Finite sample properties of tests based on prewhitened nonparametric covariance estimators

David Preinerstorfer

CREATES and Department of Economics and Business Economics, Aarhus University
Fuglesangs Allé 4, 8210 Aarhus V, Denmark
e-mail: david.preinerstorfer@econ.au.dk

Abstract: We analytically investigate size and power properties of a popular family of procedures for testing linear restrictions on the coefficient vector in a linear regression model with temporally dependent errors. The tests considered are autocorrelation-corrected F-type tests based on prewhitened nonparametric covariance estimators that possibly incorporate a data-dependent bandwidth parameter, e.g., estimators as considered in Andrews and Monahan (1992), Newey and West (1994), or Rho and Shao (2013). For design matrices that are generic in a measure theoretic sense we prove that these tests either suffer from extreme size distortions or from strong power deficiencies. Despite this negative result we demonstrate that a simple adjustment procedure based on artificial regressors can often resolve this problem.

MSC 2010 subject classifications: Primary 62F03; Secondary 62J05, 62F35, 62M10, 62M15.
Keywords and phrases: Autocorrelation robustness, prewhitening, size distortion, power deficiency, artificial regressors.

Received March 2016.

Contents

1. Introduction

The construction of tests for hypotheses on the coefficient vector in linear regression models with dependent errors is highly practically relevant and has received lots of attention in the statistics and econometrics literature. The main challenge is to obtain tests with good size and power properties in situations where the nuisance parameter governing the dependence structure of the errors is high- or possibly infinite-dimensional and allows for strong correlations. The large majority of available procedures are autocorrelation-corrected F-type tests, based on nonparametric covariance estimators trying to take into account the autocorrelation in the disturbances. The tests currently used can roughly be categorized into two groups, the distinction depending on the choice of the critical values. The first group of such tests is based on critical values obtained from an asymptotic framework in which the nonparametric covariance estimators are consistent, and where the asymptotic distribution of the F-type test statistic under the null hypothesis is a χ^2 distribution. Quantiles of this limiting distribution are then used for testing. Concerning these tests, important contributions in the econometrics literature are Newey and West (1987), Andrews (1991), Andrews and Monahan (1992), and Newey and West (1994). It is safe to say that using F-type tests based on χ^2 critical values and the covariance estimators introduced in the latter two articles currently constitutes the gold standard for the testing problem under consideration. In contrast to the estimator suggested earlier by Newey and West (1987) - structurally 2π times a standard kernel spectral density estimator (Bartlett (1950), Jowett (1955), Hannan (1957), and Grenander and Rosenblatt (1957) Section 7.9) evaluated at frequency 0 - the covariance estimators suggested in Andrews and Monahan (1992) and Newey and West (1994) both incorporate an additional prewhitening step based on an auxiliary vector autoregressive (VAR) model, as well as a data-dependent bandwidth parameter. A distinguishing feature of the estimators introduced by Andrews and Monahan (1992) on the one hand and Newey and West (1994) on the other hand is the choice of the bandwidth parameter: Andrews and Monahan (1992) used an approach introduced by Andrews (1991), where the bandwidth parameter is chosen based on auxiliary parametric models. In contrast to that, Newey and West (1994) suggested a nonparametric approach for choosing the bandwidth parameter. Even though simulation studies have shown that the inclusion of a prewhitening step and the data-dependent choice of the bandwidth parameter can improve the finite sample properties of the tests obtained, these more so-

phisticated tests still suffer from size distortions and power deficiencies. For this reason Kiefer, Vogelsang and Bunzel (2000), Kiefer and Vogelsang (2002), and Kiefer and Vogelsang (2005) suggested a different asymptotic framework for obtaining critical values. Their framework, in which the bandwidth parameter is a fixed proportion of the sample size, leads to inconsistent covariance estimators and to a non-standard limiting distribution of the corresponding test statistic under the null hypothesis, the quantiles of which are used to obtain tests. In simulation studies it has been observed that these tests still suffer from size distortions in finite samples, but less so than tests based on χ^2 critical values. However, this is at the expense of some loss in power. Furthermore, simulation results in Kiefer and Vogelsang (2005) and Rho and Shao (2013) suggest that the finite sample properties of tests based on non-standard critical values can be improved by incorporating a prewhitening step. In the latter paper it was also shown that the asymptotic distribution under the null of the test suggested by Kiefer, Vogelsang and Bunzel (2000) is the same whether or not prewhitening is used.

A number of recent studies (Velasco and Robinson (2001), Jansson (2004), Sun, Phillips and Jin (2008, 2011), Zhang and Shao (2013)) tried to use higher order expansions to uncover the mechanism leading to size distortions and power deficiencies of the above mentioned tests. These higher-order asymptotic results (and also the first-order results discussed above) are pointwise in the sense that they are obtained under the assumption of a fixed underlying data-generating-process. Hence, while they inform us about the limit of the rejection probability and the rate of convergence to this limit for a fixed underlying data-generating-process, they do not inform us about the *size* of the test or its limit as sample size increases, nor about the *power function* or its asymptotic behavior. Size and power properties of tests in regression models with dependent errors were recently studied in Preinerstorfer and Pötscher (2016): In a general finite sample setup and under high-level conditions on the structure of the test and the covariance model, they derived conditions on the design matrix under which a concentration mechanism due to strong dependencies leads to extreme size distortions or power deficiencies. Furthermore, they suggested an adjustment-procedure to obtain a modified test with improved size and power properties. Specializing their general theory to a covariance model that includes at least all covariance matrices corresponding to stationary autoregressive processes of order one (AR(1)), they investigated finite sample properties of F-type tests based on *non-prewhitened* covariance estimators with *data-independent* bandwidth parameters (covering *inter alia* the procedures in Newey and West (1987), Sections 3-5 of Andrews (1991), Hansen (1992), Kiefer, Vogelsang and Bunzel (2000), Kiefer and Vogelsang (2002, 2005), Jansson (2002, 2004), but *not* the methods considered by Andrews and Monahan (1992), Newey and West (1994) or Rho and Shao (2013)). In this setup Preinerstorfer and Pötscher (2016) demonstrated that these tests break down in terms of their size or power behavior for generic design matrices. Despite this negative result, they also showed that the adjustment procedure can often solve these problems, if elements of the covariance model which are close to being singular can be well approximated by AR(1) covariance matrices.

Preinerstorfer and Pötscher (2016), however, did not consider tests based on prewhitened covariance estimators or data-dependent bandwidth parameters. Therefore the question remains, whether the more sophisticated tests typically used in practice, i.e., tests based on χ^2 critical values and the estimators by Andrews and Monahan (1992) or Newey and West (1994), and the prewhitened tests based on non-standard critical values and data-independent bandwidth parameters, i.e., tests as considered in Rho and Shao (2013), also suffer from extreme size distortions and power deficiencies, or if prewhitening and the use of data-dependent bandwidth parameters can indeed resolve or at least substantially alleviate these problems. In the present paper we investigate finite sample properties of tests based on prewhitened covariance estimators or data-dependent bandwidth parameters. In particular our analysis covers tests based on prewhitened covariance estimators using auxiliary AR(1) models for the construction of the bandwidth parameter as discussed in Andrews and Monahan (1992), tests based on prewhitened covariance estimators as discussed in Newey and West (1994), and prewhitened tests based on non-standard critical values as discussed in Rho and Shao (2013). We show that the tests considered, albeit being structurally much more complex, exhibit a similar behavior as their *non-prewhitened* counterparts with *data-independent* bandwidth parameters: First, we establish conditions on the design matrix under which the tests considered have (i) size equal to one, or (ii) size not smaller than one half, or (iii) nuisance-minimal power equal to zero, respectively. We then demonstrate that at least one of these conditions is generically satisfied, showing that the tests considered break down for generic design matrices.

It is important to stress that this generic negative result does not only apply to tests based on χ^2 critical values, or to tests based on one of the non-standard critical values mentioned above. The result is applicable to every F-type test based on one of the nonparametric covariance estimators considered and combined with *any* (data-independent) critical value $0 < C < \infty$. Hence, the problem described by our generic negative result can not be resolved by simply adjusting (data-independently) the critical value used.

Motivated by this negative result, we introduce an adjustment procedure. Under the assumption that elements of the covariance model which are close to being singular can be well approximated by AR(1) covariance matrices, we show that the adjustment procedure, if applicable, leads to tests that do not suffer from extreme size distortions or power deficiencies. Finally, it is shown that the adjustment procedure is applicable under generic conditions on the design matrix, unless the regression includes the intercept *and* the hypothesis to be tested restricts the corresponding coefficient. On a technical level we employ the general theory developed in Preinerstorfer and Pötscher (2016). We remark, however, that the genericity results in particular do not follow from this general theory. Rather they are obtained by studying and carefully exploiting the specific structure of the procedures under consideration.

The paper is organized as follows: The framework is introduced in Section 2. In Section 3 we introduce the test statistics, covariance estimators, and bandwidth parameters we analyze. In Section 4 we establish our negative result and

its genericity. In Section 5 we discuss the adjustment-procedure and its generic applicability. Numerical results are presented in Section 6. Section 7 concludes. The proofs are collected in Appendices B-D. Appendix E contains tables for the numerical results.

2. The framework

Consider the linear regression model

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

where X is a (real) $n \times k$ dimensional non-stochastic design matrix satisfying $n > 2$, $\mathrm{rank}(X) = k$ and $1 \leq k < n$. Here, $\beta \in \mathbb{R}^k$ denotes the unknown regression parameter vector, and the disturbance vector $\mathbf{U} = (\mathbf{u}_1, \ldots, \mathbf{u}_n)'$ is Gaussian, has mean zero and its unknown covariance matrix is given by $\sigma^2 \Sigma$. The parameter σ^2 satisfies $0 < \sigma^2 < \infty$ and Σ is assumed to be an element of a prescribed (non-void) set of positive definite and symmetric $n \times n$ matrices $\mathfrak{C}$, which we shall refer to as the *covariance model*. Throughout we impose the assumption on $\mathfrak{C}$ that the parameters σ^2 and Σ can be uniquely determined from $\sigma^2 \Sigma$.

Remark 2.1. The leading case we have in mind is the situation where $\mathbf{u}_1, \ldots, \mathbf{u}_n$ are n consecutive elements of a weakly stationary process. In such a setup a co-variance model is typically obtained from a prescribed (non-void) set of spectral densities $\mathcal{F}$. Assuming that no element of $\mathcal{F}$ vanishes identically almost everywhere, the covariance model corresponding to $\mathcal{F}$ is then given by

$$\mathfrak{C}(\mathcal{F}) = \{\Sigma(f) : f \in \mathcal{F}\},$$

with

$$\Sigma(f) = \left(\int_{-\pi}^{\pi} \exp(-\iota\lambda(i-j))f(\lambda)d\lambda \Big/ \int_{-\pi}^{\pi} f(\lambda)d\lambda \right)_{i,j=1}^{n}, \tag{2}$$

and where ι denotes the imaginary unit. Every such $\Sigma(f)$ is positive definite and symmetric. Furthermore, since $\Sigma(f)$ is a correlation matrix, σ^2 and $\Sigma(f)$ can uniquely be determined from $\sigma^2 \Sigma(f)$. As outlined in the Introduction the tests we focus on in this article are particularly geared towards setups where $\mathcal{F}$ is a nonparametric class of spectral densities, i.e., where the corresponding set $\mathfrak{C}(\mathcal{F})$ is rich. A typical example is the class $\mathcal{F}_\xi$, which consists of all spectral densities of linear processes the coefficients of which satisfy a certain summability condition, i.e., spectral densities of the form

$$f(\lambda) = (2\pi)^{-1} \left| \sum_{j=0}^{\infty} c_j \exp(-\iota j\lambda) \right|^2,$$

where, for a fixed $\xi \geq 0$, the summability condition $0 < \sum_{j=0}^{\infty} j^\xi |c_j| < \infty$ is satisfied. We observe that $\mathfrak{C}(\mathcal{F}_\xi)$ contains in particular all correlation matrices corresponding to spectral densities of stationary autoregressive moving average models of arbitrary large order.

The linear model described in (1) induces a collection of distributions on $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$, the sample space of $\mathbf{Y}$. Denoting a Gaussian probability measure with mean $\mu \in \mathbb{R}^n$ and covariance matrix $\sigma^2 \Sigma$ by $P_{\mu,\sigma^2\Sigma}$ and denoting the regression manifold by $\mathfrak{M} = \operatorname{span}(X)$, the induced collection of distributions is given by

$$\left\{ P_{\mu,\sigma^2\Sigma} : \mu \in \mathfrak{M}, 0 < \sigma^2 < \infty, \Sigma \in \mathfrak{C} \right\}. \tag{3}$$

Since every $\Sigma \in \mathfrak{C}$ is positive definite by definition, each element $P_{\mu,\sigma^2\Sigma}$ of the set in the previous display is absolutely continuous with respect to (w.r.t.) Lebesgue measure on $\mathbb{R}^n$.

In this setup we shall consider the problem of testing a linear 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 of rank $q \geq 1$ and $r \in \mathbb{R}^q$. Define the affine space

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

and let

$$\mathfrak{M}_1 = \mathfrak{M} \backslash \mathfrak{M}_0 = \{\mu \in \mathfrak{M} : \mu = X\beta \text{ and } R\beta \neq r\}.$$

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

$$H_0 : \mu \in \mathfrak{M}_0, \ 0 < \sigma^2 < \infty, \ \Sigma \in \mathfrak{C} \quad \text{vs.} \quad H_1 : \mu \in \mathfrak{M}_1, \ 0 < \sigma^2 < \infty, \ \Sigma \in \mathfrak{C}, \tag{4}$$

where it is emphasized that the testing problem is a compound one. It is immediately clear that size and power properties of tests in this setup depend in a crucial way on the richness of the covariance model $\mathfrak{C}$.

Before we close this section by introducing some further terminological and notational conventions, we comment on how the fixed-design-assumption and the Gaussianity-assumption above can be relaxed:

1. We remark that even though our setup assumes a non-stochastic design matrix, the results immediately carry over to a setting where the data generating processes of the design and the disturbances are independent of each other. In such a setup, which covers many relevant scenarios, e.g., most simulation examples in our key references Andrews and Monahan (1992), Newey and West (1994), and Rho and Shao (2013), our results then deliver *size and power properties conditional on the design*, which, in the tradition of conditional inference, might be considered as the more relevant criterion, because X is observable (e.g., Robinson (1979)).

2. The Gaussianity assumption might seem to be restrictive. However, as in Section 5.5 of Preinerstorfer and Pötscher (2016), we mention that the negative results given in Section 4 of the present paper immediately extend in a trivial way without imposing the Gaussianity assumption on the error vector $\mathbf{U}$ in (1), as long as the assumptions on the feasible error distributions are weak enough to ensure that the implied set of distributions for $\mathbf{Y}$ contains the set in Equation (3), but possibly contains also other distributions. Furthermore, by applying an invariance argument (explained in Preinerstorfer and Pötscher (2016) Section 5.5) one can easily

show that all statements about the null-behavior of the procedures under consideration derived in the present paper carry over to the more general distributional setup where $\mathbf{U}$ is assumed to be elliptically distributed. This is to be understood as $\mathbf{U}$ having the same distribution as $\mathbf{m}\sigma\Sigma^{1/2}\mathbf{E}$, where $0 < \sigma < \infty$, $\Sigma \in \mathfrak{C}$, $\mathbf{E}$ is a random vector uniformly distributed on the unit sphere S^{n-1}, and $\mathbf{m}$ is a random variable distributed independently of $\mathbf{E}$ and which is positive with probability one.

We next collect some further terminology and notation used throughout the whole paper. A (non-randomized) *test* is the indicator function of a set $W \in \mathcal{B}(\mathbb{R}^n)$, i.e., 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 $\hat{u}_X(y) = y - X\hat{\beta}_X(y)$. The subscript X is omitted whenever this does not cause confusion. Random vectors and random variables are always written in bold capital and bold lower case letters, respectively. We use Pr as a generic symbol for a probability measure and denote by E the corresponding expectation operator. Lebesgue measure on $\mathbb{R}^n$ will be denoted by $\lambda_{\mathbb{R}^n}$. The Euclidean norm is denoted by $\|\cdot\|$, while $d(x, A)$ denotes the Euclidean distance of the point $x \in \mathbb{R}^n$ to the set $A \subseteq \mathbb{R}^n$. For a vector x in Euclidean space we define the symbol $\langle x \rangle$ to denote $\pm x$ for $x \neq 0$, the sign being chosen in such a way that the first nonzero component of $\langle x \rangle$ is positive, and we set $\langle 0 \rangle = 0$. The j-th standard basis vector in $\mathbb{R}^n$ is denoted by $e_j(n)$. Let B' denote the transpose of a matrix B and let $\text{span}(B)$ denote the space spanned by its columns. 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 $\mathcal{L}$. The set of real matrices of dimension $m \times n$ is denoted by $\mathbb{R}^{m \times n}$. Lebesgue measure on this set equipped with its Borel σ-algebra is denoted by $\lambda_{\mathbb{R}^{m \times n}}$. We use the convention that the adjoint of a 1×1 dimensional matrix D, i.e., $\text{adj}(D)$, equals one. Given a vector $v \in \mathbb{R}^m$ the symbol $\text{diag}(v)$ denotes the $m \times m$ diagonal matrix with main diagonal v. We define

$$\mathfrak{X}_0 = \left\{ X \in \mathbb{R}^{n \times k} : \text{rank}(X) = k \right\},$$

i.e., the set of $n \times k$ design matrices of full rank, and whenever $k \geq 2$ we define

$$\tilde{\mathfrak{X}}_0 = \left\{ \tilde{X} \in \mathbb{R}^{n \times (k-1)} : \text{rank}((e_+, \tilde{X})) = k \right\},$$

which is canonically identified (as a set) with the set of $n \times k$ design matrices of full column rank the first column of which is the intercept $e_+ = (1, \ldots, 1)' \in \mathbb{R}^n$.

3. Tests based on prewhitened covariance estimators

In the present section we formally describe the construction of tests based on prewhitened covariance estimators. These tests (cf. Remark 3.4 below and the discussion preceding it) reject for large values of a statistic

$$T(y) = \begin{cases} (R\hat{\beta}(y) - r)'\hat{\Omega}^{-1}(y)(R\hat{\beta}(y) - r) & \text{if } y \notin N^*(\hat{\Omega}), \\ 0 & \text{else,} \end{cases} \tag{5}$$

where

$$\hat{\Omega}(y) = nR(X'X)^{-1}\hat{\Psi}(y)(X'X)^{-1}R',$$

and

$$N^*(\hat{\Omega}) = \left\{ y \in \mathbb{R}^n : \hat{\Omega}(y) \text{ is not invertible or not well defined} \right\}.$$

The quantity $\hat{\Psi}$ appearing in the definition of $\hat{\Omega}$ above denotes a (VAR-) prewhitened nonparametric estimator of $n^{-1}E(X'\mathbf{U}\mathbf{U}'X)$ that incorporates a bandwidth parameter which might depend on the data. Such an estimator is completely specified by three core ingredients: First, a *kernel* $\kappa : \mathbb{R} \to \mathbb{R}$, i.e., an even function satisfying $\kappa(0) = 1$, such as, e.g., the Bartlett or Parzen kernel; second, a (non-negative) possibly data-dependent *bandwidth parameter* M; and third, a deterministic *prewhitening order* p, i.e., an integer satisfying $1 \leq p \leq n/(k+1)$ (cf. Remark 3.2). Specific choices of M are discussed in detail in Section 3.1. All possible combinations of κ, M and p we analyze are specified in Assumption 1 of Section 3.2. Once these core ingredients have been chosen, one obtains a prewhitened estimator $\hat{\Psi}$, which is computed at an observation y following the Steps (1) - (3) outlined subsequently (cf. also den Haan and Levin (1997)). We here assume that the quantities involved (e.g., inverse matrices) are well defined, cf. Remark 3.1 below, and follow the convention in the literature and leave the estimator undefined at y else. Using this convention $\hat{\Psi}(y)$ is obtained as follows:

1. To prewhiten the data a VAR(p) model is fitted via ordinary least squares to the columns of $\hat{V}(y) = X' \operatorname{diag}(\hat{u}(y))$. One so obtains the VAR(p) residual matrix $\hat{Z}(y) \in \mathbb{R}^{k \times (n-p)}$ with columns

$$\hat{Z}_{\cdot(j-p)}(y) = \hat{V}_{\cdot j}(y) - \sum_{l=1}^{p} \hat{A}_l^{(p)}(y)\hat{V}_{\cdot(j-l)}(y) \qquad \text{for } j = p+1, \ldots, n.$$

The $k \times (kp)$-dimensional VAR(p)-OLS estimator is given by

$$\hat{A}^{(p)}(y) = \left(\hat{A}_1^{(p)}(y), \ldots, \hat{A}_p^{(p)}(y) \right) = \hat{V}_p(y)\hat{V}_0'(y) \left(\hat{V}_0(y)\hat{V}_0'(y) \right)^{-1},$$

where $\hat{V}_p(y) = \left(\hat{V}_{\cdot(p+1)}(y), \ldots, \hat{V}_{\cdot n}(y) \right) \in \mathbb{R}^{k \times (n-p)}$ and the j-th column of $\hat{V}_0(y) \in \mathbb{R}^{kp \times (n-p)}$ equals $\left(\hat{V}'_{\cdot j+p-1}(y), \ldots, \hat{V}'_{\cdot j+1}(y), \hat{V}'_{\cdot j}(y) \right)' \in \mathbb{R}^{kp}$ for $j = 1, \ldots, n-p$. In matrix form we clearly have $\hat{Z}(y) = \hat{V}_p(y) - \hat{A}^{(p)}(y)\hat{V}_0(y)$.

