* $e_n(n) \notin \operatorname{span}(X)^\perp$. *Then,* $\tilde{N} = \operatorname{span}(X)$, *and the following holds:*

1. *If U is an $n \times m$-dimensional matrix with $m \geq 1$ such that (X, U) is of full column-rank $k + m < n$, then the estimators $\hat{\beta}$ and $\check{\Omega}^{\mathrm{BV},J}_{U,a,\bar{a},h,p}$ satisfy Assumption 1 with $N = N_{\mathrm{BV},U}$, where $N_{\mathrm{BV},U} = \operatorname{span}(X, U) \cup \{y \in \mathbb{R}^n \backslash \operatorname{span}(X, U) : \hat{\rho}(y) \in \{\bar{a}_1, \ldots, \bar{a}_{m'}\}\}$ in case $\hat{\rho}$ attains at least two different values on $\mathbb{R}^n \backslash \operatorname{span}(X)$, and $N_{\mathrm{BV},U} = \operatorname{span}(X, U)$ else. Furthermore, $\check{\Omega}^{\mathrm{BV},J}_{U,a,\bar{a},h,p}$ satisfies Assumption 2, and $\check{\Omega}^{\mathrm{BV},J}_{U,a,\bar{a},h,p}(y)$ is positive definite for every $y \in \mathbb{R}^n \backslash N_{\mathrm{BV},U}$ (in fact, for $y \in \mathbb{R}^n \backslash \operatorname{span}(X, U)$).*

2. *The estimators $\hat{\beta}$ and $\check{\Omega}^{\mathrm{BV}}_{a,\bar{a},h,p}$ satisfy Assumption 1 with $N = N_{\mathrm{BV}}$, where $N_{\mathrm{BV}} = \operatorname{span}(X) \cup \{y \in \mathbb{R}^n \backslash \operatorname{span}(X) : \hat{\rho}(y) \in \{\bar{a}_1, \ldots, \bar{a}_{m'}\}\}$ in case $\hat{\rho}$ attains at least two different values on $\mathbb{R}^n \backslash \operatorname{span}(X)$, and $N_{\mathrm{BV}} = \operatorname{span}(X)$ else. Furthermore, $\check{\Omega}^{\mathrm{BV}}_{a,\bar{a},h,p}$ satisfies Assumption 2, and $\check{\Omega}^{\mathrm{BV}}_{a,\bar{a},h,p}(y)$ is positive definite for every $y \in \mathbb{R}^n \backslash N_{\mathrm{BV}}$ (in fact, for $y \in \mathbb{R}^n \backslash \operatorname{span}(X)$).*

Proof: The assumption $\operatorname{span}(e_n(n)) \notin \operatorname{span}(X)^{\perp}$ implies non-existence of a $y \in \mathbb{R}^n \backslash \operatorname{span}(X)$ so that $\sum_{i=1}^{n-1} \hat{u}_i^2(y) = 0$, showing that $\hat{\rho}$ is well-defined everywhere on $\mathbb{R}^n \backslash \operatorname{span}(X)$, i.e., that $\tilde{N} = \operatorname{span}(X)$. We consider two cases: First, assume that the design matrix X is such that $\hat{\rho} = \rho$ holds everywhere on $\mathbb{R}^n \backslash \operatorname{span}(X)$ for some fixed $\rho \in \mathbb{R}$. Then, the statements in 1. and 2., except for the positive definiteness claims, follow from Lemma D.5, because $b_{\mathrm{BV}}(., a, A)$ and $c_{\mathrm{BV}}(., p)$ are then constant equal to b and c, say, respectively, on $\mathbb{R}^n \backslash \operatorname{span}(X)$ and thus $\check{\Omega}^{\mathrm{BV},J}_{U,a,\bar{a},h,p}(y) = \check{\Omega}^{\mathrm{BV},J}_{\mathcal{W},U,c}(y)$ holds for every $y \notin \operatorname{span}(X, U)$, and $\check{\Omega}^{\mathrm{BV}}_{a,\bar{a},h,p}(y) = \check{\Omega}^{\mathrm{BV}}_{\mathcal{W},c}(y)$ holds for every $y \notin \operatorname{span}(X)$ where the matrix $\mathcal{W} = (\mathcal{W}_{ij}) = (\kappa_D(|i - j| / \max(bn, 2)))$. Observe here that $\mathcal{W}$ is constant in y, is symmetric, and is positive definite. The positive definiteness claims in 1. and 2. finally follow since $\hat{\omega}_{\mathcal{W}}(y) = 0$ holds for $y \in \mathbb{R}^n \backslash \operatorname{span}(X)$ in view of positive definiteness of $\mathcal{W}$.