2. Then, one computes the quantities

$$
\check{\Gamma}_i(y) = \begin{cases} \frac{1}{n-p}\sum_{j=i+1}^{n-p} \hat{Z}_{\cdot j}(y)\hat{Z}'_{\cdot(j-i)}(y) & \text{if } 0 \le i \le n-p-1, \\ \check{\Gamma}'_{-i}(y) & \text{if } 0 < -i \le n-p-1, \end{cases}
$$

and defines the preliminary estimate

$$
\check{\Psi}(y) = \sum_{i=-(n-p-1)}^{n-p-1} \kappa(i/M(y))\check{\Gamma}_i(y),
$$

where in case $M(y) = 0$ one sets $\kappa(i/M(y)) = 0$ for $i \ne 0$ and $\kappa(i/M(y)) = \kappa(0)$ for $i = 0$.

3. Finally, the preliminary estimate $\check{\Psi}(y)$ is 'recolored' using the transformation

$$
\hat{\Psi}(y) = \left(I_k - \sum_{l=1}^{p} \hat{A}_l^{(p)}(y)\right)^{-1} \check{\Psi}(y) \left[\left(I_k - \sum_{l=1}^{p} \hat{A}_l^{(p)}(y)\right)^{-1}\right]'.
$$

Remark 3.1. The construction of $\hat{\Psi}(y)$ outlined above clearly assumes that (i) $\hat{A}^{(p)}(y)$ is well defined, which is equivalent to $\text{rank}(\hat{V}_0(y)) = kp$; that (ii) $M(y)$ is well defined, which depends on the specific choice of M (cf. Section 3.1); and that (iii) $I_k - \sum_{i=1}^{p} \hat{A}_i^{(p)}(y)$ is invertible.

Remark 3.2. By assumption, all possible VAR orders p we consider must satisfy $p \le n/(k+1)$. This is done to rule out degenerate cases: for if $p > n/(k+1)$, then $\text{rank}(\hat{V}_0(y)) < kp$ would follow because of $\hat{V}_0(y) \in \mathbb{R}^{kp \times (n-p)}$. Hence the covariance estimator would nowhere be well defined for such a choice, because (i) in Remark 3.1 would then clearly be violated at every observation y.

Remark 3.3. In the present paper we focus on VAR prewhitening based on the OLS estimator. This is in line with the original suggestions by Newey and West (1994), as well as with Rho and Shao (2013). Alternatively, for $p = 1$, Andrews and Monahan (1992) suggested to use an eigenvalue adjusted version of the OLS estimator, the adjustment being applied if the matrix $I_k - \hat{A}_1^{(1)}(y)$ is close to being singular. We shall focus on the unadjusted OLS estimator for the following reasons: Newey and West (1994) reported that the finite sample properties show little sensitivity to this eigenvalue adjustment. Furthermore, it is the unadjusted estimator that is often used in implementations of the method suggested by Andrews and Monahan (1992) in software packages for statistical and econometric computing (e.g., its implementation in the R (R Core Team (2016)) package **sandwich** by Zeileis (2004), or its implementation in EViews, e.g., Schwert (2009), p. 784.). We remark, however, that one can obtain a negative result similar to Theorem 4.2, and a positive result concerning an adjustment procedure similar to Theorem 5.4, also for tests based on prewhitened estimators with eigenvalue adjustment. Furthermore, we conjecture that it is

possible to prove (similar to Proposition 4.5) the genericity of such a negative result, and to show that one can (similar to Proposition 5.5) generically resolve this problem by using the adjustment procedure. We leave the question of which estimator to choose for prewhitening to future research.

In a typical asymptotic analysis of tests based on prewhitened covariance estimators the event $N^*(\hat{\Omega})$ is asymptotically negligible (since $\hat{\Omega}$ converges to a positive definite, or almost everywhere positive definite matrix). Hence there is no need to be specific about the definition of the test statistic for $y \in N^*(\hat{\Omega})$, and one can work directly with the statistic

$$y \mapsto (R\hat{\beta}(y) - r)'\hat{\Omega}^{-1}(y)(R\hat{\beta}(y) - r), \tag{6}$$

which is left undefined for $y \in N^*(\hat{\Omega})$. In a finite sample setup, however, one has to think about the definition of the test statistic also for $y \in N^*(\hat{\Omega})$. Our decision to assign the value 0 to the test statistic for $y \in N^*(\hat{\Omega})$ is of course completely arbitrary. That this assignment does not affect our results at all is discussed in detail in the following remark.

Remark 3.4. Given that the estimator $\hat{\Omega}$ is based on a triple κ, M, p that satisfies Assumption 1 introduced below (which is assumed in all of our main results, and which is satisfied for covariance estimators using auxiliary AR(1) models for the construction of the bandwidth parameter as considered in Andrews and Monahan (1992), for covariance estimators as considered in Newey and West (1994), and for covariance estimators as considered in Rho and Shao (2013)), it follows from Lemma 3.10 that $N^*(\hat{\Omega})$ is either a $\lambda_{\mathbb{R}^n}$-null set, or that it coincides with $\mathbb{R}^n$. In the first case, which is generic under weak dimensionality constraints as shown in Lemma 3.11, the definition of the test statistic on $N^*(\hat{\Omega})$ does hence not influence the rejection probabilities, because our model is dominated by $\lambda_{\mathbb{R}^n}$ ($\mathfrak{C}$ contains only positive definite matrices). Therefore, size and power properties are not affected by the definition of the test statistic for $y \in N^*(\hat{\Omega})$. In the second case, i.e., if $N^*(\hat{\Omega})$ coincides with $\mathbb{R}^n$, the statistic in (6) is nowhere well defined, and hence, regardless of which value is assigned to it for observations $y \in N^*(\hat{\Omega})$, the resulting test statistic is constant, and thus any test based on it breaks down trivially.

3.1. Bandwidth parameters

In the following we describe bandwidth parameters M that are typically used in Step 2 in the construction of the prewhitened estimator $\hat{\Psi}$ as discussed above: The parametric approach (based on auxiliary AR(1) models) suggested by Andrews (1991) and Andrews and Monahan (1992), the nonparametric approach introduced by Newey and West (1994), and a data-independent approach which was already investigated in Kiefer and Vogelsang (2005) in simulation studies and which has recently been theoretically investigated by Rho and Shao (2013). Since the bandwidth parameter M is computed in Step 2 in the construction of $\hat{\Psi}(y)$, we assume that κ, p and y are given and that Step 1 has already been

successfully completed, i.e., all operations in Step 1 are well defined at y, in particular $\hat{Z}(y)$ is available for the construction of M. If not, we leave the bandwidth parameter (and hence the covariance estimator) undefined at y. We also implicitly assume that the quantities and operations appearing in the procedures outlined subsequently are well defined and leave the bandwidth parameter (and hence the covariance estimator) undefined else. A detailed structural analysis of the subset of the sample space where a prewhitened estimator $\hat{\Omega}$ is well defined is then later given in Lemma 3.9 in Section 3.3. Finally, we emphasize that the bandwidth parameters discussed subsequently all require the choice of additional tuning parameters. These tuning parameters are typically chosen independently of y and X, an assumption we shall maintain throughout the whole paper (but see Remark 3.8 for some generalizations).

3.1.1. The parametric approach of Andrews and Monahan (1992)

Let $\omega \in \mathbb{R}^k$ be such that $\omega \neq 0$ and $\omega_i \geq 0$ for $i = 1, \ldots, k$, i.e., ω is a *weights vector*. Based on this weights vector the bandwidth parameter is now obtained as follows: First, univariate AR(1) models are fitted via OLS to $\hat{Z}_{i\cdot}(y)$ for $i = 1, \ldots, k$, giving

$$\hat{\rho}_i(y) = \sum_{j=2}^{n-p} \hat{Z}_{ij}(y)\hat{Z}_{i(j-1)}(y) \bigg/ \sum_{j=1}^{n-p-1} \hat{Z}_{ij}(y)^2 \qquad \text{for } i = 1, \ldots, k,$$

$$\hat{\sigma}_i^2(y) = (n-p-1)^{-1} \sum_{j=2}^{n-p} \left(\hat{Z}_{ij}(y) - \hat{\rho}_i(y)\hat{Z}_{i(j-1)}(y) \right)^2 \qquad \text{for } i = 1, \ldots, k,$$

where we note that $n - p - 1 > 0$ holds as a consequence of $n > 2$ and $1 \leq p \leq \frac{n}{k+1}$. Then, one calculates

$$\hat{\alpha}_1(y) = \sum_{i=1}^{k} \omega_i \frac{4\hat{\rho}_i^2(y)\hat{\sigma}_i^4(y)}{(1-\hat{\rho}_i(y))^6(1+\hat{\rho}_i(y))^2} \bigg/ \sum_{i=1}^{k} \omega_i \frac{\hat{\sigma}_i^4(y)}{(1-\hat{\rho}_i(y))^4},$$

$$\hat{\alpha}_2(y) = \sum_{i=1}^{k} \omega_i \frac{4\hat{\rho}_i(y)^2\hat{\sigma}_i^4(y)}{(1-\hat{\rho}_i(y))^8} \bigg/ \sum_{i=1}^{k} \omega_i \frac{\hat{\sigma}_i^4(y)}{(1-\hat{\rho}_i(y))^4}.$$

Finally, bandwidth parameters are obtained via

$$M_{AM,j,\omega,c}(y) = c_1 \left(\hat{\alpha}_j(y)n \right)^{c_2} \quad \text{for} \quad j = 1, 2,$$

where to obtain a bandwidth parameter, one has to fix the constants $c_1 > 0$, $c_2 > 0$ and j and where $c = (c_1, c_2)$. Typically the choice of these constants and the choice of j depends on certain characteristics of κ (for specific choices see Andrews (1991), Section 6, in particular p. 834). For example, if κ is the Bartlett kernel one uses $c_1 = 1.1447$, $c_2 = 1/3$ and $j = 1$, or if κ is the Quadratic-Spectral kernel one would use $c_1 = 1.13221$, $c_2 = 1/5$ and $j = 2$. Since we do

not need such a specific dependence to derive our theoretical results, we do not impose any further assumptions on these constants beyond being positive (and independent of y and X). We shall denote by $\mathbb{M}_{AM}$ the set of all bandwidth parameters that can be obtained as special cases of the method in the present section, by appropriately choosing - functionally independently of y and X - a weights vector ω, constants $c_1 > 0$, $c_2 > 0$ and a $j \in \{1, 2\}$.

Remark 3.5. Since n, k and q are fixed quantities, the tuning parameters ω, c_i for $i = 1, 2$ and j might also depend on them, although we do not signify this in our notation. A similar remark applies to the constants appearing in Section 3.1.2 and in Section 3.1.3. Although we do not provide any details, we furthermore remark that one can extend our analysis to bandwidth parameters as above, but based on estimators other than $\hat{\rho}_i$, e.g., all estimators satisfying Assumption 4 of Preinerstorfer and Pötscher (2016) such as the Yule-Walker estimator or variants of the OLS estimator.

3.1.2. The non-parametric approach of Newey and West (1994)

Let $\omega \in \mathbb{R}^k$ be as in Section 3.1.1 and let $w(i) \geq 0$ for $|i| = 0, \ldots, n-p-1$ be real numbers such that $w(0) = 1$. For example, Newey and West (1994) suggested to use rectangular weights, i.e.,

$$
w^*(i) = \begin{cases} 1 & \text{if } |i| \leq \lfloor 4(n/100)^{2/9} \rfloor, \\ 0 & \text{else,} \end{cases}
$$

where $\lfloor . \rfloor$ denotes the floor function. Define for every $|i| = 0, \ldots n - p - 1$

$$
\bar{\sigma}_i(y) = \omega' \check{\Gamma}_i(y)\omega = (n-p)^{-1} \sum_{j=|i|+1}^{n-p} \omega' \hat{Z}_{\cdot j}(y) \hat{Z}'_{\cdot(j-|i|)}(y)\omega.
$$

A bandwidth parameter is then obtained via

$$
M_{NW,\omega,w,\bar{c}}(y) = \bar{c}_2 \left(\left[\sum_{i=-(n-p-1)}^{n-p-1} |i|^{\bar{c}_1} w(i)\bar{\sigma}_i(y) \Big/ \sum_{i=-(n-p-1)}^{n-p-1} w(i)\bar{\sigma}_i(y) \right]^2 n \right)^{\bar{c}_3}
$$

where $\bar{c}_1$ is a positive integer, where $\bar{c}_2$ and $\bar{c}_3$ are positive real numbers and where $\bar{c} = (\bar{c}_1, \bar{c}_2, \bar{c}_3)$. These numbers are constants independent of y and X and have to be chosen by the user. The choice typically depends on the kernel (for the specific choices we refer the reader to Newey and West (1994), Section 3). As in the previous section, we do not impose any assumptions beyond positivity (and independence of y and X) on the constants. Furthermore, we shall denote by $\mathbb{M}_{NW}$ the set of all bandwidth parameters that can be obtained as special cases of the method in the present section, by appropriately choosing - functionally independently of y and X - a weights vector, numbers $w(i) \geq 0$ for $|i| = 0, \ldots, n - p - 1$, $\bar{c}_1$ a positive integer, $\bar{c}_2 > 0$ and $\bar{c}_3 > 0$.

Remark 3.6. (i) The method described here is the 'real-bandwidth' approach suggested in Newey and West (1994), as opposed to the 'integer-bandwidth' approach. In the latter approach one would use $1 + \lfloor M_{NW,\omega,w,\bar{c}}(y) \rfloor$ instead of $M_{NW,\omega,w,\bar{c}}(y)$. Both approaches are asymptotically equivalent (Newey and West (1994), Theorem 2) for most kernels (including the Bartlett kernel which is suggested in Newey and West (1994)). Therefore, they are equally plausible in terms of their theoretical foundation. For the sake of simplicity and comparability with the bandwidth parameter as suggested by Andrews and Monahan (1992), which is not an integer in general, we have chosen to focus on the 'real-bandwidth' approach.

(ii) Newey and West (1994), p. 637, in principle also allow for $\bar{c}_1 = 0$ ($q = 0$ in their notation) in the definition of their estimator. We do not allow for such a choice. However, note that $\bar{c}_1 = 0$ implies $M_{NW,\omega,w,\bar{c}}(y) \equiv \bar{c}_2 n^{\bar{c}_3}$. This is a data-independent bandwidth parameter. These parameters are separately treated in Section 3.1.3.

3.1.3. Data-independent bandwidth parameters

Kiefer and Vogelsang (2005) and Rho and Shao (2013) studied properties of tests based on prewhitened covariance estimators and data-independent bandwidth parameters. Here one sets $M \equiv b(n - p)$ where $b \in (0, 1]$ is functionally independent of y and X. For example, in Rho and Shao (2013) the choice $b = 1$ is studied. These approaches all lead to bandwidth parameters $M_{KV} > 0$, that are functionally independent of both X and y. We denote the set of such bandwidth parameters by $\mathbb{M}_{KV}$.

3.2. Assumptions on κ, M and p

Different combinations of kernels κ, bandwidth parameters M and VAR orders p obviously lead to different estimators. We indicate the dependence of the estimator on these quantities by writing $\hat{\Omega}_{\kappa,M,p}$. In the present paper we shall consider estimators $\hat{\Omega}_{\kappa,M,p}$ based on a triple κ, M, p which satisfies the following assumption:

Assumption 1. The triple κ, M, p satisfies:

1. $\kappa : \mathbb{R} \to \mathbb{R}$ is an even function and $\kappa(0) = 1$. Furthermore, κ is continuous, satisfies $\lim_{x \to \infty} \kappa(x) = 0$, and for every real number $s > 0$ and every positive integer J the $J \times J$ symmetric Toeplitz matrix with ij-th coordinate $\kappa((i - j)/s)$ is positive definite.
2. $M \in \mathbb{M}_{AM} \cup \mathbb{M}_{NW} \cup \mathbb{M}_{KV}$.
3. p is an integer satisfying $1 \leq p \leq n/(k + 1)$.

Remark 3.7. First, we remark that the positive definiteness assumption in Part 1 of Assumption 1 is natural in our context, because it guarantees that $\hat{\Omega}_{\kappa,M,p}$ is nonnegative definite whenever it is well defined. Furthermore, it allows us to

derive simple conditions for positive definiteness of $\hat{\Omega}_{\kappa,M,p}$ (cf. Lemma 3.10). It is well known that many kernels used in practice satisfy the positive definiteness assumption, e.g., the Bartlett, Parzen, and Quadratic-Spectral kernel. Secondly, we note that in principle Assumption 1 does not prohibit a combination of $M_{AM,1,\omega,c} \in \mathbb{M}_{AM}$ with a second order kernel, or the combination of $M_{AM,2,\omega,c} \in \mathbb{M}_{AM}$ with a first order kernel. It also allows for a combination of elements of $\mathbb{M}_{NW}$ with a prewhitening order $p > 1$ and for the combination of elements of $\mathbb{M}_{KV}$ with a kernel other than the Bartlett kernel. This goes well beyond the original suggestions in Andrews and Monahan (1992), Newey and West (1994) and Rho and Shao (2013), but we include these additional possibilities for convenience. We also remark that since we assume throughout that $n > k$, the set of VAR orders satisfying the third part of Assumption 1 always includes the order $p = 1$.

Remark 3.8 (*Tuning parameters depending on the design*)**.** The tuning parameters used in the construction of $M \in \mathbb{M}_{AM} \cup \mathbb{M}_{NW} \cup \mathbb{M}_{KV}$, e.g., the weights vector ω used in the construction of $M \in \mathbb{M}_{AM} \cup \mathbb{M}_{NW}$, are by definition functionally independent of y *and* X. Requiring that the tuning parameters are independent of X is not a restriction in all results of the present paper in which the design matrix X is *fixed* (i.e., Theorem 4.2, Proposition 5.2, and Theorem 5.4). To see this, suppose that a design matrix X as in (1) is given, that κ and p satisfy the first and third part of Assumption 1, respectively, and that M is constructed as in one of the Sections 3.1.1, 3.1.2, 3.1.3, but with a vector of tuning parameters $c^*(.)$, say, that is not constant on $\mathfrak{X}_0$. The triple κ, M, p hence does not satisfy Assumption 1. Let $\tilde{M}$ be the bandwidth parameter that is obtained from M by replacing the vector of tuning parameters $c^*(.)$ by $\tilde{c} \equiv c^*(X)$. Clearly, $\kappa, \tilde{M}, p$ satisfies Assumption 1, and the test statistics as in Equation (5) based on $\hat{\Omega}_{\kappa,\tilde{M},p}$ and $\hat{\Omega}_{\kappa,M,p}$, respectively, coincide for this specific X.

3.3. Structural properties of prewhitened covariance estimators

The study of finite sample properties of a test based on the statistic in Equation (5) with $\hat{\Omega} = \hat{\Omega}_{\kappa,M,p}$ requires a detailed understanding of definiteness properties of the covariance estimator $\hat{\Omega}_{\kappa,M,p}$, and of the structure of the set $N^*(\hat{\Omega}_{\kappa,M,p})$. Denoting the subset of the sample space $\mathbb{R}^n$ where $\hat{\Omega}_{\kappa,M,p}$ is not well defined by $N(\hat{\Omega}_{\kappa,M,p})$, we can write

$$N^*(\hat{\Omega}_{\kappa,M,p}) = N(\hat{\Omega}_{\kappa,M,p}) \cup \left\{ y \in \mathbb{R}^n \backslash N(\hat{\Omega}_{\kappa,M,p}) : \det(\hat{\Omega}_{\kappa,M,p}(y)) = 0 \right\}.$$

As a first step we study $N(\hat{\Omega}_{\kappa,M,p})$ in the subsequent lemma, where it is shown that $N(\hat{\Omega}_{\kappa,M,p})$ is algebraic. The lemma also characterizes the dependence of $N(\hat{\Omega}_{\kappa,M,p})$ on the design matrix, which will later be useful for obtaining our genericity results.

Lemma 3.9. *Assume that the triple κ, M, p satisfies Assumption 1. Then,*

$$N(\hat{\Omega}_{\kappa,M,p}) = \{y \in \mathbb{R}^n : g_{\kappa,M,p}(y,X) = 0\},$$

where $g_{\kappa,M,p} : \mathbb{R}^n \times \mathbb{R}^{n \times k} \to \mathbb{R}$ *is a multivariate polynomial (explicitly constructed in the proof). As a consequence* $N(\hat{\Omega}_{\kappa,M,p})$ *is an algebraic set. Furthermore,* $g_{\kappa,M,p}$ *does not depend on the hypothesis* (R, r).

The subsequent lemma discusses definiteness and regularity properties of $\hat{\Omega}_{\kappa,M,p}$ and shows that $N^*(\hat{\Omega}_{\kappa,M,p})$ is an algebraic subset of $\mathbb{R}^n$. Again the dependence of this algebraic set on the design is clarified. Given a prewhitening order p satisfying Part 3 of Assumption 1, we define for every $y \in \mathbb{R}^n$ such that $\hat{A}^{(p)}(y)$ is well defined and such that $I_k - \sum_{l=1}^{p} \hat{A}_l^{(p)}(y)$ is invertible the matrix

$$B_p(y) = R(X'X)^{-1} \left(I_k - \sum_{l=1}^{p} \hat{A}_l^{(p)}(y) \right)^{-1} \hat{Z}(y). \tag{7}$$

Lemma 3.10. *Assume that the triple* κ, M, p *satisfies Assumption 1. Then the following holds:*

1. $\hat{\Omega}_{\kappa,M,p}(y)$ *is nonnegative definite if and only if* $g_{\kappa,M,p}(y, X) \neq 0$.
2. $\hat{\Omega}_{\kappa,M,p}(y)$ *is singular if and only if* $g_{\kappa,M,p}(y, X) \neq 0$ *and* $\mathrm{rank}(B_p(y)) < q$.
3. $\hat{\Omega}_{\kappa,M,p}(y) = 0$ *if and only if* $g_{\kappa,M,p}(y, X) \neq 0$ *and* $B_p(y) = 0$.
4. $\hat{\Omega}_{\kappa,M,p}(y)$ *is positive definite if* $g_{\kappa,M,p}(y, X) \neq 0$ *and* $\mathrm{rank}(\hat{Z}(y)) = k$.
5. *We have*

$$N^*(\hat{\Omega}_{\kappa,M,p}) = \left\{ y \in \mathbb{R}^n : g_{\kappa,M,p}^*(y, X, R) = 0 \right\},$$

where $g_{\kappa,M,p}^* : \mathbb{R}^n \times \mathbb{R}^{n \times k} \times \mathbb{R}^{q \times k} \to \mathbb{R}$ *is a multivariate polynomial (explicitly constructed in the proof). As a consequence* $N^*(\hat{\Omega}_{\kappa,M,p})$ *is an algebraic set. Furthermore,* $g_{\kappa,M,p}^*$ *is independent of* r.

It is a well known fact that an algebraic subset of $\mathbb{R}^n$ is either a closed $\lambda_{\mathbb{R}^n}$-null set, or coincides with $\mathbb{R}^n$ (for a proof see, e.g., Okamoto (1973)). The latter case occurs if and only if a (multivariate) polynomial defining the algebraic set vanishes everywhere. Together with Part 5 of Lemma 3.10 this implies that $N^*(\hat{\Omega}_{\kappa,M,p})$ is either a closed $\lambda_{\mathbb{R}^n}$-null set, or coincides with $\mathbb{R}^n$, depending on whether $g_{\kappa,M,p}^*(., X, R) \not\equiv 0$ or $g_{\kappa,M,p}^*(., X, R) \equiv 0$ holds, respectively. In the latter case, every test based on the test statistic defined in Equation (5) with $\hat{\Omega} = \hat{\Omega}_{\kappa,M,p}$ trivially breaks down, because in this case the test statistic vanishes identically on $\mathbb{R}^n$. Obviously, studying size and power properties of tests based on this test statistic in a sample of size n is only interesting, if we can guarantee that $g_{\kappa,M,p}^*(., X, R) \not\equiv 0$ holds for a sufficiently large set of design matrices. That this is indeed the case is the content of the subsequent lemma. More precisely it is shown that $g_{\kappa,M,p}^*(., X, R) \not\equiv 0$ is generically satisfied whenever n exceeds a certain threshold. It is also shown that the threshold we give can not be substantially improved. The notion of genericity employed is further discussed in Remark 3.12 following the lemma.

Lemma 3.11. *Assume that the triple* κ, M, p *satisfies Assumption 1. Then the following holds:*

1. *If $n < k(p+1) + p$ and $q = k$, then*

$$g^*_{\kappa,M,p}(.,X,R) \equiv 0 \text{ for every } X \in \mathfrak{X}_0.$$

2. *If $k(p+1) + p + \mathbf{1}_{\mathbb{M}_{AM}}(M) \leq n$, then*

$$g^*_{\kappa,M,p}(.,X,R) \not\equiv 0 \text{ for } \lambda_{\mathbb{R}^{n \times k}}\text{-almost every } X \in \mathfrak{X}_0;$$

   if $k = 1$ we have in particular $g^*_{\kappa,M,p}(.,e_+,R) \not\equiv 0$.
3. *If $k \geq 2$ and $k(p+1) + p^* + \mathbf{1}_{\mathbb{M}_{AM}}(M) \leq n$, where $p^* = p + (p \bmod 2)$, then*

$$g^*_{\kappa,M,p}(.,(e_+,\tilde{X}),R) \not\equiv 0 \text{ for } \lambda_{\mathbb{R}^{n \times (k-1)}}\text{-almost every } \tilde{X} \in \tilde{\mathfrak{X}}_0.$$

Remark 3.12. (1) Part 1 demonstrates that if n is too small in the sense that $n < k(p+1) + p$, then for every $X \in \mathfrak{X}_0$ the test statistic in Equation (5) with $\hat{\Omega} = \hat{\Omega}_{\kappa,M,p}$ vanishes identically if $q = k$ holds, because the estimator $\hat{\Omega}_{\kappa,M,p}$ is either not well defined or singular at every observation y. This shows that one can in general not expect that $N^*(\hat{\Omega}_{\kappa,M,p})$ is generically a $\lambda_{\mathbb{R}^n}$-null set in case $n < k(p+1) + p$.

(2) Under the assumption that $k(p+1) + p + \mathbf{1}_{\mathbb{M}_{AM}}(M) \leq n$ holds, Part 2 establishes genericity of $g^*_{\kappa,M,p}(.,X,R) \not\equiv 0$ in that it shows that the statement holds for $\lambda_{\mathbb{R}^{n \times k}}$- almost every $X \in \mathfrak{X}_0$. This notion of genericity is obviously related to situations, where the data-generating process underlying the design matrix X is assumed to be absolutely continuous w.r.t. $\lambda_{\mathbb{R}^{n \times k}}$. In this situation, a bandwidth parameter $M \in \mathbb{M}_{AM} \cup \mathbb{M}_{NW}$ would typically be based on the weights vector $\omega = (1,\ldots,1)' \in \mathbb{R}^k$. As a specific result of independent interest it is also shown that if $k = 1$ then $g^*_{\kappa,M,p}(.,e_+,R) \not\equiv 0$ holds, which means that in the location model the set $N^*(\hat{\Omega}_{\kappa,M,p})$ is a $\lambda_{\mathbb{R}^n}$- null set.

(3) Under the assumption that $k \geq 2$ and $k(p+1) + p^* + \mathbf{1}_{\mathbb{M}_{AM}}(M) \leq n$ holds, Part 3 establishes genericity of $g^*_{\kappa,M,p}(.,(e_+,\tilde{X}),R) \not\equiv 0$ by showing that the statement holds for $\lambda_{\mathbb{R}^{n \times (k-1)}}$ almost every $\tilde{X} \in \tilde{\mathfrak{X}}_0$. This is a genericity statement concerning design matrices the first column of which is the intercept. In contrast to (2) this notion of genericity is related to situations, where the first column of the design matrix is fixed and the data-generating process underlying the remaining columns is absolutely continuous w.r.t. $\lambda_{\mathbb{R}^{n \times (k-1)}}$. In such a setup the construction of a bandwidth parameter $M \in \mathbb{M}_{AM} \cup \mathbb{M}_{NW}$ would typically be based on the weights vector $\omega = (0,1\ldots,1)' \in \mathbb{R}^k$.

4. A negative result and its generic applicability

In the first part of this section we obtain our main negative result concerning finite sample properties of tests based on prewhitened nonparametric covariance estimators. For this result to hold, we have to impose a richness assumption on the covariance model $\mathfrak{C}$. Let $\mathfrak{C}_{AR(1)}$ denote the set of all correlation matrices corresponding to stationary autoregressive processes of order one, i.e., $\mathfrak{C}_{AR(1)} = \{\Lambda(\rho) : \rho \in (-1,1)\}$, where $\Lambda(\rho)_{ij} = \rho^{|i-j|}$ for $1 \leq i,j \leq n$. The assumption is as follows.

Assumption 2. $\mathfrak{C}_{AR(1)} \subseteq \mathfrak{C}$.

Remark 4.1. Assumption 2 implies in particular that the singular boundary of $\mathfrak{C} \subseteq \mathbb{R}^{n \times n}$, i.e., the set of singular matrices in bd $\mathfrak{C}$, contains at least the two elements $e_+ e'_+$ and $e_- e'_-$, where $e_+ = (1, \ldots, 1)'$ and $e_- = (-1, 1, \ldots, (-1)^n)'$. We note that these two singular matrices can be approximated by sequences $\Lambda(\rho_m) \in \mathfrak{C}$ with $\rho_m \to 1$ and $\rho_m \to -1$, respectively, where $\rho_m \in (-1, 1)$.

Since the procedures we study in the present paper are geared towards situations such as $\mathfrak{C} \supseteq \mathfrak{C}_\xi$ for some $\xi \geq 0$ (cf. Remark 2.1), covariance models which clearly satisfy the above assumption, Assumption 2 is mild in our context (cf. also the discussion in Section 3.2.2 of Preinerstorfer and Pötscher (2016)). Under this assumption and given a hypothesis (R, r), the subsequent theorem provides four sufficient conditions on the design matrix under which a test based on a test statistic as in Equation (5) with $\hat{\Omega} = \hat{\Omega}_{\kappa,M,p}$, together with an arbitrary (but data-independent) critical value $0 < C < \infty$, breaks down in terms of its finite sample size and/or power properties. More precisely, Conditions (1) and (4) imply that the test has size equal to one, Condition (3) implies that the test has size not smaller than $1/2$, and Condition (2) implies that the nuisance-minimal rejection probability equals zero at every point $\mu_1 \in \mathfrak{M}_1$.

Theorem 4.2. *Suppose that the triple κ, M, p satisfies Assumption 1 and that $\mathfrak{C}$ satisfies Assumption 2. Let T be the test statistic defined in (5) with $\hat{\Omega} = \hat{\Omega}_{\kappa,M,p}$. Let $W(C) = \{y \in \mathbb{R}^n : T(y) \geq C\}$ be the rejection region, where C is a real number satisfying $0 < C < \infty$. Then the following holds:*

1. *Suppose $g^*_{\kappa,M,p}(e_+, X, R) \neq 0$ and $T(e_+ + \mu_0^*) > C$ holds for some (and hence all) $\mu_0^* \in \mathfrak{M}_0$, or $g^*_{\kappa,M,p}(e_-, X, R) \neq 0$ and $T(e_- + \mu_0^*) > C$ holds for some (and hence all) $\mu_0^* \in \mathfrak{M}_0$. Then*

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

   holds for every $\mu_0 \in \mathfrak{M}_0$ and every $0 < \sigma^2 < \infty$. In particular, the size of the test is equal to one.

2. *Suppose $g^*_{\kappa,M,p}(e_+, X, R) \neq 0$ and $T(e_+ + \mu_0^*) < C$ holds for some (and hence all) $\mu_0^* \in \mathfrak{M}_0$, or $g^*_{\kappa,M,p}(e_-, X, R) \neq 0$ and $T(e_- + \mu_0^*) < C$ holds for some (and hence all) $\mu_0^* \in \mathfrak{M}_0$. Then*

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

   holds for every $\mu_0 \in \mathfrak{M}_0$ and every $0 < \sigma^2 < \infty$, and hence

$$\inf_{\mu_1 \in \mathfrak{M}_1} \inf_{\Sigma \in \mathfrak{C}} P_{\mu_1, \sigma^2 \Sigma} (W(C)) = 0$$

   holds for every $0 < \sigma^2 < \infty$. In particular, the test is biased. Furthermore, the nuisance-infimal rejection probability at every point $\mu_1 \in \mathfrak{M}_1$ is zero, i.e.,

$$\inf_{0 < \sigma^2 < \infty} \inf_{\Sigma \in \mathfrak{C}} P_{\mu_1, \sigma^2 \Sigma} (W(C)) = 0.$$

   In particular, the infimal power of the test is equal to zero.

3. *Suppose $g^*_{\kappa,M,p}(e_+, X, R) \neq 0$, $T(e_+ + \mu^*_0) = C$ and $\operatorname{grad} T(e_+ + \mu^*_0)$ exists for some (and hence all) $\mu^*_0 \in \mathfrak{M}_0$, or $g^*_{\kappa,M,p}(e_-, X, R) \neq 0$, $T(e_- + \mu^*_0) = C$ and $\operatorname{grad} T(e_- + \mu^*_0)$ exists for some (and hence all) $\mu^*_0 \in \mathfrak{M}_0$. Then*

$$\sup_{\Sigma \in \mathfrak{C}} P_{\mu_0, \sigma^2 \Sigma}\left(W\left(C\right)\right) \geq 1/2$$

holds for every $\mu_0 \in \mathfrak{M}_0$ and every $0 < \sigma^2 < \infty$. In particular, the size of the test is at least $1/2$.

4. *Suppose that $g^*_{\kappa,M,p}(., X, R) \not\equiv 0$. Suppose further that $e_+ \in \mathfrak{M}$ and $R\hat{\beta}(e_+) \neq 0$ holds, or $e_- \in \mathfrak{M}$ and $R\hat{\beta}(e_-) \neq 0$ holds. Then*

$$\sup_{\Sigma \in \mathfrak{C}} P_{\mu_0, \sigma^2 \Sigma}\left(W\left(C\right)\right) = 1$$

holds for every $\mu_0 \in \mathfrak{M}_0$ and every $0 < \sigma^2 < \infty$. In particular, the size of the test is equal to one.

Remark 4.3. (i) Lemma C.1 in Appendix C shows that the rejection probabilities $P_{\mu, \sigma^2 \Sigma}(W(C))$ depend on (μ, σ^2, Σ) only through $(\langle(R\beta - r)/\sigma\rangle, \Sigma)$, where β is uniquely determined by $X\beta = \mu$.

(ii) Obviously, the conclusions of the preceding theorem also apply to any rejection region $W^* \in \mathcal{B}(\mathbb{R}^n)$ which differs from $W(C)$ only by a $\lambda_{\mathbb{R}^n}$-null set.

(iii) In Part 1 of the theorem the condition $g^*_{\kappa,M,p}(e_+, X, R) \neq 0$ ($g^*_{\kappa,M,p}(e_-, X, R) \neq 0$) is superfluous, because it is already implicit in $T(e_+ + \mu^*_0) > C > 0$ ($T(e_- + \mu^*_0) > C > 0$), which is readily seen from the definition of T in Equation (5). A similar comment applies to Part 3 of the theorem, where the condition $g^*_{\kappa,M,p}(e_+, X, R) \neq 0$ ($g^*_{\kappa,M,p}(e_-, X, R) \neq 0$) is already implicit in $T(e_+ + \mu^*_0) = C > 0$ ($T(e_- + \mu^*_0) = C > 0$). The conditions are included for the sake of comparability with Part 2 of the theorem.

(iv) In case $M \in \mathbb{M}_{KV}$, the assumption concerning the existence of the gradient can be dropped in Part 3 of the theorem. This follows from Lemma C.2 in Appendix C, where it is shown that if $M \in \mathbb{M}_{KV}$, then the existence of $\operatorname{grad} T(e_+ + \mu^*_0)$ and $\operatorname{grad} T(e_- + \mu^*_0)$ is already implied by $g^*_{\kappa,M,p}(e_+, X, R) \neq 0$ and $g^*_{\kappa,M,p}(e_-, X, R) \neq 0$, respectively.

(v) Throughout the theorem, Assumption 2 can be replaced by the weaker assumption that there exist two sequences $\Lambda(\rho^{(1)}_m)$ and $\Lambda(\rho^{(2)}_m)$ of AR(1) correlation matrices in $\mathfrak{C}$, such that $\rho^{(1)}_m \to -1$ and $\rho^{(2)}_m \to 1$. In Parts 1 and 2 of the theorem it is even enough to assume that there exist sequences $\Sigma^{(i)}_m \in \mathfrak{C}$ for $i = 1, 2$ with $\Sigma^{(1)}_m \to e_+ e'_+$ and $\Sigma^{(2)}_m \to e_- e'_-$. Therefore, in these parts it is only important that - and not how - these singular matrices can be approximated from within $\mathfrak{C}$.