Next, we consider the case where X is such that $\hat{\rho}$ attains at least two different values on $\mathbb{R}^n \backslash \operatorname{span}(X)$. We start with the statement in 1.: First of all, $N_{\mathrm{BV},U}$ is easily seen to be $G(\mathfrak{M})$-invariant (because $\hat{\rho}$ is so). Second, we can rewrite

$$N_{\mathrm{BV},U} = \bigcup_{i=1}^{m'} \left\{ y \in \mathbb{R}^n : \sum_{i=2}^{n} \hat{u}_i(y)\hat{u}_{i-1}(y) - \bar{a}_i \sum_{i=1}^{n-1} \hat{u}_i^2(y) = 0 \right\} \cup \operatorname{span}(X, U).$$

From that we see that $N_{\mathrm{BV},U}$ is a finite union of algebraic sets, and hence an algebraic set. Thus, $N_{\mathrm{BV},U}$ is closed. Since we also work under the hypothesis that $\hat{\rho}$ attains at least two different values on $\mathbb{R}^n \backslash \operatorname{span}(X)$, we can conclude that

$$\left\{ y \in \mathbb{R}^n : \sum_{i=2}^{n} \hat{u}_i(y)\hat{u}_{i-1}(y) - \bar{a}_i \sum_{i=1}^{n-1} \hat{u}_i^2(y) = 0 \right\} \neq \mathbb{R}^n$$

holds for every $i = 1, \ldots, m'$. It follows that the algebraic set in the previous display is a $\lambda_{\mathbb{R}^n}$-null set for every $i = 1, \ldots, m'$. Hence $N_{\mathrm{BV},U}$ is a closed $\lambda_{\mathbb{R}^n}$-null set as $\operatorname{span}(X, U) \neq \mathbb{R}^n$. To prove the statements of 1., we now verify (a)-(d) in Lemma D.1 for $H = I_n$, $\nu(.) = \hat{\omega}_{\mathcal{W}_{\mathrm{BV}}}(.) \exp(c_{\mathrm{BV}}(., p) J^1_{n,U}(.))$, and $N' = N_{\mathrm{BV},U}$. We have already verified (a). Furthermore, note that $b_{\mathrm{BV}}(.,a,\bar{a})$ is continuous on $\mathbb{R}^n \backslash N_{\mathrm{BV},U}$. As a consequence, $c_{\mathrm{BV}}(.,p)$, and $\mathcal{W}_{\mathrm{BV}}(.)$, and thus $\hat{\omega}_{\mathcal{W}_{\mathrm{BV}}}$ are continuous on $\mathbb{R}^n \backslash N_{\mathrm{BV},U}$. We already know from the proof of Lemma D.4 that $J_{n,U}^1$ is continuous on the complement of $\mathrm{span}(X,U) \subseteq N_{\mathrm{BV},U}$. It thus follows that $y \mapsto \hat{\omega}_{\mathcal{W}_{\mathrm{BV}}}(y) \exp(c_{\mathrm{BV}}(y,p) J_{n,U}^1(y))$ is continuous on $\mathbb{R}^n \backslash N_{\mathrm{BV},U}$. Hence, we have verified (c) in Lemma D.1. To verify (b) we recall from above that $N_{\mathrm{BV},U}$ is $G(\mathfrak{M})$-invariant. Furthermore, the required equivariance property in (b) holds as a consequence of $G(\mathfrak{M})$-invariance of $\hat{\rho}$ and $J_{n,U}^1$ (cf. (34)), and hence of $c_{\mathrm{BV}}(.,p)$ and $\mathcal{W}_{\mathrm{BV}}(.)$ on $\mathbb{R}^n \backslash N_{\mathrm{BV},U}$, together with $\hat{u}(\delta y + X\eta) = \delta \hat{u}(y)$ for every $\delta \neq 0$, $y \in \mathbb{R}^n$ and $\eta \in \mathbb{R}^k$. That $\nu(y) = \hat{\omega}_{\mathcal{W}_{\mathrm{BV}}}(y) \exp(c_{\mathrm{BV}}(y,p) J_{n,U}^1(y))$ is even positive on $\mathbb{R}^n \backslash N_{\mathrm{BV},U}$ follows because $\mathcal{W}_{\mathrm{BV}}(y)$ is positive definite on $\mathbb{R}^n \backslash \tilde{N}$ and $\tilde{N} = \mathrm{span}(X)$ holds. This implies (d) in Lemma D.1, and also the sufficient condition for positive definiteness in the same lemma. The statements in 2. for the case where $\hat{\rho}$ attains at least two different values on $\mathbb{R}^n \backslash \mathrm{span}(X)$ are almost identical, and we skip the details. ■