We shall now provide some intuition for Theorem 4.2 (cf. also the discussion preceding Theorem 5.7 in Preinerstorfer and Pötscher (2016)). The repeated appearance of the vectors e_+ and e_- in the theorem stems from the fact that both $e_+ e'_+$ and $e_- e'_-$ are elements of the singular boundary of $\mathfrak{C} \supseteq \mathfrak{C}_{AR(1)}$ (cf. Remark 4.1). Furthermore, for every $\mu^*_0 \in \mathfrak{M}_0$ and every $0 < \sigma^2 < \infty$ we have that $P_{\mu^*_0, \sigma^2 \Sigma} \to P_{\mu^*_0, \sigma^2 e_+ e'_+}$ weakly as $\Sigma \to e_+ e'_+$ with $\Sigma \in \mathfrak{C}$, and similarly that

$P_{\mu_0^*,\sigma^2\Sigma} \to P_{\mu_0^*,\sigma^2 e_- e_-'}$ weakly as $\Sigma \to e_- e_-'$ with $\Sigma \in \mathfrak{C}$. These limiting measures are absolutely continuous w.r.t. $\lambda_{\mu_0^*+\mathrm{span}(e_+)}$ and $\lambda_{\mu_0^*+\mathrm{span}(e_-)}$, respectively. As a consequence we see that the mass of $P_{\mu_0^*,\sigma^2\Sigma}$ concentrates on 'neighborhoods' of certain one-dimensional affine spaces as Σ approximates $e_+ e_+'$ or $e_- e_-'$ from within $\mathfrak{C}$. From that it is intuitively clear that size and power properties crucially depend on the behavior of the tests on 'neighborhoods' of these spaces. The first and second part of the theorem provide sufficient conditions under which these spaces are almost surely (w.r.t. $\lambda_{\mu_0^*+\mathrm{span}(e_+)}$ and $\lambda_{\mu_0^*+\mathrm{span}(e_-)}$) contained in the interior or exterior of the rejection region, respectively. The former case then leads to size distortions, the latter to power deficiencies. The situation in the third part of the theorem is quite different and more complex. In this case the one-dimensional affine space supporting the respective limiting measure is neither almost surely contained in the interior, nor almost surely contained in the exterior of the rejection region. Rather it is almost surely contained in the boundary of the rejection region. Therefore, in contrast to Parts 1 and 2, it is not only important that the measures concentrate on the respective one-dimensional space, but also *how* they concentrate (cf. Remark 4.3 (v)). The concentration turns out to be such that eventually the measures put roughly equal weight onto the rejection region and onto its complement, resulting in rejection probabilities as large as $1/2$ under the null. We point out that the proof idea used to establish Part 3 is inspired by the proof of Theorem 2.18 in Preinerstorfer and Pötscher (2017). The last part of the theorem considers the case where one of the vectors e_+ or e_- is an element of $\mathfrak{M}$ that is also 'involved' in the hypothesis. It is then shown that the size of the test is one if the global condition $g_{\kappa,M,p}^*(.,X,R) \not\equiv 0$ is satisfied. We recall that if this condition fails to hold, then the test T based on $\hat{\Omega}_{\kappa,M,p}$ breaks down in a trivial way, because T is then zero everywhere. Therefore we see that under Assumption 2 one simply can not test a hypothesis involving $e_+ \in \mathfrak{M}$ or $e_- \in \mathfrak{M}$ by means of a test T based on $\hat{\Omega}_{\kappa,M,p}$ with κ, M, p satisfying Assumption 1 (this in particular covers the location model where $X = e_+$, cf. also Lemma 3.11, Part 2).

Remark 4.4. Suppose that it is *known a priori* that for some (fixed) $\varepsilon \in (0,1]$ the covariance model $\mathfrak{C}$ does not contain AR(1) correlation matrices $\Lambda(\rho)$ with $\rho \leq -1 + \varepsilon$; i.e., instead of Assumption 2 the covariance model $\mathfrak{C}$ satisfies

$$\mathfrak{C}_{AR(1)}(\varepsilon) = \{\Lambda(\rho) : \rho \in (-1 + \varepsilon, 1)\} \subseteq \mathfrak{C}.$$

Inspection of the proof of Theorem 4.2 then shows that a version of Theorem 4.2 holds, in which all references to e_- are deleted in Parts 1-4. For example, Part 4 of this version of Theorem 4.2 reads as follows:

"Suppose that $g_{\kappa,M,p}^*(.,X,R) \not\equiv 0$. Suppose further that $e_+ \in \mathfrak{M}$ and $R\hat{\beta}(e_+) \neq 0$ holds. Then

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

holds for every $\mu_0 \in \mathfrak{M}_0$ and every $0 < \sigma^2 < \infty$. In particular, the size of the test is equal to one."

This statement covers (in particular) the important special case of testing a restriction on the mean in a location model. We make the following observations concerning this version of Theorem 4.2:

- Since $e_- e'_-$ is not necessarily an element of the singular boundary of the covariance model considered here, the result just described does not contain "size equal to one"- or "nuisance-minimal-power equal to zero"-statements that arise from covariance matrices approaching $e_- e'_-$. Note, however, that the original Theorem 4.2 implies by a continuity argument that if ε is small (compared to sample size), then considerable size distortions or power deficiencies will nevertheless be present for covariance matrices in $\mathfrak{C}$ that are close to $e_- e'_-$.
- Consider the case where $e_+ \in \mathfrak{M}$, i.e., the regression contains an intercept, and where the hypothesis does not involve the intercept, i.e., $R\hat{\beta}(e_+) = 0$: Then we see that Parts 1-4 of the version of Theorem 4.2 just obtained do not apply. In fact, in this case we can establish a positive result concerning a test based on T with $\hat{\Omega} = \hat{\Omega}_{\kappa,M,p}$, and based on a non-standard critical value that depends on ε. This positive result, together with its restrictions, is discussed in Remark 5.3.

Given a hypothesis (R, r) the four sufficient conditions provided in the preceding theorem are conditions on the design matrix X. They depend on observable quantities only. How these conditions can be checked is discussed in the subsequent paragraph: The first three parts of the theorem operate under the local assumption that the multivariate polynomial $g^*_{\kappa,M,p}(., X, R)$ does not vanish at the point e_+ or e_-, respectively. The multivariate polynomial $g^*_{\kappa,M,p}(., X, R)$ is explicitly constructed in the proof of Lemma 3.10. Therefore, the condition that it does not vanish at specific data points can readily be checked. Some additional conditions needed in Parts 1-3 of the theorem are formulated in terms of $T(e_+ + \mu_0^*)$ and $T(e_- + \mu_0^*)$, which are in fact independent of the specific $\mu_0^* \in \mathfrak{M}_0$ chosen and therefore easy to calculate. Part 3 of the theorem requires the existence of $\operatorname{grad} T(e_+ + \mu_0^*)$ or $\operatorname{grad} T(e_- + \mu_0^*)$ (which is immaterial if $M \in \mathbb{M}_{KV}$ as discussed in the preceding Remark). Again the existence of the gradients is independent of the specific choice of $\mu_0^* \in \mathfrak{M}_0$. Sufficient conditions for the existence of the gradient, under the assumption that κ is continuously differentiable on the complement of a finite number of points, are provided in Lemma C.2 in Appendix C. These conditions amount to checking whether or not $M(e_+)$ or $M(e_-)$, respectively, is an element of a certain set determined by κ consisting of finitely many points. In contrast to Parts 1-3, the fourth part of the theorem operates under the global assumption that the multivariate polynomial $g^*_{\kappa,M,p}(., X, R)$ is not the zero polynomial. Since the polynomial $g^*_{\kappa,M,p}(., X, R)$ is explicitly constructed in the proof of Lemma 3.10, the global assumption $g^*_{\kappa,M,p}(., X, R) \not\equiv 0$ can either be checked analytically, or by using standard algorithms for polynomial identity testing. In addition to this global assumption, the fourth part needs additional assumptions on the structure of $\mathfrak{M}$ and the hypothesis (R, r) which can of course be easily checked by the user.

The preceding theorem has given sufficient conditions on the design matrix, under which the test considered breaks down in terms of its size and/or power behavior. However, for a given hypothesis (R, r) there exist elements of $\mathfrak{X}_0 \subseteq \mathbb{R}^{n \times k}$ to which the theorem is not applicable. As a consequence, the question remains to 'how many' elements of $\mathfrak{X}_0$ the theorem can be applied once (R, r) has been fixed. This question is studied subsequently. It is shown that generically in the space of all design matrices at least one of the four conditions of Theorem 4.2 applies. The first part of the proposition establishes this genericity result in the class of all design matrices of full column rank, i.e., $\mathfrak{X}_0$. The remaining parts establish the genericity result in case $k \geq 2$ and the first column of X is the intercept, i.e., $X = (e_+, \tilde{X})$ with $\tilde{X} \in \tilde{\mathfrak{X}}_0$. Before we state the proposition, we introduce two assumptions on the kernel κ. The first assumption is satisfied by all kernels typically used in practice.

Assumption 3. The kernel κ is continuously differentiable on the complement of $\Delta(\kappa) \subseteq \mathbb{R}$, a set consisting of finitely many elements.

The second assumption, which is used in some statements of the second part of the genericity result, imposes compactness of the support of the kernel. This is satisfied by many kernels used in practice, e.g., the Bartlett kernel or the Parzen kernel, but is not satisfied by the Quadratic-Spectral kernel.

Assumption 4. The support of κ is compact.

The genericity result is now as follows, where several quantities are equipped with the additional subindex X to stress their dependence on the design matrix.

Proposition 4.5. *Fix a hypothesis (R, r) such that* $\operatorname{rank}(R) = q$. *Let κ, M, p satisfy Assumption 1. For $X \in \mathfrak{X}_0$ let T_X be the test statistic defined in (5) with $\hat{\Omega} = \hat{\Omega}_{\kappa, M, p, X}$ and let $\mu_{0,X}^* \in \mathfrak{M}_{0,X} = \{\mu \in \operatorname{span}(X) : \mu = X\beta, R\beta = r\}$ be arbitrary (the sets defined below do not depend on the choice of $\mu_{0,X}^*$). Fix a critical value C such that $0 < C < \infty$. Then, the following holds.*

1. *Suppose that $k(p+1) + p + \mathbf{1}_{\mathbb{M}_{AM}}(M) \leq n$, define*

$$\mathfrak{X}_1(e_+) = \left\{ X \in \mathfrak{X}_0 : g_{\kappa, M, p}^*(e_+, X, R) = 0 \right\}$$
$$\mathfrak{X}_2(e_+) = \left\{ \begin{matrix} X \in \mathfrak{X}_0 \backslash \mathfrak{X}_1(e_+) : & \nexists \, (\operatorname{grad} T_X(.))|_{e_+ + \mu_{0,X}^*} \\ & \text{and } T_X(e_+ + \mu_{0,X}^*) = C \end{matrix} \right\},$$

   and similarly define $\mathfrak{X}_1(e_-)$ and $\mathfrak{X}_2(e_-)$. Then, $\mathfrak{X}_1(e_+)$ and $\mathfrak{X}_1(e_-)$ are $\lambda_{\mathbb{R}^{n \times k}}$-null sets. If $M \in \mathbb{M}_{KV}$ or if κ satisfies Assumption 3, then $\mathfrak{X}_2(e_+)$ and $\mathfrak{X}_2(e_-)$ are $\lambda_{\mathbb{R}^{n \times k}}$-null sets. If Assumption 2 holds, then the set of all design matrices $X \in \mathfrak{X}_0$ for which the first three parts of Theorem 4.2 do not apply is a subset of $(\mathfrak{X}_1(e_+) \cup \mathfrak{X}_2(e_+)) \cap (\mathfrak{X}_1(e_-) \cup \mathfrak{X}_2(e_-))$ and hence is a $\lambda_{\mathbb{R}^{n \times k}}$-null set if $M \in \mathbb{M}_{KV}$ or if κ satisfies Assumption 3; it thus is a 'negligible' subset of $\mathfrak{X}_0$ in view of the fact that $\mathfrak{X}_0$ differs from $\mathbb{R}^{n \times k}$ only by a $\lambda_{\mathbb{R}^{n \times k}}$-null set.

2. *Let $k \geq 2$ and assume further that $k(p+1) + p^* + \mathbf{1}_{\mathbb{M}_{AM}}(M) \leq n$, where $p^* = p + (p \bmod 2)$. Define*

$$\tilde{\mathfrak{X}}_1\left(e_-\right) = \left\{\tilde{X} \in \tilde{\mathfrak{X}}_0 : g^*_{\kappa,M,p}(e_-, (e_+, \tilde{X}), R) = 0\right\},$$

$$\tilde{\mathfrak{X}}_2\left(e_-\right) = \left\{\begin{array}{ll} \tilde{X} \in \tilde{\mathfrak{X}}_0 \backslash \tilde{\mathfrak{X}}_1\left(e_-\right) : & \nexists \left(\operatorname{grad} T_{(e_+,\tilde{X})}(.)\right)|_{e_- + \mu^*_{0,(e_+,\tilde{X})}} \\ & and\ T_{(e_+,\tilde{X})}(e_- + \mu^*_{0,(e_+,\tilde{X})}) = C \end{array}\right\}.$$

Then, $\tilde{\mathfrak{X}}_1\left(e_-\right)$ is a $\lambda_{\mathbb{R}^{n \times (k-1)}}$-null set. Furthermore, $\tilde{\mathfrak{X}}_2\left(e_-\right)$ is a $\lambda_{\mathbb{R}^{n \times (k-1)}}$-null set under each of the following conditions:

(a) $M \in \mathbb{M}_{KV}$.

(b) $M \in \mathbb{M}_{AM}$ and κ satisfies Assumptions 3 and 4.

(c) $M \in \mathbb{M}_{NW}$, p is odd, $\omega_i > 0$ for some $i > 1$ and κ satisfies Assumptions 3 and 4.

(d) κ satisfies Assumption 3 and $\tilde{X} \mapsto T_{(e_+,\tilde{X})}(e_- + \mu^*_{0,(e_+,\tilde{X})}) \not\equiv C$ on $\tilde{\mathfrak{X}}_0 \backslash \tilde{\mathfrak{X}}_1\left(e_-\right)$.

Suppose that the first column of R consists of zeros and that Assumption 2 holds. Then, the set of all matrices $\tilde{X} \in \tilde{\mathfrak{X}}_0$ such that the first three parts of Theorem 4.2 do not apply to the design matrix $X = (e_+, \tilde{X})$ is a subset of $\tilde{\mathfrak{X}}_1\left(e_-\right) \cup \tilde{\mathfrak{X}}_2\left(e_-\right)$ and hence is a $\lambda_{\mathbb{R}^{n \times (k-1)}}$-null set if one of the conditions in (a)-(d) holds; it thus is a 'negligible' subset of $\tilde{\mathfrak{X}}_0$ in view of the fact that $\tilde{\mathfrak{X}}_0$ differs from $\mathbb{R}^{n \times (k-1)}$ only by a $\lambda_{\mathbb{R}^{n \times (k-1)}}$-null set.

3. *Suppose $k \geq 2$, that the first column of R is nonzero and that Assumption 2 holds. Then Theorem 4.2 (Part 4) applies to the design matrix $X = \left(e_+, \tilde{X}\right)$ for every $\tilde{X} \in \tilde{\mathfrak{X}}_0$ satisfying $g^*_{\kappa,M,p}(., X, R) \not\equiv 0$.*

Remark 4.6. (i) If $n < k(p+1)+p$ and $q = k$ holds, the first part of Lemma 3.11 shows that the test trivially breaks down, since for every element X of $\mathfrak{X}_0$ the test statistic T_X is then constant on $\mathbb{R}^n$. Therefore, the assumption on n in the first two parts of the proposition can in general not be substantially improved.

(ii) In the second part of the proposition, the analogously defined sets $\tilde{\mathfrak{X}}_1\left(e_+\right)$ and $\tilde{\mathfrak{X}}_2\left(e_+\right)$ clearly satisfy $\tilde{\mathfrak{X}}_1\left(e_+\right) = \tilde{\mathfrak{X}}_0$ and $\tilde{\mathfrak{X}}_2\left(e_+\right) = \emptyset$.

(iii) In the third part of the proposition, if $X = (e_+, \tilde{X})$ does not satisfy $g^*_{\kappa,M,p}(., X, R) \not\equiv 0$, then the test breaks down in a trivial way, since T_X is then constant.

The first part of the preceding genericity result shows that if $M \in \mathbb{M}_{KV}$, or if the kernel satisfies Assumption 3, then Theorem 4.2 can be applied to generic elements of $\mathfrak{X}_0$, i.e., to all elements besides a $\lambda_{\mathbb{R}^{n \times k}}$-null set. Since all kernels used in practice, in particular the kernels emphasized in Andrews and Monahan (1992) and Newey and West (1994), i.e., the Quadratic-Spectral kernel and the Bartlett kernel, respectively, satisfy Assumption 3, this additional restriction on κ is practically immaterial. The second part of the proposition considers the situation where the first column of the design matrix is the intercept, which in addition is assumed not to be involved in the hypothesis in the sense that the first column of R is zero. In this situation it is shown that Theorem 4.2 can generically be applied to design matrices of the form $(e_+, \tilde{X})$ with $\tilde{X} \in \tilde{\mathfrak{X}}_0$, under certain sets of conditions on the triple κ, M, p. We first discuss Conditions (a)-(c):

(a) In case $M \in \mathbb{M}_{KV}$ no additional condition is needed for establishing generic applicability of Theorem 4.2.

(b) If $M \in \mathbb{M}_{AM}$, generic applicability of the negative result follows if the kernel satisfies Assumptions 3 and 4, which applies to many kernels used in practice, but not to the Quadratic-Spectral kernel which is emphasized in Andrews and Monahan (1992).

(c) In case $M \in \mathbb{M}_{NW}$, the result shows that the procedure breaks down generically if p is odd, $\omega_i > 0$ for some $i > 1$ and κ satisfies Assumptions 3 and 4. This seems to be restrictive. However, the recommended procedure in Newey and West (1994) is obtained by choosing κ the Bartlett kernel, $p = 1$ and $\omega = (0, 1, \ldots, 1)'$, because in Part 2 the first column of X is the intercept. Therefore, we see that the recommended procedure in Newey and West (1994) satisfies this condition.

Summarizing, we see that the Conditions (a)-(c) in the proposition cover the recommended choices of κ, M and p in Newey and West (1994) and Rho and Shao (2013). For all procedures that are not covered by Conditions (a)-(c), e.g., the procedure in Andrews and Monahan (1992) based on the Quadratic-Spectral kernel, one can typically obtain the genericity result by applying Condition (d), which (under Assumption 3) is always satisfied apart from at most one exceptional critical value C^*. This is seen as follows: Clearly, Condition (d) depends on the critical value C. We see that if Assumption 3 is satisfied, then (d) can be violated for at most a single $0 < C^* < \infty$. If this C^* happens to coincide with C, the condition is not satisfied and we can not draw the desired conclusion for this specific value of C. Moreover, we immediately see that the condition must then be satisfied for any other choice C', say. Therefore, generic applicability of the negative result follows for any value $C' \neq C$ in that case. This shows that even if one chooses a triple κ, M, p that does not allow for an application of (a)-(c), one can not expect to obtain a procedure that has good finite sample size and power properties, because for all but at most one exceptional critical value the corresponding test is guaranteed to break down generically. The third part of the proposition considers the case where the first column of the design matrix is the intercept, and where the coefficient corresponding to the intercept is restricted by the hypothesis. In this case it follows that one can either apply Part 4 of Theorem 4.2, or the test statistic is constant and hence the test breaks down in a trivial way (cf. Remark 4.6).

5. A positive result, an adjustment procedure and its generic applicability

In the previous section we have established a (generically applicable) negative result concerning tests based on (5) and a prewhitened covariance estimator $\hat{\Omega}_{\kappa,M,p}$. In the present section we first present a positive result concerning these tests under a *non-generic* condition on the design matrix. Then we introduce an adjustment procedure and establish a condition on the design matrix under which the adjustment procedure leads to improved tests. Finally we prove

that this condition holds generically in the set of all design matrices. Both, the positive result concerning tests as in (5) based on a prewhitened covariance estimator $\hat{\Omega}_{\kappa,M,p}$, and the results concerning the adjustment procedure are established under the following assumption on the covariance model $\mathfrak{C}$.

Assumption 5. The set $\mathfrak{C} \subseteq \mathbb{R}^{n \times n}$ is norm-bounded and satisfies $\mathfrak{C}_{AR(1)} \subseteq \mathfrak{C}$. Furthermore, for every sequence $\Sigma_m \in \mathfrak{C}$ that converges to $\bar{\Sigma} \in \mathrm{bd}(\mathfrak{C})$ satisfying $\mathrm{rank}(\bar{\Sigma}) < n$ there exists a corresponding sequence $\rho_m \in (-1, 1)$ such that $\Lambda(\rho_m)^{-1/2} \Sigma_m \Lambda(\rho_m)^{-1/2} \to I_n$ as $m \to \infty$.

Remark 5.1. (i) We first note that Assumption 5 is stronger than Assumption 2. Therefore, under the former assumption the negative result established in Section 4 concerning tests as in (5) based on a prewhitened covariance estimator $\hat{\Omega}_{\kappa,M,p}$ *does* apply a fortiori. As a consequence, if $\mathfrak{C}$ satisfies Assumption 5, then positive results concerning size and power properties of tests of the form (5) can only be established under non-generic assumptions on the design matrix. However, as we shall show, positive results can generically be established for an *adjusted* version of such tests.

(ii) Boundedness of $\mathfrak{C}$ is typically satisfied in our setup, as it is always satisfied if $\mathfrak{C}$ consists only of correlation matrices.

(iii) The last part of the assumption states that elements of $\mathfrak{C}$ that are 'close' to being singular can be well approximated by AR(1) correlation matrices. This, together with $\mathfrak{C}_{AR(1)}$ being a subset of $\mathfrak{C}$, readily implies that the singular boundary of $\mathfrak{C}$ must coincide with $\{e_+ e'_+, e_- e'_-\}$. Therefore, we see that the assumption rules out the existence of rank deficient elements of $\mathrm{bd}(\mathfrak{C})$ with rank strictly greater than one. As an example, this rules out the case where $\mathfrak{C}$ is the correlation model corresponding to all stationary autoregressive processes of order less than or equal to two (cf. Lemma G.2 in Preinerstorfer and Pötscher (2016)). If this is not ruled out, however, further obstructions to good size and power properties can arise along suitable sequences approximating these boundary points (cf. Section 3.2.3 in Preinerstorfer and Pötscher (2016)). The possibility of establishing positive results in settings like that is beyond the scope of the present paper and will be discussed elsewhere.

(iv) We note that Assumption 5 is clearly satisfied for every covariance model of the form $\mathfrak{C} = \mathfrak{C}_{AR(1)} \cup \mathfrak{C}^\sharp$, where $\mathfrak{C}^\sharp \subseteq \mathbb{R}^{n \times n}$ is a closed set consisting of positive definite correlation matrices. As an example, let $d \in \mathbb{N}$ be fixed and let $\mathfrak{C}_{MA(d)}$ denote the set of all correlation matrices corresponding to stationary moving average processes of an order not exceeding d, i.e.,

$$\mathfrak{C}_{MA(d)} = \left\{ \Sigma(f_{\alpha,\delta}) : \alpha = (1, \alpha_1, \ldots, \alpha_d)' \in \mathbb{R}^{d+1}, \delta > 0 \right\},$$

where $\Sigma(f_{\alpha,\delta})$ denotes the $n \times n$-dimensional correlation matrix corresponding to the spectral density $f_{\alpha,\delta}(\lambda) = \frac{\delta^2}{2\pi} |\sum_{j=0}^{d} \alpha_j \exp(-\iota\lambda j)|^2$ (cf. Equation (2)). Then $\mathfrak{C} = \mathfrak{C}_{AR(1)} \cup \mathrm{cl}(\mathfrak{C}_{MA(d)})$ satisfies Assumption 5, because every element of the closure of $\mathfrak{C}_{MA(d)}$ is a positive definite correlation matrix (the latter statement follows from Equation (2), compactness of the unit sphere in $\mathbb{R}^{d+1}$, and the Dominated Convergence Theorem).

Under Assumption 5 we shall subsequently establish a positive result concerning tests based on a test statistic T as in (5) with $\hat{\Omega} = \hat{\Omega}_{\kappa,M,p}$. In light of Part (i) of the preceding remark we already know that such a positive result can only be established under non-generic conditions on the design matrix. In particular, the subsequent positive result considers the non-generic case where - besides $g^*_{\kappa,M,p}(.,X,R) \not\equiv 0$, a condition that is generically satisfied under a mild constraint on n (cf. Lemma 3.11) - the column span of the design matrix includes the vectors e_+ and e_- and where $R\hat{\beta}(e_+) = R\hat{\beta}(e_-) = 0$ holds.

Proposition 5.2. *Suppose that the triple κ, M, p satisfies Assumption 1, and that $\mathfrak{C}$ satisfies Assumption 5. Let T be the test statistic defined in Equation (5) with $\hat{\Omega} = \hat{\Omega}_{\kappa,M,p}$. Let $W(C) = \{y \in \mathbb{R}^n : T(y) \geq C\}$ be the rejection region, where C is a real number satisfying $0 < C < \infty$. Suppose further that $e_+, e_- \in \mathfrak{M}$, $R\hat{\beta}(e_+) = R\hat{\beta}(e_-) = 0$ and $g^*_{\kappa,M,p}(.,X,R) \not\equiv 0$. Then, the following holds:*

1. *The size of the rejection region $W(C)$ is strictly less than 1, i.e.,*

$$\sup_{\mu_0 \in \mathfrak{M}_0} \sup_{0 < \sigma^2 < \infty} \sup_{\Sigma \in \mathfrak{C}} P_{\mu_0, \sigma^2 \Sigma}(W(C)) < 1.$$

   Furthermore,

$$\inf_{\mu_0 \in \mathfrak{M}_0} \inf_{0 < \sigma^2 < \infty} \inf_{\Sigma \in \mathfrak{C}} P_{\mu_0, \sigma^2 \Sigma}(W(C)) > 0.$$

2. *The infimal power is bounded away from zero, i.e.,*

$$\inf_{\mu_1 \in \mathfrak{M}_1} \inf_{0 < \sigma^2 < \infty} \inf_{\Sigma \in \mathfrak{C}} P_{\mu_1, \sigma^2 \Sigma}(W(C)) > 0.$$

3. *For every $0 < c < \infty$*

$$\inf_{\substack{\mu_1 \in \mathfrak{M}_1, 0 < \sigma^2 < \infty \\ d(\mu_1, \mathfrak{M}_0)/\sigma \geq c}} P_{\mu_1, \sigma^2 \Sigma_m}(W(C)) \to 1$$

   holds for $m \to \infty$ and for any sequence $\Sigma_m \in \mathfrak{C}$ satisfying $\Sigma_m \to \bar{\Sigma}$ with $\bar{\Sigma}$ a singular matrix. Furthermore, for every sequence $0 < c_m < \infty$

$$\inf_{\substack{\mu_1 \in \mathfrak{M}_1, \\ d(\mu_1, \mathfrak{M}_0) \geq c_m}} \inf_{\Sigma \in \mathfrak{C}^*} P_{\mu_1, \sigma^2_m \Sigma}(W(C)) \to 1$$

   holds for $m \to \infty$ whenever $0 < \sigma^2_m < \infty$, $c_m/\sigma_m \to \infty$, and $\mathfrak{C}^*$ is a (nonempty) closed subset of $\mathfrak{C}$. [The very last statement even holds if one of the conditions $e_+, e_- \in \mathfrak{M}$ and $R\hat{\beta}(e_+) = R\hat{\beta}(e_-) = 0$ is violated.]

4. *For every δ, $0 < \delta < 1$, there exists a $C(\delta)$, $0 < C(\delta) < \infty$, such that*

$$\sup_{\mu_0 \in \mathfrak{M}_0} \sup_{0 < \sigma^2 < \infty} \sup_{\Sigma \in \mathfrak{C}} P_{\mu_0, \sigma^2 \Sigma}(W(C(\delta))) \leq \delta.$$

Under the maintained assumptions on the hypothesis and the design Proposition 5.2 shows that given any level of significance $0 < \delta < 1$, a critical value can

be chosen in such a way that the test obtained holds its size, while its nuisance-minimal power at every point μ_1 in the alternative is bounded away from zero. As Theorem 4.2 in combination with Proposition 4.5 shows, this is impossible for generic elements of the space of all design matrices. Additionally, Part 3 of the proposition shows that the power approaches one in certain parts of the parameter space corresponding to the alternative hypothesis. These parts are characterized by $\|(R\beta^{(1)} - r)/\sigma\|$ being bounded away from zero and $\Sigma \to \bar{\Sigma}$ with $\bar{\Sigma}$ being singular, or $\|(R\beta^{(1)} - r)/\sigma\| \to \infty$ and $\Sigma \to \bar{\Sigma}$ with $\bar{\Sigma}$ positive definite, and where in both cases $\beta^{(1)}$ is the parameter vector corresponding to μ_1 (note that $d(\mu_1, \mathfrak{M}_0)$ is bounded from above and below by multiples of $\|R\beta^{(1)} - r\|$, where the constants involved are positive and depend only on X, R and r).

Remark 5.3. Suppose that instead of Assumption 5 it is known that the covariance model satisfies the following variant of Assumption 5 that rules out AR(1) correlation matrices $\Lambda(\rho)$ with ρ arbitrarily close to -1:

> The covariance model $\mathfrak{C} \subseteq \mathbb{R}^{n \times n}$ is norm-bounded and there exists an $\varepsilon \in (0, 1]$ such that $\mathfrak{C}_{AR(1)}(\varepsilon) \subseteq \mathfrak{C}$ (cf. Remark 4.4). Furthermore, for every sequence $\Sigma_m \in \mathfrak{C}$ that converges to $\bar{\Sigma} \in \mathrm{bd}(\mathfrak{C})$ satisfying $\mathrm{rank}(\bar{\Sigma}) < n$ there exists a corresponding sequence $\rho_m \in (-1 + \varepsilon, 1)$ such that $\Lambda(\rho_m)^{-1/2} \Sigma_m \Lambda(\rho_m)^{-1/2} \to I_n$ as $m \to \infty$.

Let T and $W(C)$ be defined as in Proposition 5.2 above, and suppose further that $e_+ \in \mathfrak{M}$, i.e., the regression contains an intercept, that $R\hat{\beta}(e_+) = 0$, i.e., the hypothesis does not involve the intercept, and that $g^*_{\kappa,M,p}(., X, R) \not\equiv 0$. Then one can show (using essentially the same argument as in the proof of Proposition 5.2) that the Conclusions 1-4 of Proposition 5.2 hold in this setup. In particular, for any given $\delta \in (0, 1)$ there exists a critical value $C(\delta) = C(\delta, \varepsilon)$ such that the test with rejection region $W(C(\delta, \varepsilon))$ has size not greater than δ. This establishes a *positive result* concerning a test based on the test statistic T with $\hat{\Omega} = \hat{\Omega}_{\kappa,M,p}$, and based on the non-standard critical value $C(\delta, \varepsilon)$, which in practice can be obtained as explained in the discussion following Theorem 5.4. However, the test obtained critically depends on ε, which in practice will typically be difficult to choose. If one uses a test with critical region $W(C(\delta, \varepsilon^*))$ where $\varepsilon^* > \varepsilon$, then the size of this test might exceed δ. In light of this drawback, it is important to stress that autocorrelation-robust-testing is possible *without* imposing such an artificial condition that rules out AR(1) correlation matrices $\Lambda(\rho)$ with ρ arbitrarily close to -1: Theorems 5.4 and 5.5 in the present section show that (at the small cost of including the artificial regressor e_-) a generic positive result can be obtained under the more natural Assumption 5. Theorem 5.4 furthermore shows that including the artificial regressor e_- leads to good power properties of the resulting test for covariance matrices close to $\Lambda(-1)$.

Proposition 5.2 assumes, among others, that X satisfies $e_+, e_- \in \mathrm{span}(X)$, an assumption which is satisfied only by non-generic elements of the set of all design matrices. In the following we shall now consider the (generic) situation where $e_+, e_- \in \mathrm{span}(X)$ is *violated*. Proposition 5.2 is then clearly not applicable. However, as we shall see, a positive result similar to Proposition 5.2 can be

established for an *adjusted* test statistic. To explain how the adjustment procedure works, suppose that we have a triple κ, M, p satisfying Assumption 1 which we want to use for covariance estimation. Suppose further that $1 \leq p \leq \frac{n}{k+3}$ holds, which is typically satisfied. The following theorem now shows that under certain conditions on the design matrix - which are shown to be generically satisfied in Proposition 5.5 below, and which in particular require $e_+, e_- \in \text{span}(X)$ to be violated - one can work with an *adjusted* test statistic that has improved size and power properties, and which is constructed as follows: instead of basing the construction of the test statistic on the true design matrix X and on R, we first construct an artificial design matrix $\bar{X}$ of full column rank satisfying $\text{span}(\bar{X}) = \text{span}(X, e_+, e_-)$ by adding the vectors e_+ and/or e_- to X. Furthermore, we construct a corresponding matrix $\bar{R} = (R, 0)$ by appending zero vectors to R such that $\bar{R}$ and $\bar{X}$ have the same number of columns. Then we construct a test statistic $\bar{T}$ as in Equation (5), but with X and R replaced by $\bar{X}$ and $\bar{R}$, respectively, and where the covariance estimator is based on κ, M and p as above besides some minor updates in the construction of M described below. The subsequent theorem shows that if $e_+, e_- \in \text{span}(X)$ is violated, if every $e \in \{e_+, e_-\} \cap \text{span}(X)$ satisfies $R\hat{\beta}(e) = 0$, if $\text{rank}(\bar{X}) < n$, and if an assumption on $\bar{X}$ and $\bar{R}$ analogous to the assumption $g^*_{\kappa,M,p}(., X, R) \not\equiv 0$ in Proposition 5.2 is satisfied, then for every critical value $0 < C < \infty$ the test with critical region $\bar{W}(C) = \{y \in \mathbb{R}^n : \bar{T}(y) \geq C\}$ has the same Properties (1)-(4) as $W(C)$ in Proposition 5.2.

Theorem 5.4. *Suppose that the triple κ, M, p satisfies Assumption 1, that p additionally satisfies $1 \leq p \leq \frac{n}{k+3}$, and that $\mathfrak{C}$ satisfies Assumption 5. Suppose that one of the following (mutually exclusive) scenarios applies:*

1. *$e_+ \in \mathfrak{M}$ with $R\hat{\beta}_X(e_+) = 0$, $e_- \notin \mathfrak{M}$ and $\bar{k} = k+1 < n$. Let $\bar{X} = (X, e_-)$ and define $\bar{R} = (R, 0) \in \mathbb{R}^{q \times \bar{k}}$.*
2. *$e_+ \notin \mathfrak{M}$, $e_- \in \mathfrak{M}$ with $R\hat{\beta}_X(e_-) = 0$ and $\bar{k} = k+1 < n$. Let $\bar{X} = (X, e_+)$ and define $\bar{R} = (R, 0) \in \mathbb{R}^{q \times \bar{k}}$.*
3. *$e_+ \notin \mathfrak{M}$, $e_- \notin \mathfrak{M}$ with $\text{rank}(X, e_+, e_-) = k+2$ and $\bar{k} = k+2 < n$. Let $\bar{X} = (X, e_+, e_-)$ and define $\bar{R} = (R, 0, 0) \in \mathbb{R}^{q \times \bar{k}}$.*
4. *$e_+ \notin \mathfrak{M}$, $e_- \notin \mathfrak{M}$ with $\text{rank}(X, e_+, e_-) = k+1$ and $\bar{k} = k+1 < n$. Let $\bar{X} = (X, e_+)$ and define $\bar{R} = (R, 0) \in \mathbb{R}^{q \times \bar{k}}$.*

Then in all cases $\bar{X}$ is a matrix of full column rank. Define

$$\bar{T}(y) = \begin{cases} (\bar{R}\hat{\beta}_{\bar{X}}(y) - r)' \hat{\Omega}^{-1}_{\kappa,\bar{M},p,\bar{X}}(y)(\bar{R}\hat{\beta}_{\bar{X}}(y) - r) & \text{if } y \notin N^*(\hat{\Omega}_{\kappa,\bar{M},p,\bar{X}}), \\ 0 & \text{else}, \end{cases}$$

where $\hat{\Omega}_{\kappa,\bar{M},p,\bar{X}}(y)$ is the estimator one would obtain following Steps 1-3 in Section 3 if $\bar{X}$ was the underlying design matrix, $(\bar{R}, r)$ was the hypothesis to be tested and where $\bar{M}$ is defined as follows: in case $M \in \mathbb{M}_{KV}$ we set $\bar{M} \equiv M$; in case $M \in \mathbb{M}_{AM}$ we compute $\bar{M}$ as outlined in Section 3.1.1 (using as input $\hat{Z}_{\bar{X}}(y)$ as opposed to $\hat{Z}_X(y)$, and replacing k by $\bar{k}$), with the same constants c_1 and c_2 and j as used in the construction of M, but with ω replaced by $\bar{\omega} = $

$(\omega, 0)' \in \mathbb{R}^{\bar{k}}$; in case $M \in \mathbb{M}_{NW}$ we compute $\bar{M}$ as outlined in Section 3.1.2 (using as input $\hat{Z}_{\bar{X}}(y)$ as opposed to $\hat{Z}_X(y)$), with the same constants $\bar{c}_i$ for $i = 1, 2, 3$ and the same weights w as used in the construction of M, but with ω replaced by $\bar{\omega} = (\omega, 0)' \in \mathbb{R}^{\bar{k}}$. Let $\bar{W}(C) = \{y \in \mathbb{R}^n : \bar{T}(y) \geq C\}$ be the rejection region where C is a real number satisfying $0 < C < \infty$. If $g^*_{\kappa, \bar{M}, p}(., \bar{X}, \bar{R}) \not\equiv 0$ (where $g^*_{\kappa, \bar{M}, p}$ is the function obtained from Lemma 3.10 applied to κ, $\bar{M}$ and p and acting as if $\bar{X}$ was the underlying design matrix and $(\bar{R}, r)$ was the hypothesis to be tested), or equivalently if $N^*(\hat{\Omega}_{\kappa, \bar{M}, p, \bar{X}}) \neq \mathbb{R}^n$, then in each of the four scenarios above the Conclusions (1)-(4) of Proposition 5.2 hold with $W(C)$ replaced by $\bar{W}(C)$.

The procedure outlined in the preceding theorem is based on an artificial design matrix $\bar{X}$ which is obtained from X by adding either one or both elements of the set $\{e_+, e_-\}$ to the columns of X. If the so-obtained matrix $\bar{X}$ satisfies $g^*_{\kappa, \bar{M}, p}(., \bar{X}, \bar{R}) \not\equiv 0$, then the results from Proposition 5.2 carry over to the rejection regions derived from $\bar{T}$. In particular the adjusted test statistic $\bar{T}$ leads to rejection regions the size of which is bounded away from one and such that the nuisance minimal power is bounded away from zero. Besides these improvements, the adjustment procedure is extremely convenient from a computational perspective, as the adjusted test statistic $\bar{T}$ does not require any additional implementations. It is based on the same algorithm as the calculation of T, only with a different design matrix.

The theorem shows that for every level of significance $0 < \delta < 1$, there exists a critical value $C(\delta)$ such that the rejection region $\bar{W}(C(\delta))$ has size smaller than δ. The critical value can be determined as follows: First of all, due to certain invariance properties of $\bar{T}$ (cf. the proof of Theorem 5.4), the probabilities $P_{\mu_0, \sigma^2 \Sigma}(\bar{W}(C))$ do not depend on μ_0 and σ^2. Hence, for any fixed $0 < C < \infty$, the maximal rejection probability under the null can be approximated numerically by simulating the rejection probabilities from a finite subset of $\mathfrak{C}$, and then doing a grid search. In a second step $C(\delta)$ can be approximated by a line search exploiting monotonicity of $P_{\mu_0, \sigma^2 \Sigma}(\bar{W}(C))$ in the critical value.