Proof of Corollary 5.8: From Assumption 3 it follows that the last row of X is not equal to zero, i.e., $e_n(n) \notin \mathrm{span}(X)^{\perp}$ must hold. Hence, all assumptions of Lemma D.6 are satisfied. Combining this lemma with Theorem 5.2 proves the claims with $C_{\mathrm{BV}}(y,h)$ replaced by an arbitrary constant critical value C (noting that the lower bound obtained via Theorem 5.2 equals 1 due to nonnegative definiteness of $\check{\Omega}_{U,a,\bar{a},h,p}^{\mathrm{BV},J}$ ($\check{\Omega}_{a,\bar{a},h,p}^{\mathrm{BV}}$, respectively) on the complement of the $\lambda_{\mathbb{R}^n}$-null set $N_{\mathrm{BV},U}$ (N_{BV}, respectively)). But now we observe that $y \mapsto C_{\mathrm{BV}}(y,h)$ is well-defined on $\mathbb{R}^n$ (recall the convention preceding Corollary 5.8), and by construction takes on only finitely many real numbers $C_1 < \ldots < C_l$, say. Hence, for every $f \in \mathfrak{F}$, every $\mu_0 \in \mathfrak{M}_0$, every $\sigma^2 \in (0,\infty)$ we can conclude that

$$P_{\mu_0,\sigma^2\Sigma(f)}(\{y \in \mathbb{R}^n : T(y) \geq C_{\mathrm{BV}}(y,h)\}) \geq P_{\mu_0,\sigma^2\Sigma(f)}(\{y \in \mathbb{R}^n : T(y) \geq C_l\}).$$

Now apply what has been established before with $C = C_l$. This completes the proof ■

Proof of Theorem 5.10: For any $i = 1,2$ with $R_{\cdot i} \neq 0$, the i-th column of $E_{n,0}(\omega)$ does not belong to $\mathfrak{M}_0^{lin}$. Hence $\mathrm{span}(E_{n,0}(\omega)) \not\subseteq \mathfrak{M}_0^{lin}$, implying that $\rho(\omega)$ must be zero. However, $\mathrm{span}(E_{n,0}(\omega) \subseteq \mathrm{span}(X)$ clearly holds. All the other assumptions being obviously satisfied, Theorem 3.7 completes the proof. ■

Proof of Theorem 5.11: We apply Theorem 3.9. It suffices to verify that $\gamma = \omega$ satisfies the assumption $\mathrm{span}(E_{n,\rho(\gamma)}(\gamma)) \subseteq \mathrm{span}(X)$ in that theorem. But this can be established exactly in the same way as in the proof of Theorem 5.10. ■

References

Andrews, D. W. K. (1991). Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica*, **59** 817–858.