The adjustment procedure described in Theorem 5.4 is applicable and yields an improved test under the assumption that $e_+, e_- \in \text{span}(X)$ is violated (and hence the positive result in Proposition 5.2 concerning the *unadjusted* test does not apply), that every $e \in \{e_+, e_-\} \cap \text{span}(X)$ satisfies $R\hat{\beta}(e) = 0$, that $\bar{k} < n$ and that $g^*_{\kappa, \bar{M}, p}(., \bar{X}, \bar{R}) \not\equiv 0$. Given a hypothesis (R, r) these are conditions on the design matrix X. Our final result now shows (under mild constraints on n) that these conditions are generically satisfied in the set of all design matrices $\mathfrak{X}_0$; and also in $\tilde{\mathfrak{X}}_0$, the set of all design matrices the first column of which is the intercept, under the additional condition that the first column of R is zero. Under Assumption 5 we hence see that although rejection regions based on T generically break down as a consequence of Proposition 4.5, this problem can generically be resolved by using rejection regions based on the adjusted test statistic $\bar{T}$, unless the regression includes an intercept and the first column of R is nonzero.

Proposition 5.5. *Fix a hypothesis* (R, r) *with* $\mathrm{rank}(R) = q$, *suppose that the triple* κ, M, p *satisfies Assumption 1, that* p *additionally satisfies* $1 \leq p \leq \frac{n}{k+3}$ *and that* $\mathfrak{C}$ *satisfies Assumption 5. Then the following holds, where* $p^* = p + (p \bmod 2)$.

1. *If* $(k+3)(p^* + 2) + p - 1 + \mathbf{1}_{\mathbb{M}_{AM}}(M) \leq n$, *then for* $\lambda_{\mathbb{R}^{n \times k}}$-*almost every design matrix* $X \in \mathfrak{X}_0 \subseteq \mathbb{R}^{n \times k}$ *Scenario 3 in Theorem 5.4 applies, and the Conclusions (1)-(4) of Proposition 5.2 hold for any critical value* $0 < C < \infty$ *with* $W(C)$ *replaced by* $\bar{W}(C) = \{ y \in \mathbb{R}^n : \bar{T}(y) \geq C \}$, *where* $\bar{T}$ *is constructed as outlined in Theorem 5.4.*

2. *Suppose that the first column of* R *is zero, that* $k \geq 2$ *and assume that* $(k+2)(p^* + 2) + p - 1 + \mathbf{1}_{\mathbb{M}_{AM}}(M) \leq n$ *holds. Then for* $\lambda_{\mathbb{R}^{n \times (k-1)}}$-*almost every* $\tilde{X} \in \tilde{\mathfrak{X}}_0$ *Scenario 1 of Theorem 5.4 applies to* $X = (e_+, \tilde{X})$, *and the Conclusions (1)-(4) of Proposition 5.2 hold for any critical value* $0 < C < \infty$ *with* $W(C)$ *replaced by* $\bar{W}(C) = \{ y \in \mathbb{R}^n : \bar{T}(y) \geq C \}$, *where* $\bar{T}$ *is constructed as outlined in Theorem 5.4.*

6. Numerical results

In this section we discuss numerical results to further illustrate our theoretical findings. In both of the subsequent examples sample size n equals 100 and the level of significance is 0.05. We study properties of the following tests (based on the respective design matrix and hypothesis (R, r)): (i) the test that rejects if (5) with κ the Quadratic-Spectral kernel, M the corresponding bandwidth parameter based on auxiliary AR(1) models as suggested in Andrews and Monahan (1992) (cf. the discussion in Section 3.1.1), and prewhitening order $p = 1$ exceeds the χ^2 critical value; (ii) the test that rejects if (5) with κ the Bartlett kernel, M the corresponding bandwidth parameter as suggested in Newey and West (1994) (cf. the discussion in Section 3.1.2), and prewhitening order $p = 1$ exceeds the χ^2 critical value; (iii) the test that rejects if (5) with κ the Bartlett kernel, data-independent bandwidth parameter $M = n - p$ (cf. the discussion in Section 3.1.3 and Rho and Shao (2013)), and prewhitening order $p = 1$ exceeds the corresponding non-standard critical value (which was obtained from Table 1 in Kiefer and Vogelsang (2002)). In Example 2 we also illustrate the effect of applying our adjustment procedure to the tests (i) - (iii) (the adjustment is not applicable in the testing problem considered in Example 1). For the actual computations we used the implementations of the tests (i) - (iii) (and of their adjustments, which are computationally of the same structure, apart from the computation of the critical values, cf. the discussion after Theorem 5.4) provided by the R (R Core Team (2016)) packages **lmtest** (Zeileis and Hothorn (2002)) and **sandwich** (Zeileis (2004)).

6.1. Example 1

The first example concerns testing a zero restriction on the mean in a location model with stationary and Gaussian AR(1) errors, i.e., testing problem (4) with
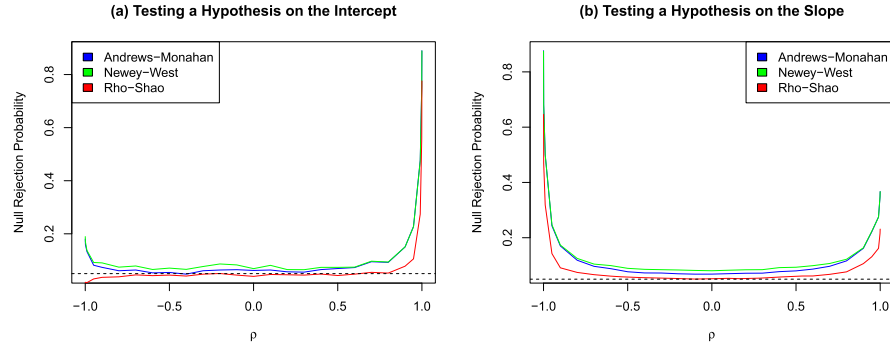
FIG 1. *(a) Example 1: Null rejection probabilities for the tests (i), (ii), and (iii), corresponding to "Andrews-Monahan", "Newey-West", and "Rho-Shao", respectively. (b) Example 2: Null rejection probabilities for the tests (i), (ii), and (iii), corresponding to "Andrews-Monahan", "Newey-West", and "Rho-Shao", respectively. In both (a) and (b) the dashed black line corresponds to* 0.05.

$X = e_+$, $R = 1$, $r = 0$, and covariance model $\mathfrak{C} = \mathfrak{C}_{AR(1)}$. We observe that the triples κ, M, p used in the tests (i) - (iii) satisfy Assumption 1 (cf. Remark 3.7), and that Part 2 of Lemma 3.11 can be used to verify the applicability of Part 4 of Theorem 4.2. This then shows that the size of each test under consideration is 1. To illustrate this numerically, we obtained Monte Carlo approximations of the rejection probabilities of the tests (i) - (iii) under $P_{e_+\beta_1,\Lambda(\rho)}$, recall that $\Lambda(\rho)$ denotes the correlation matrix corresponding to a stationary AR(1) model with parameter ρ, for $\rho \in \{0, \pm0.1, \pm0.2, \pm0.3, \dots, \pm0.9, \pm0.95, \pm0.99, \pm0.999, \pm0.9999\}$ and $\beta_1 \in \{0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1, 1.5, 2\}$, based on 2500 replications in each scenario. The results for $\beta_1 = 0$, i.e., the null rejection probabilities illustrating the above quoted negative result, are shown in Figure 1(a). As expected from the theoretical results (cf. also the explanations after Theorem 4.2), the rejection probabilities are very large when the AR(1) parameter ρ is close to 1. Furthermore, Figure 1(a) shows that for the tests (i) and (ii) the rejection probabilities are clearly above 0.05 when ρ is close to -1. We also observe that while the tests (i) and (ii) show an almost identical behavior in terms of their null rejection probabilities for ρ close to 1, the null rejection probabilities of test (iii), although they eventually get much too large, remain stable for a wider range of values of ρ. Tables 1, 2, and 3 in Appendix E contain all simulated rejection probabilities. These tables also contain Monte Carlo approximations of the rejection probabilities under the alternative (i.e., the simulation results for $\beta_1 > 0$).

6.2. *Example 2*

The second example concerns testing a zero restriction on the slope parameter in a regression model with an intercept, one additional regressor, and with stationary and Gaussian AR(1) errors, i.e., testing problem (4) with $X =$

(e_+, x) (the choice of x is outlined below), $R = (0, 1)$, $r = 0$, and covariance model $\mathfrak{C} = \mathfrak{C}_{AR(1)}$. We proceeded as follows: For each value of $\rho \in \{0, \pm 0.1, \pm 0.2, \pm 0.3, \ldots, \pm 0.9, \pm 0.95, \pm 0.99, \pm 0.999, \pm 0.9999\}$ a regressor x was generated by drawing a random vector from the distribution of n consecutive observations of a stationary Gaussian AR(1) process with parameter ρ. Then rejection probabilities for the tests (i) - (iii) under $P_{x\beta_2, \Lambda(\rho)}$ were obtained for $\beta_2 \in \{0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1, 1.5, 2\}$. In each scenario, i.e., for each combination of ρ and β_2, the Monte Carlo approximations were based on 2500 replications. Since the regressor x was drawn randomly for each value of ρ, the whole procedure was repeated 28 times, and the rejection probabilities were then averaged. The average null rejection probabilities obtained for each $\rho \in \{0, \pm 0.1, \pm 0.2, \pm 0.3, \ldots, \pm 0.9, \pm 0.95, \pm 0.99, \pm 0.999, \pm 0.9999\}$ are given in Figure 1(b). This figure clearly shows that the tests (i) - (iii) under consideration do not control size, not even approximately. The null rejection probabilities become very large for $|\rho|$ close to 1. This effect is more pronounced for ρ close to -1, which is due to the fact that an intercept is included in the model. A complete summary of the rejection probabilities (which also contains rejection probabilities under the alternative) can be found in Tables 4, 6 and 8 in Appendix E. In addition to the tests (i) - (iii) we also studied the behavior of their adjusted versions (for the same set of x vectors), the adjusted versions being obtained following the description in Part 1 of Theorem 5.4. The critical values were obtained by applying the grid-based Monte Carlo procedure explained after Theorem 5.4, where we have chosen an equally spaced grid from $-.99$ to $.99$ with 20 grid points, and 1000 replications for simulating the corresponding rejection probabilities. Of course, in a specific application, where one only needs to obtain a critical value for a single design matrix, the grid can be chosen much finer. The null rejection probabilities of the adjusted versions of the tests (i) - (iii) are shown in Figure 2(a). We see that the rejection probabilities of the adjusted versions of the tests, although they are slightly above 0.05 for large values of $|\rho|$ (due to the coarseness of the grid used), are mostly between 0.03 and 0.04. Comparing these results to the null rejection probabilities of the unadjusted versions of the tests in Figure 1(b), the adjustment procedure, albeit leading to rejection probabilities clearly below 0.05 in most scenarios considered, leads to a substantial improvement concerning the null rejection probabilities. Power properties of the adjusted versions of the tests (i) - (iii) are presented in Figures 2(b)-(d) in the form of contour plots. It can be seen that the three adjusted tests are very similar concerning their power properties. Furthermore, fixing ρ and treating β_2 as a function argument, the power functions become "flatter" for larger values of $|\rho|$. A complete summary of rejection probabilities is provided in Tables 5, 7, and 9 in Appendix E.

7. Conclusion

We have shown that tests for (4) based on prewhitened covariance estimators and possibly data-dependent bandwidth parameters break down in finite samples in terms of their size or power properties. This breakdown arises already for
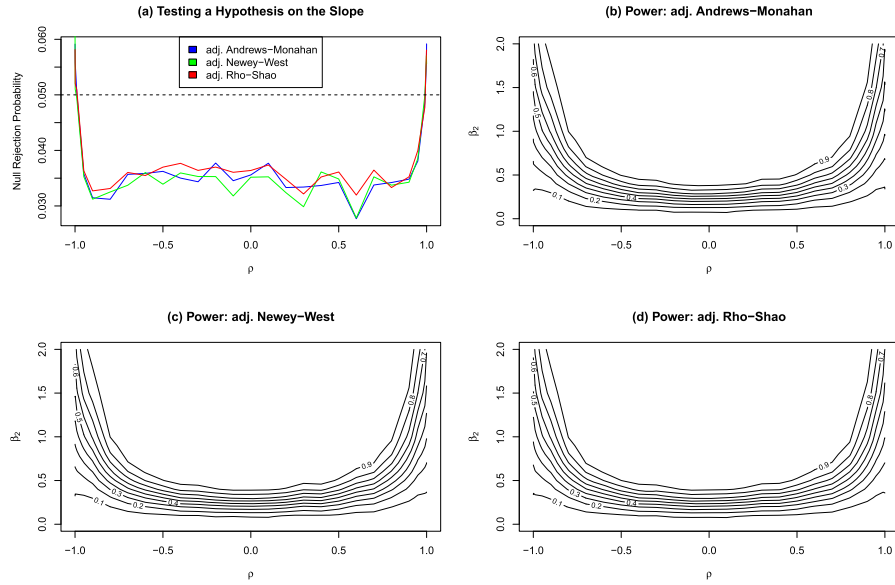
FIG 2. (a) Example 2: Null rejection probabilities for the adjusted versions of the tests (i), (ii), and (iii), corresponding to "adj. Andrews-Monahan", "adj. Newey-West", and "adj. Rho-Shao", respectively. The dashed black line corresponds to 0.05. (b) Example 2: Contour plot of rejection probabilities of the adjusted version of test (i). (c) Example 2: Contour plot of rejection probabilities of the adjusted version of test (ii). (d) Example 2: Contour plot of rejection probabilities of the adjusted version of test (iii).

comparably simple covariance models such as $\mathfrak{C} = \mathfrak{C}_{AR(1)}$. We have also shown how a simple adjustment procedure can generically solve this problem in many cases. The test statistic obtained by applying the adjustment procedure is of the same structural form as the test statistic based on estimators suggested by Andrews and Monahan (1992) and Newey and West (1994) and the test statistic in Rho and Shao (2013), but it is based on an artificial design matrix. Therefore, the adjustment procedure does not only lead to improved size and power properties, but is also convenient from a computational point of view. For the adjustment procedure to work, Assumption 5 has to be satisfied, which requires that elements of the covariance model $\mathfrak{C}$ that are close to being singular are well approximated by AR(1) correlation matrices. If and how the adjustment procedure can be extended to settings where this approximation condition is not satisfied is currently under investigation.

Appendices

Additional notation: For the sake of clarity we shall repeatedly stress the dependence of $\hat{V}$, $\hat{V}_0$, $\hat{V}_p$, $\hat{Z}$, $\hat{A}^{(p)}$, $\hat{u}$, $\hat{\Omega}_{\kappa,M,p}$ and B_p on the design matrix X by writing $\hat{V}_X$, $\hat{V}_{0,X}$, $\hat{V}_{p,X}$, $\hat{Z}_X$, $\hat{A}_X^{(p)}$, $\hat{u}_X$, $\hat{\Omega}_{\kappa,M,p,X}$ and $B_{p,X}$ in the following

proofs. At various places we shall use the following notation: Given a matrix $M \in \mathbb{R}^{m_1 \times m_2}$ and indices $1 \leq i \leq m_1$ and $1 \leq j \leq m_2$ we denote by $[M]_{ij} = M_{ij}$ the ij-th coordinate of M, by $[M]_{.j} = M_{.j}$ the $j - th$ column of M and by $[M]_{i.} = M_{i.}$ the $i - th$ row of M. In case $m_2 = 1$ we write $[M]_i = M_i$ instead of $[M]_{i1}$.

Appendix B: Proofs of results in Section 3.3

Proof of Lemma 3.9. Since X is a matrix of full column rank by assumption, we clearly have $\det(X'X) \neq 0$. From the definition of $\hat{\Omega}_{\kappa,M,p,X}$ we see that $y \in N(\hat{\Omega}_{\kappa,M,p,X})$, i.e., $\hat{\Omega}_{\kappa,M,p,X}(y)$ is not well defined, if and only if one of the following conditions is satisfied (cf. Remark 3.1):

(I) $\det\left(\hat{V}_{0,X}(y)\hat{V}'_{0,X}(y)\right) = 0$;

(II) $\det\left(\hat{V}_{0,X}(y)\hat{V}'_{0,X}(y)\right) \neq 0$ and $\det\left(I_k - \sum_{l=1}^{p} \hat{A}^{(p)}_{l,X}(y)\right) = 0$;

(III) $\det\left(\hat{V}_{0,X}(y)\hat{V}'_{0,X}(y)\right) \neq 0$ and $\det\left(I_k - \sum_{l=1}^{p} \hat{A}^{(p)}_{l,X}(y)\right) \neq 0$ and $M(y)$ is not well defined.

Using $\hat{u}_X(y) = (I_n - \det(X'X)^{-1}X \operatorname{adj}(X'X)X')y$ we see that the coordinates of $\bar{V}_{0,X}(y) := \det(X'X)\hat{V}_{0,X}(y)$ and of $\bar{V}_{p,X}(y) := \det(X'X)\hat{V}_{p,X}(y)$ are values of certain multivariate polynomials defined on $\mathbb{R}^n \times \mathbb{R}^{n \times k}$ evaluated at the point (y, X). Since (I) is equivalent to

$$\det(\det(X'X)^2 \hat{V}_{0,X}(y)\hat{V}'_{0,X}(y)) = \det(\bar{V}_{0,X}(y)\bar{V}'_{0,X}(y)) = 0,$$

this shows that (I) is equivalent to $g_1(y, X) = 0$, say, where $g_1 : \mathbb{R}^n \times \mathbb{R}^{n \times k} \to \mathbb{R}$ is a multivariate polynomial which is clearly independent of (R, r). Using this equivalence, Condition (II) is seen to be equivalent to $g_1(y, X) \neq 0$ and $\det\left(I_k - \sum_{l=1}^{p} \hat{A}^{(p)}_{l,X}(y)\right) = 0$. Because of $g_1(y, X) \neq 0$ we have

$$I_k - \sum_{l=1}^{p} \hat{A}^{(p)}_{l,X}(y)$$

$$= I_k - \hat{V}_{p,X}(y)\hat{V}'_{0,X}(y)\left(\hat{V}_{0,X}(y)\hat{V}'_{0,X}(y)\right)^{-1} D(p),$$

$$= I_k - \bar{V}_{p,X}(y)\bar{V}'_{0,X}(y)\left(\bar{V}_{0,X}(y)\bar{V}'_{0,X}(y)\right)^{-1} D(p),$$

$$= I_k - \det\left(\bar{V}_{0,X}(y)\bar{V}'_{0,X}(y)\right)^{-1} \bar{V}_{p,X}(y)\bar{V}'_{0,X}(y) \operatorname{adj}\left(\bar{V}_{0,X}(y)\bar{V}'_{0,X}(y)\right) D(p),$$

where $D(p) = (I_k, \ldots, I_k)' \in \mathbb{R}^{kp \times k}$. Using this together with similar arguments as above we see that pre-multiplying $I_k - \sum_{l=1}^{p} \hat{A}^{(p)}_{l,X}(y)$ by $\det\left(\bar{V}_{0,X}(y)\bar{V}'_{0,X}(y)\right)$ results in a matrix, the entries of which are values of certain multivariate polynomials, defined on $\mathbb{R}^n \times \mathbb{R}^{n \times k}$, evaluated at the point (y, X). It follows that the second equation in (II) can be replaced by

$$g_2(y, X) := \left[\det\left(\bar{V}_{0,X}(y)\bar{V}'_{0,X}(y)\right)\right]^k \det\left(I_k - \sum_{l=1}^{p} \hat{A}^{(p)}_{l,X}(y)\right) = 0,$$

where $g_2 : \mathbb{R}^n \times \mathbb{R}^{n \times k} \to \mathbb{R}$ is a multivariate polynomial which is independent of (R, r) either. Summarizing our observations concerning (I) and (II) we see that

$$N(\hat{\Omega}_{\kappa,M,p,X}) = \{y \in \mathbb{R}^n : g_1(y, X)g_2(y, X) = 0\}$$

$$\cup \{y \in \mathbb{R}^n : g_1(y, X)g_2(y, X) \neq 0 \text{ and } M(y) \text{ not w.d.}\}.$$

The set in the second line of the previous display depends on the specific bandwidth M. Hence, we have to distinguish three cases: Suppose first that $M \equiv M_{KV} \in \mathbb{M}_{KV}$, i.e., M is a constant which is functionally independent of y, X and thus everywhere well defined on $\mathbb{R}^n$. Define $g_{\kappa,M_{KV},p} \equiv g_1 g_2$, so that $g_{\kappa,M_{KV},p} : \mathbb{R}^n \times \mathbb{R}^{n \times k} \to \mathbb{R}$ is a multivariate polynomial. Noting that

$$N(\hat{\Omega}_{\kappa,M_{KV},p,X}) = \{y \in \mathbb{R}^n : g_{\kappa,M_{KV},p}(y, X) = 0\}$$

then proves the statement in case $M \in \mathbb{M}_{KV}$, because $g_{\kappa,M_{KV},p}$ is independent of (R, r). Next we consider the case $M = M_{AM,j,\omega,c} \in \mathbb{M}_{AM}$, and we start with $j = 1$ (the case $j = 2$ is handled almost identically: omit the square in the definition of D_1 below, the squares from the expression in Equation (9), and adapt the discussion following Equation (9) and the exponents of the pre-multiplying factors accordingly). We shall drop the subindices j and c from $M_{AM,j,\omega,c}$ for notational convenience and write $M_{AM,\omega}$ instead. We partition

$$\{y \in \mathbb{R}^n : g_1(y, X)g_2(y, X) \neq 0 \text{ and } M_{AM,\omega}(y) \text{ not w.d.}\} = D_1 \cup D_2, \quad (8)$$

where D_1 and D_2 are disjoint and defined as

$$D_1 = \left\{y \in \mathbb{R}^n : g_1(y, X)g_2(y, X) \neq 0, \exists i^* : \hat{\rho}_{i^*}(y) \text{ not w.d. or } \hat{\rho}_{i^*}(y)^2 = 1\right\},$$
$$D_2 = \left\{y \in \mathbb{R}^n \backslash D_1 : g_1(y, X)g_2(y, X) \neq 0, \forall i \text{ s.t. } \omega_i \neq 0 : \hat{\sigma}_i^2(y) = 0\right\}.$$

The equality in (8) is readily seen from the definition of $M_{AM,\omega}$. We want to obtain more suitable characterizations of D_1 and D_2 and proceed in two steps: (i) First, we claim that $y \in D_1$ if and only if

$$g_1(y, X)g_2(y, X) \neq 0 \quad \text{and}$$
$$\prod_{i=1}^k \left(\left[\sum_{j=2}^{n-p} [\hat{Z}_X(y)]_{ij} [\hat{Z}_X(y)]_{i(j-1)} \right]^2 - \left[\sum_{j=1}^{n-p-1} [\hat{Z}_X(y)]_{ij}^2 \right]^2 \right) = 0. \quad (9)$$

To see this assume that $g_1(y, X)g_2(y, X) \neq 0$ holds: Suppose that $\hat{\rho}_{i^*}(y)$ is not well defined. The latter occurs if and only if $\sum_{j=1}^{n-p-1} [\hat{Z}_X(y)]_{i^*j}^2 = 0$, i.e., all summands are zero, which immediately implies $\sum_{j=2}^{n-p} [\hat{Z}_X(y)]_{i^*j} [\hat{Z}_X(y)]_{i^*(j-1)} = 0$. Therefore, the factor corresponding to index i^* vanishes and thus the product defining the second equation in (9) vanishes. That $\hat{\rho}_{i^*}^2(y) = 1$ implies that the product vanishes is obvious. To prove the other direction assume that $g_1(y, X)g_2(y, X) \neq 0$ and that the product vanishes. This implies that at least one factor with index i^*, say, equals zero, which implies that either $\hat{\rho}_{i^*}(y)$ is

not well defined or $\hat{\rho}_{i*}^2(y) = 1$ holds. This proves the claim. Secondly, we recall that if $g_1(y, X)g_2(y, X) \neq 0$, then $\hat{Z}_X(y) = \hat{V}_{p,X}(y) - \hat{A}_X^{(p)}(y)\hat{V}_{0,X}(y)$. Using an argument as above it is then easy to see that $\hat{Z}_X(y)$ pre-multiplied by

$$\det\left(\bar{V}_{0,X}(y)\bar{V}_{0,X}'(y)\right)\det(X'X) \tag{10}$$

gives a matrix, the entries of which are values of certain multivariate polynomials defined on $\mathbb{R}^n \times \mathbb{R}^{n \times k}$ evaluated at the point (y, X). Thus, if we multiply the second equation in (9) by the $(4k)$-th power of the expression in the previous display we see that Equation (9) can equivalently be written as

$$g_1(y, X)g_2(y, X) \neq 0 \text{ and } g_{AM,1}(y, X) = 0,$$

where $g_{AM,1} : \mathbb{R}^n \times \mathbb{R}^{n \times k} \to \mathbb{R}$ is a multivariate polynomial that is independent of (R, r). Summarizing, we have shown that

$$D_1 = \{y \in \mathbb{R}^n : g_1(y, X)g_2(y, X) \neq 0, g_{AM,1}(y, X) = 0\}.$$

(ii) First we observe that $y \in D_2$ if and only if

$$g_1(y, X)g_2(y, X)g_{AM,1}(y, X) \neq 0 \quad \text{and}$$
$$\sum_{i=1}^{k} \omega_i \sum_{j=2}^{n-p} \left([\hat{Z}_X(y)]_{ij} - \hat{\rho}_i(y)[\hat{Z}_X(y)]_{i(j-1)}\right)^2 = 0, \tag{11}$$

where we recall that by assumption ω is functionally independent of y and X. Because $g_{AM,1}(y, X) \neq 0$ implies that $\hat{\rho}_i(y)$ is well defined for $i = 1, \ldots, k$, which is equivalent to $\sum_{l=1}^{n-p-1}[\hat{Z}_X(y)]_{il}^2 \neq 0$ for $i = 1, \ldots, k$, the second equation in the previous display can be replaced by

$$\sum_{i=1}^{k} \omega_i \sum_{j=2}^{n-p} \left(\sum_{l=1}^{n-p-1} [\hat{Z}_X(y)]_{il}^2 [\hat{Z}_X(y)]_{ij} - \sum_{l=2}^{n-p}[\hat{Z}_X(y)]_{il}[\hat{Z}_X(y)]_{i(l-1)}[\hat{Z}_X(y)]_{i(j-1)}\right)^2 = 0.$$

We now multiply the function defining this equation by the 6-th power of the expression in Equation (10) and denote the resulting function by $g_{AM,\omega,2}(y, X)$. The statement in Equation (11) is then seen to be equivalent to

$$g_1(y, X)g_2(y, X)g_{AM,1}(y, X) \neq 0 \text{ and } g_{AM,\omega,2}(y, X) = 0,$$

where $g_{AM,\omega,2} : \mathbb{R}^n \times \mathbb{R}^{n \times k} \to \mathbb{R}$ is a multivariate polynomial. We also see that $g_{AM,\omega,2}$ is independent of (R, r). We conclude that

$$D_2 = \{y \in \mathbb{R}^n : g_1(y, X)g_2(y, X)g_{AM,1}(y, X) \neq 0, g_{AM,\omega,2}(y, X) = 0\}.$$

Now let $g_{\kappa, M_{AM,\omega}, p} \equiv g_1 g_2 g_{AM,1} g_{AM,\omega,2}$. By what has been shown above $g_{\kappa, M_{AM,\omega}, p} : \mathbb{R}^n \times \mathbb{R}^{n \times k} \to \mathbb{R}$ is a multivariate polynomial. Furthermore, $g_{\kappa, M_{AM,\omega}, p}$ does not depend on (R, r). We observe that

$$N(\hat{\Omega}_{\kappa, M_{AM,\omega}, p, X}) = \{y \in \mathbb{R}^n : g_{\kappa, M_{AM,\omega}, p}(y, X) = 0\}.$$

This proves the lemma in case $M \in \mathbb{M}_{AM}$. Finally, we consider $M \equiv M_{NW,\omega,w} \in \mathbb{M}_{NW}$, where we write $M_{NW,\omega,w}$ instead of $M_{NW,\omega,w,\bar{c}}$, because the argument and the resulting polynomial are independent of $\bar{c}$. We use a similar argument as in the previous case. We observe that if $g_1(y, X)g_2(y, X) \neq 0$, then the function $M_{NW,\omega,w}$ is not well defined if and only if

$$\sum_{i=-(n-p-1)}^{n-p-1} w(i)\bar{\sigma}_i(y) = 0, \tag{12}$$

where

$$\bar{\sigma}_i(y) = (n-p)^{-1} \sum_{j=|i|+1}^{n-p} \omega'[\hat{Z}_X(y)]_{\cdot j} \left([\hat{Z}_X(y)]_{\cdot(j-|i|)}\right)' \omega \text{ for } |i| = 0, \ldots n-p-1.$$

Since ω and w are both functionally independent of X and y, we can premultiply Equation (12) by the square of the expression in Equation (10) to see that the statement $g_1(y, X)g_2(y, X) \neq 0$ and $M_{NW,\omega,w}$ not being well defined is equivalent to

$$g_1(y, X)g_2(y, X) \neq 0 \text{ and } g_{NW,\omega,w}(y, X) = 0,$$

where $g_{NW,\omega,w} : \mathbb{R}^n \times \mathbb{R}^{n \times k} \to \mathbb{R}$ is a multivariate polynomial. The function $g_{NW,\omega,w}$ is independent of (R, r). Using these properties, defining $g_{\kappa,M_{NW,\omega,w},p} = g_1 g_2 g_{NW,\omega,w}$, a function which does not depend on (R, r), and noting that

$$N(\hat{\Omega}_{\kappa,M_{NW,\omega,w},p,X}) = \left\{y \in \mathbb{R}^n : g_{\kappa,M_{NW,\omega,w},p}(y, X) = 0\right\},$$

then proves the claim in case $M \in \mathbb{M}_{NW}$. $\square$

Proof of Lemma 3.10. To establish Parts 1-4 of the lemma we apply a similar argument as in the proof of Lemma 3.1 in Preinerstorfer and Pötscher (2016). We observe that if $y \notin N(\hat{\Omega}_{\kappa,M,p,X})$, or equivalently $g_{\kappa,M,p}(y, X) \neq 0$, we can write $\hat{\Omega}_{\kappa,M,p,X}(y)$ as

$$\hat{\Omega}_{\kappa,M,p,X}(y) = \frac{n}{n-p} B_{p,X}(y) \mathcal{W}_{n-p}(y) B'_{p,X}(y), \tag{13}$$

where $\mathcal{W}_{n-p}(y) \in \mathbb{R}^{(n-p) \times (n-p)}$ is the symmetric Toeplitz matrix with ones on the main diagonal, and where for $i \neq j$ its ij-th coordinate is given by $\kappa((i-j)/M(y))$ whenever $M(y) \neq 0$, and by 0 else. Recall that $M(y) \geq 0$. If $M(y) = 0$ we have $\mathcal{W}_{n-p}(y) = I_n$. If $M(y) > 0$ the matrix $\mathcal{W}_{n-p}(y)$ is positive definite by Assumption 1. Therefore, in both cases the matrix $\mathcal{W}_{n-p}(y)$ is positive definite. This immediately establishes Parts 1-4, where $\text{rank}(R) = q$ is used in proving Part 4 (we emphasize that $\hat{\Omega}_{\kappa,M,p,X}(y)$ can be nonnegative definite, singular, zero or positive definite only if it is well defined, i.e., only if $g_{\kappa,M,p}(y, X) \neq 0$ holds). It remains to prove Part 5. We recall that

$$\begin{aligned} &N^*(\hat{\Omega}_{\kappa,M,p,X}) \\ = \quad &N(\hat{\Omega}_{\kappa,M,p,X}) \cup \left\{y \in \mathbb{R}^n \backslash N(\hat{\Omega}_{\kappa,M,p,X}) : \det\left[\hat{\Omega}_{\kappa,M,p,X}(y)\right] = 0\right\}. \end{aligned} \tag{14}$$

From Part 2 of the present lemma we know that we can rewrite the second set to the right as

$$\left\{ y \in \mathbb{R}^n \backslash N(\hat{\Omega}_{\kappa,M,p,X}) : \det \left[B_{p,X}(y) B'_{p,X}(y) \right] = 0 \right\}. \tag{15}$$

For every $y \in \mathbb{R}^n \backslash N(\hat{\Omega}_{\kappa,M,p,X})$ we have with $D(p) = (I_k, \ldots, I_k)' \in \mathbb{R}^{(kp) \times k}$ that (using the same notation as in the proof of Lemma 3.9) $B_{p,X}(y)$ can be written as

$$R(X'X)^{-1} \left(I_k - \bar{V}_{p,X}(y) \bar{V}'_{0,X}(y) \left[\bar{V}_{0,X}(y) \bar{V}'_{0,X}(y) \right]^{-1} D(p) \right)^{-1}$$
$$\times \left(\hat{V}_{p,X}(y) - \bar{V}_{p,X}(y) \bar{V}'_{0,X}(y) \left[\bar{V}_{0,X}(y) \bar{V}'_{0,X}(y) \right]^{-1} \hat{V}_{0,X}(y) \right)$$

$$= R(X'X)^{-1} \det(X'X)^{-1}$$
$$\times \left(\det([\bar{V}_{0,X}(y) \bar{V}'_{0,X}(y)]) I_k - \bar{V}_{p,X}(y) \bar{V}'_{0,X}(y) \operatorname{adj}\left[\bar{V}_{0,X}(y) \bar{V}'_{0,X}(y) \right] D(p) \right)^{-1}$$
$$\times \left(\det\left[\bar{V}_{0,X}(y) \bar{V}'_{0,X}(y) \right] \bar{V}_{p,X}(y) - \bar{V}_{p,X}(y) \bar{V}'_{0,X}(y) \operatorname{adj}\left[\bar{V}_{0,X}(y) \bar{V}'_{0,X}(y) \right] \bar{V}_{0,X}(y) \right)$$

$$= \det(X'X)^{-2}$$
$$\times \det \left(\det([\bar{V}_{0,X}(y) \bar{V}'_{0,X}(y)]) I_k - \bar{V}_{p,X}(y) \bar{V}'_{0,X}(y) \operatorname{adj}\left[\bar{V}_{0,X}(y) \bar{V}'_{0,X}(y) \right] D(p) \right)^{-1}$$
$$\times R \operatorname{adj}(X'X)$$
$$\times \operatorname{adj} \left(\det([\bar{V}_{0,X}(y) \bar{V}'_{0,X}(y)]) I_k - \bar{V}_{p,X}(y) \bar{V}'_{0,X}(y) \operatorname{adj}\left[\bar{V}_{0,X}(y) \bar{V}'_{0,X}(y) \right] D(p) \right)$$
$$\times \left(\det\left[\bar{V}_{0,X}(y) \bar{V}'_{0,X}(y) \right] \bar{V}_{p,X}(y) - \bar{V}_{p,X}(y) \bar{V}'_{0,X}(y) \operatorname{adj}\left[\bar{V}_{0,X}(y) \bar{V}'_{0,X}(y) \right] \bar{V}_{0,X}(y) \right)$$

We therefore see that the coordinates of the matrix $\bar{B}_{p,X}(y)$, say, which is obtained by pre-multiplying $B_{p,X}(y)$ by the factor $F_p(y, X)$, defined as the product of $\det(X'X)^2$ and

$$\det \left(\det([\bar{V}_{0,X}(y) \bar{V}'_{0,X}(y)]) I_k - \bar{V}_{p,X}(y) \bar{V}'_{0,X}(y) \operatorname{adj}\left[\bar{V}_{0,X}(y) \bar{V}'_{0,X}(y) \right] D(p) \right),$$

(for later reference we note that $F_p : \mathbb{R}^n \times \mathbb{R}^{n \times k} \to \mathbb{R}$ is a multivariate polynomial), are values of certain multivariate polynomials defined on $\mathbb{R}^n \times \mathbb{R}^{n \times k}$ evaluated at (y, X). Furthermore, we can replace $B_{p,X}(y)$ in Equation (15) by $\bar{B}_{p,X}(y)$ without changing the set. This follows because $y \notin N(\hat{\Omega}_{\kappa,M,p})$ implies

$$F_p(y, X) = \det(X'X)^2 \det(\bar{V}_{0,X}(y) \bar{V}'_{0,X}(y))^k \det \left(I_k - \sum_{l=1}^{p} \hat{A}_{l,X}^{(p)}(y) \right) \neq 0.$$

If we combine this equivalent expression for (15) with (14) and Lemma 3.9 we obtain

$$N^*(\hat{\Omega}_{\kappa,M,p,X}) = \left\{ y \in \mathbb{R}^n : g_{\kappa,M,p}(y, X) \det \left[\bar{B}_{p,X}(y) \bar{B}'_{p,X}(y) \right] = 0 \right\}.$$

We next define $g^*_{\kappa,M,p}(y, X, R) \equiv g_{\kappa,M,p}(y, X) \det[\bar{B}_{p,X}(y) \bar{B}'_{p,X}(y)]$. By Lemma 3.9 we see that $g^*_{\kappa,M,p} : \mathbb{R}^n \times \mathbb{R}^{n \times k} \times \mathbb{R}^{q \times k} \to \mathbb{R}$ is a multivariate polynomial that does not depend on r. $\square$

The subsequent technical lemma plays a key role in several constructions in the proofs of the genericity results.

Lemma B.1. *Let $1 \leq k < n$, $n > 2$ and let (R, r) be a hypothesis. Suppose that the triple κ, M, p satisfies Assumption 1. Assume that the tuple $(y, X) \in \mathbb{R}^n \times \mathfrak{X}_0$ satisfies for some $t \geq k$:*

(A1) $\hat{V}_X(y)$ has exactly $t + 1$ nonzero columns with indices $1 = j_1 < j_2 < \ldots < j_{t+1} \leq n$.

(A2) $j_{i+1} - j_i \geq p + 1$ for $i = 1, \ldots, t$, and $n - j_{t+1} \geq p - 1$.

(A3) If $t = k$, then $\mathrm{rank}(\hat{V}_X(y)) = k$. Otherwise,

$$\mathrm{span}(\left\{ [\hat{V}_X(y)]_{\cdot j_i} : i = 1, \ldots, t \right\})$$
$$= \mathrm{span}(\left\{ [\hat{V}_X(y)]_{\cdot j_i} : i = 2, \ldots, t+1 \right\}) = \mathbb{R}^k.$$

Then, the following holds:

1. *$\hat{A}_X^{(p)}(y) = 0$ and $\mathrm{rank}(\hat{Z}_X(y)) = k$.*
2. *Under each of the following three conditions we have $g_{\kappa,M,p}^*(y, X, R) \neq 0$ (or equivalently $y \notin N^*(\hat{\Omega}_{\kappa,M,p,X})$):*

(CKV) $M \in \mathbb{M}_{KV}$;

(CAM) $M \in \mathbb{M}_{AM}$, and every row vector of the matrix obtained from $\hat{Z}_X(y)$ by deleting its last column is nonzero [this is in particular satisfied if $n - j_{t+1} > p - 1$];

(CNW) $M \in \mathbb{M}_{NW}$, and either each coordinate of $\omega' \hat{Z}_X(y)$ is non-negative, or each coordinate of $\omega' \hat{Z}_X(y)$ is non-positive.