Andrews, D. W. K. and Monahan, J. C. (1992). An improved heteroskedasticity and autocorrelation consistent covariance matrix estimator. *Econometrica*, **60** 953–966.

Bunzel, H. and Vogelsang, T. J. (2005). Powerful trend function tests that are robust to strong serial correlation, with an application to the Prebisch-Singer hypothesis. *Journal of Business & Economic Statistics*, **23** 381–394.

Davies, R. B. (1980). Algorithm AS 155: The distribution of a linear combination of χ^2 random variables. *Journal of the Royal Statistical Society. Series C*, **29** 323–333.

Eicker, F. (1967). Limit theorems for regressions with unequal and dependent errors. In *Proc. Fifth Berkeley Sympos. Math. Statist. and Probability (Berkeley, Calif., 1965/66)*. Univ. California Press, Berkeley, Calif., Vol. I: Statistics, pp. 59–82.

Grenander, U. (1954). On the estimation of regression coefficients in the case of an autocorrelated disturbance. *The Annals of Mathematical Statistics*, **25** 252–272.

Harvey, D. I., Leybourne, S. J. and Taylor, A. R. (2007). A simple, robust and powerful test of the trend hypothesis. *Journal of Econometrics*, **141** 1302 – 1330.

Kiefer, N. M. and Vogelsang, T. J. (2002a). Heteroskedasticity-autocorrelation robust standard errors using the Bartlett kernel without truncation. *Econometrica*, **70** 2093–2095.

Kiefer, N. M. and Vogelsang, T. J. (2002b). Heteroskedasticity-autocorrelation robust testing using bandwidth equal to sample size. *Econometric Theory*, **18** 1350–1366.

Kiefer, N. M. and Vogelsang, T. J. (2005). A new asymptotic theory for heteroskedasticity-autocorrelation robust tests. *Econometric Theory*, **21** 1130–1164.

Kiefer, N. M., Vogelsang, T. J. and Bunzel, H. (2000). Simple robust testing of regression hypotheses. *Econometrica*, **68** 695–714.

Long, J. S. and Ervin, L. H. (2000). Using heteroscedasticity consistent standard errors in the linear regression model. *The American Statistician*, **54** 217–224.

Newey, W. K. and West, K. D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, **55** 703–708.

Newey, W. K. and West, K. D. (1994). Automatic lag selection in covariance matrix estimation. *The Review of Economic Studies*, **61** 631–653.

Perron, P. and Yabu, T. (2009). Estimating deterministic trends with an integrated or stationary noise component. *Journal of Econometrics*, **151** 56 – 69.

Pötscher, B. M. and Preinerstorfer, D. (2016). Controlling the size of autocorrelation robust tests. *arXiv:1612.06127*.

Preinerstorfer, D. (2016). acrt: Autocorrelation robust testing. Version 1.0. URL `https://CRAN.R-project.org/package=acrt`.

Preinerstorfer, D. (2017). Finite sample properties of tests based on prewhitened nonparametric covariance estimators. *Electron. J. Statist.*, **11** 2097–2167. URL `http://dx.doi.org/10.1214/17-EJS1281`.

Preinerstorfer, D. and Pötscher, B. M. (2016). On size and power of heteroskedasticity and autocorrelation robust tests. *Econometric Theory*, **32** 261–358.

Robinson, G. (1979). Conditional properties of statistical procedures. *Annals of Statistics*, **7** 742–755.

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