3. *For every $Q \in \mathbb{R}^{k \times k}$ such that $\mathrm{rank}(Q) = k$, the tuple (y, XQ) is an element of $\mathbb{R}^n \times \mathfrak{X}_0$ that satisfies (A1), (A2) and (A3).*
4. *If $k \geq 2$ and either $[\hat{V}_X(y)]_{1j_i} > 0$ for $i = 2, \ldots, t+1$ or $[\hat{V}_X(y)]_{1j_i} < 0$ for $i = 2, \ldots, t+1$ holds, then there exists a regular matrix $Q \in \mathbb{R}^{k \times k}$ such that the first columns of X and XQ, respectively, coincide and such that $g_{\kappa,M,p}^*(y, XQ, R) \neq 0$ (or equivalently $y \notin N^*(\hat{\Omega}_{\kappa,M,p,XQ})$).*

Proof. Denote the column vectors of $\hat{V}_X(y)$ by v_i for $i = 1, \ldots, n$. If $t > k$, then by (A3) the set $\{v_{j_1}, \ldots, v_{j_t}\}$ and $\{v_{j_2}, \ldots, v_{j_{t+1}}\}$, respectively, spans $\mathbb{R}^k$. Using (A3), we now show that this is automatically satisfied in case $t = k$. To prove this claim, we first recall that $v_j = [\hat{u}_X(y)]_j X_{j\cdot}'$. We see that $\hat{u}_X(y) \perp \mathrm{span}(X)$ implies

$$0 = \sum_{j=1}^{n} [\hat{u}_X(y)]_j X_{j\cdot}' = \sum_{j=1}^{n} v_j = \sum_{i=1}^{k+1} v_{j_i},$$

where the third equality follows from (A1) $(t = k)$. This shows that

$$v_{j_1} = -\sum_{i=2}^{k+1} v_{j_i}$$

$$v_{j_{k+1}} = -\sum_{i=1}^{k} v_{j_i}.$$

By (A3) $\operatorname{rank}(\hat{V}_X(y)) = k$, which together with (A1) implies that $\operatorname{span}(\{v_{j_i} : i = 1, \ldots, k+1\}) = \mathbb{R}^k$. Therefore, it follows from the two equations in the previous display that $\{v_{j_i} : i = 1, \ldots, k\}$ and $\{v_{j_i} : i = 2, \ldots, k+1\}$, respectively, spans $\mathbb{R}^k$. Hence the claim follows. We next show that $\hat{A}_X^{(p)}(y)$ is well defined. For this we have to verify that $\operatorname{rank}(\hat{V}_{0,X}(y)) = kp$ (cf. Remark 3.1). The j-th column $(j = 1, \ldots, n - p)$ of $\hat{V}_{0,X}(y)$ is given by

$$(v'_{j+p-1}, \ldots, v'_{j+1}, v'_j)' \in \mathbb{R}^{kp}, \tag{16}$$

which is to be interpreted as v_j if $p = 1$, as $(v'_{j+1}, v'_j)'$ if $p = 2$ etc. For $l = 1, \ldots, p$ we define the $(kp) \times k$ dimensional auxiliary matrix $D_l = e_l(p) \otimes I_k$, where $e_l(p)$ denotes the l-th element of the canonical basis of $\mathbb{R}^p$ (and $\otimes$ denotes the Kronecker product). The following claims are immediate consequences of the structure of $\hat{V}_X(y)$ implied by (A1) - (A2):

(I) $D_p v_{j_1}$ is the first column of $\hat{V}_{0,X}(y)$;
(II) If $t \geq 2$, then $D_l v_{j_i}$ for $i = 2, \ldots, t$ and $l = 1, \ldots, p$ are columns of $\hat{V}_{0,X}(y)$;
(III) If $p \geq 2$, then $D_l v_{j_{t+1}}$ for $l = 1, \ldots, (p-1)$ are columns of $\hat{V}_{0,X}(y)$.

To see Parts (I) and (II), we observe that (A1) and (A2) imply that for $i = 1, \ldots, t$, there are at least p zero columns between the columns v_{j_i} and $v_{j_{i+1}}$ of $\hat{V}_X(y)$. Equation (16) together with $j_1 = 1$ then immediately implies Parts (I) and (II). Now we consider Part (III) and hence assume that $p \geq 2$. We start with the case $l = (p-1)$. Every column of $\hat{V}_X(y)$ with index greater than j_{t+1} is zero by Assumption (A1). By Assumption (A2) we have $n - j_{t+1} \geq p - 1$. Together, this implies that the column $v_{j_{t+1}}$ is followed by at least $p-1$ zero columns. Since $j_{t+1} - j_t \geq p+1$ by Assumption (A2), the column $v_{j_{t+1}}$ is preceded by at least p zero columns. The assumption $n - j_{t+1} \geq p-1$ is equivalent to $n - p \geq j_{t+1} - 1$. Hence, denoting the $m_1 \times m_2$-dimensional zero matrix by $0_{m_1, m_2}$, we can use Equation (16) with $j = j_{t+1} - 1$ to see that

$$(v'_{j_{t+1}+p-2}, \ldots, v'_{j_{t+1}}, v'_{j_{t+1}-1})' = \begin{cases} (0_{1,(p-2)k}, v'_{j_{t+1}}, 0_{1,k})' & \text{if } p > 2 \\ (v'_{j_{t+1}}, 0_{1,k})' & \text{if } p = 2, \end{cases}$$

is a column of $\hat{V}_{0,X}(y)$, where in deriving the equality we made use of the already established fact that the column $v_{j_{t+1}}$ of $\hat{V}_X(y)$ is preceded by at least $p > 1$ zero columns (which implies that $v'_{j_{t+1}-1}$ is the zero vector), and followed by at least $p-1$ zero columns (which is used in case $p > 2$). This proves the statement

concerning $D_{p-1}v_{j_{t+1}}$. In case $p = 2$ we are done. If $p > 2$, then the statements concerning $D_l v_{j_{t+1}}$ for $l = 1, \ldots, p - 2$ follow from Equation (16) together with the equation in the previous display and the fact that $v_{j_{t+1}}$ is preceded by at least p zero columns.

We now use (I)-(III) and $\text{span}(\{v_{j_1}, \ldots, v_{j_t}\}) = \text{span}(\{v_{j_2}, \ldots, v_{j_{t+1}}\}) = \mathbb{R}^k$ to show that $\text{rank}(\hat{V}_{0,X}(y)) = kp$: The matrix $\hat{V}_{0,X}(y)$ is $kp \times (n-p)$ dimensional. Therefore, we must show that it has full row rank. Assume existence of a row vector $\xi = (\xi_1, \ldots, \xi_p)$, where $\xi_i' \in \mathbb{R}^k$ for $i = 1, \ldots, p$, such that $\xi \hat{V}_{0,X}(y) = 0$ holds. Part (I) shows that $0 = \xi D_p v_{j_1} = \xi_p v_{j_1}$. If $t \geq 2$, then Part (II) applied with $l = p$ shows that $0 = \xi D_p v_{j_i} = \xi_p v_{j_i}$ for $i = 2, \ldots, t$. Summarizing the cases $t = 1$ and $t \geq 2$ we obtain $\xi_p v_{j_i} = 0$ for $i = 1, \ldots, t$. Because $\{v_{j_1}, \ldots, v_{j_t}\}$ spans $\mathbb{R}^k$, it follows that $\xi_p = 0$. If $p \geq 2$, Part (III) implies that $0 = \xi D_l v_{j_{t+1}} = \xi_l v_{j_{t+1}}$ for $l = 1, \ldots, (p-1)$. If $t \geq 2$ Part (II) implies $0 = \xi D_l v_{j_i} = \xi_l v_{j_i}$ for $l = 1, \ldots, (p-1)$ and $i = 2, \ldots, t$. Summarizing again the cases $t = 1$ and $t \geq 2$ we obtain that for $l = 1, \ldots, (p-1)$ we have

$$\xi_l v_{j_i} = 0 \text{ for } i = 2, \ldots, t+1.$$

Because $\{v_{j_2}, \ldots, v_{j_{t+1}}\}$ spans $\mathbb{R}^k$, it follows from the previous display that $\xi_l = 0$ for $l = 1, \ldots, p-1$. Since we already know that $\xi_p = 0$, we obtain $\xi = 0$ and thus $\text{rank}(\hat{V}_{0,X}(y)) = kp$. Therefore $\hat{A}_X^{(p)}(y)$ is well defined. To see that $\hat{A}_X^{(p)}(y) = 0$ we observe that every nonzero column of $\hat{V}_X(y)$ besides the first one is preceded by at least p zero columns. The matrix $\hat{V}_{p,X}(y)$ is obtained from $\hat{V}_X(y)$ by deleting the first $p \geq 1$ columns. This together with Equation (16) immediately implies $\hat{V}_{p,X}(y)\hat{V}_{0,X}'(y) = 0$ and thus $\hat{A}_X^{(p)}(y) = 0$. Before we show that $y \notin N^*(\hat{\Omega}_{\kappa,M,p,X})$ under the conditions (CKV), (CAM) and (CNW), respectively, we note that $\text{rank}(\hat{Z}_X(y)) = k$. This follows, because $\hat{A}_X^{(p)}(y) = 0$ implies $\hat{Z}_X(y) = \hat{V}_{p,X}(y)$, which together with $j_2 - j_1 \geq p + 1$ shows that the vectors v_{j_i} for $i = 2, \ldots, t+1$ (which span $\mathbb{R}^k$) are column vectors of $\hat{Z}_X(y)$.

Now, as a consequence of Part 4 of Lemma 3.10 positive definiteness of $\hat{\Omega}_{\kappa,M,p,X}(y)$ and hence $y \notin N^*(\hat{\Omega}_{\kappa,M,p,X})$ follows if we can show that $\hat{\Omega}_{\kappa,M,p,X}(y)$ is well defined. Since $\hat{A}_X^{(p)}(y) = 0$ implies invertibility of $I_k - \sum_{l=1}^p \hat{A}_{i,X}^{(p)}(y)$, it remains to show that M is well defined at y (cf. Remark 3.1). This is trivially satisfied under Condition (CKV) because $M \in \mathbb{M}_{KV}$ is everywhere well defined. Suppose that Condition (CAM) holds, i.e., $M \in \mathbb{M}_{AM}$ and every row vector of the matrix obtained from $\hat{Z}_X(y)$ by deleting the last column is nonzero. That the latter condition is satisfied if $n - j_{t+1} > p - 1$ follows because in that case the last column of $\hat{Z}_X(y)$ is the zero vector and $\text{rank}(\hat{Z}_X(y)) = k$. Under the assumption that every row vector of the matrix obtained from $\hat{Z}_X(y)$ by deleting the last column is nonzero it is obvious that the denominators in the definition of $\hat{\rho}_i(y)$ for $i = 1, \ldots, k$, i.e.,

$$\sum_{j=1}^{n-p-1} [\hat{Z}_X(y)]_{ij}^2 \quad \text{for} \quad i = 1, \ldots, k,$$

do not vanish. Therefore, $\hat{\rho}_i(y)$ for $i = 1, \ldots, k$ are well defined. Using Assumptions (A1) and (A2) together with $p \geq 1$ and $\hat{Z}_X(y) = \hat{V}_{p,X}(y)$, it follows that there is always at least one zero column between two nonzero columns of $\hat{Z}_X(y)$. Therefore, it is clear that the numerators appearing in the definition of $\hat{\rho}_i(y)$ for $i = 1, \ldots, k$, i.e.,

$$\sum_{j=2}^{n-p} [\hat{Z}_X(y)]_{ij} [\hat{Z}_X(y)]_{i(j-1)} \quad \text{for} \quad i = 1, \ldots, k,$$

must vanish. It follows that $\hat{\rho}_i(y) = 0$ for $i = 1, \ldots, k$. We finally show that $\hat{\sigma}_i(y) > 0$ for $i = 1, \ldots, k$. We note that $\hat{\rho}_i(y) = 0$ for $i = 1, \ldots, k$ implies

$$\hat{\sigma}_i(y) = (n - p - 1)^{-1} \sum_{j=2}^{n-p} [\hat{Z}_X(y)]_{ij}^2 \text{ for } i = 1, \ldots, k.$$

Because of $\hat{Z}_X(y) = \hat{V}_{p,X}(y)$ it follows from Assumptions (A1) and (A2) that the first column of $\hat{Z}_X(y)$ must be zero. Furthermore, we already know that $\text{rank}(\hat{Z}_X(y)) = k$. This implies that the matrix Z_*, say, which is obtained from $\hat{Z}_X(y)$ by deleting the first column, must be of full row rank k. Consequently all rows of Z_* must be non-zero. The previous display shows that $\hat{\sigma}_i(y)$ for $i = 1, \ldots, k$ is, up to a positive factor, the squared Euclidean norm of the i-th row of Z_*. Therefore $\hat{\sigma}_i(y) > 0$ for $i = 1, \ldots, k$ must hold. Therefore, we have shown that $M(y)$ is well defined (we even see that $M(y) = 0$ holds). Now we consider the case where Condition (CNW) holds, i.e., $M \in \mathbb{M}_{NW}$ and every coordinate of $\omega' \hat{Z}_X(y)$ is non-negative (non-positive). We have to show that

$$\sum_{i=-(n-p-1)}^{n-p-1} w(i) \bar{\sigma}_i(y) \neq 0.$$

The non-negativity (non-positivity) condition immediately implies $\bar{\sigma}_i(y) \geq 0$ for $|i| = 0, 1, \ldots, n - p - 1$. Furthermore, since $\text{rank}(\hat{Z}_X(y)) = k$ and $\omega \neq 0$, by the definition of the weights vector, we have $\omega' \hat{Z}_X(y) \neq 0$. This implies $\bar{\sigma}_0(y) = (n - p)^{-1} \|\omega' \hat{Z}_X(y)\|^2 > 0$. By assumption $w(0) = 1$ and $w(i) \geq 0$ for $|i| = 1, \ldots, n - p - 1$. Therefore, the quantity in the previous display does not vanish and thus $M(y)$ is well defined. This proves the second part of the lemma.

We next prove Part 3. Let Q be a regular $k \times k$ dimensional matrix. First, we obviously have $XQ \in \mathfrak{X}_0$, because $X \in \mathfrak{X}_0$ and Q is regular. Secondly, since $\text{span}(X) = \text{span}(XQ)$, using regularity of Q, we have that $\hat{u}_X(y) = \hat{u}_{XQ}(y)$. This immediately entails

$$\hat{V}_{XQ}(y) = (XQ)' \text{diag}(\hat{u}_{XQ}(y)) = (XQ)' \text{diag}(\hat{u}_X(y)) = Q'X' \text{diag}(\hat{u}_X(y))$$
$$= Q' \hat{V}_X(y).$$

Therefore, the tuple $(y, XQ) \in \mathbb{R}^n \times \mathfrak{X}_0$ satisfies (A1), (A2) and (A3), because (y, X) does so and Q is regular.

It remains to prove Part 4. We do this by constructing a Q as in Part 3 such that the tuple $(y, XQ) \in \mathbb{R}^n \times \mathfrak{X}_0$ satisfies Condition (CKV), (CAM) or (CNW), respectively, if $M \in \mathbb{M}_{KV}$, $M \in \mathbb{M}_{AM}$ or $M \in \mathbb{M}_{NW}$, respectively. If $M \in \mathbb{M}_{KV}$ we can obviously choose $Q = I_k$. Suppose that $M \notin \mathbb{M}_{KV}$. Let $Q \in \mathbb{R}^{k \times k}$ be such that $\operatorname{rank}(Q) = k$. We specify this matrix later on. From Part 2 of the present lemma we see that the tuple (y, XQ) satisfies (A1), (A2) and (A3) and therefore we can conclude from Part 1 of the present lemma that $\hat{A}_{XQ}^{(p)}(y) = 0$, which implies $\hat{Z}_{XQ}(y) = \hat{V}_{p,XQ}(y)$. Together with the equation in the previous display, we see that

$$\hat{Z}_{XQ}(y) = \hat{V}_{p,XQ}(y) = Q'\hat{V}_{p,X}(y). \tag{17}$$

We now want to choose $Q \in \mathbb{R}^{k \times k}$ (regular) such that

(i) every row vector of the matrix obtained from $\hat{Z}_{XQ}(y)$ by deleting the last column is nonzero, and
(ii) either every coordinate of $\omega'\hat{Z}_{XQ}(y)$ is non-negative, or every coordinate of $\omega'\hat{Z}_{XQ}(y)$ is non-positive,

holds, and that additionally the first columns of X and XQ, respectively, coincide. By assumption we either have $[\hat{V}_X(y)]_{1j_i} > 0$ for $i = 2, \ldots, t+1$, or we have $[\hat{V}_X(y)]_{1j_i} < 0$ for $i = 2, \ldots, t+1$. Consider the former (latter) case: Let

$$Q(\gamma) = \begin{pmatrix} 1 & \gamma & \gamma & \gamma & \ldots & \gamma \\ 0 & 1 & 0 & 0 & \ldots & 0 \\ 0 & 0 & 1 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \ldots & 1 \end{pmatrix},$$

which is a regular matrix for every $\gamma \in \mathbb{R}$. Post-multiplying X by $Q(\gamma)$ has the same effect as adding γ-times the first column of X to all other columns, without changing the first column. We observe that since $\hat{V}_{p,X}(y)$ is obtained from $\hat{V}_X(y)$ by deleting the first p columns, the nonzero columns of $\hat{V}_{p,X}(y)$ are precisely the vectors $[\hat{V}_X(y)]_{\cdot j_i}$ for $i = 2, \ldots, t+1$ because (y, X) satisfies Assumptions (A1) and (A2). Therefore, it is obvious from Equation (17), together with the assumed $[\hat{V}_X(y)]_{1j_i} > 0$ for $i = 2, \ldots, t+1$ ($[\hat{V}_X(y)]_{1j_i} < 0$ for $i = 2, \ldots, t+1$), that by choosing $\gamma^* > 0$ large enough, we can enforce that all nonzero columns of $\hat{Z}_{XQ(\gamma^*)}(y)$ are coordinate-wise positive (negative). Consider (i): Using Equation (17) we see that $Q(\gamma^*)'[\hat{V}_X(y)]_{\cdot j_2}$ is a column of the matrix obtained by deleting the last column of $\hat{Z}_{XQ(\gamma^*)}(y)$. This follows from $j_2 \geq p+2$ (a consequence of the assumptions $j_2 - j_1 \geq p+1$ and $j_1 = 1$), which shows that $Q(\gamma^*)'[\hat{V}_X(y)]_{\cdot j_2}$ is a column of $\hat{Z}_{XQ(\gamma^*)}(y)$, together with $j_2 < j_3 \leq n$ (a consequence of $k \geq 2$), which shows that it is not the last column of $\hat{Z}_{XQ(\gamma^*)}(y)$. Since all coordinates of $Q(\gamma^*)'[\hat{V}_X(y)]_{\cdot j_2}$ are positive (negative) by construction of $Q(\gamma^*)$, it follows that $Q(\gamma^*)$ satisfies (i) above. To show (ii) we recall that ω is nonzero and coordinate-wise nonnegative. Since all nonzero

columns of $\hat{Z}_{XQ(\gamma^*)}(y)$ are coordinate-wise positive (negative), it immediately follows that every coordinate of $\omega'\hat{Z}_{XQ(\gamma^*)}(y)$ is non-negative (non-positive). By construction the first columns of X and $XQ(\gamma^*)$ coincide. This proves the claim. $\qquad\square$

Proof of Lemma 3.11. We start with the first part. Let $X \in \mathfrak{X}_0 \subseteq \mathbb{R}^{n \times k}$ and $y \in \mathbb{R}^n$ be arbitrary but fixed. We show that $y \in N^*(\hat{\Omega}_{\kappa,M,p,X})$, which is equivalent to $g^*_{\kappa,M,p}(y, X, R) = 0$ by Part 5 of Lemma 3.10. If $y \in N(\hat{\Omega}_{\kappa,M,p,X}) \subseteq N^*(\hat{\Omega}_{\kappa,M,p,X})$ we are done. Suppose $y \notin N(\hat{\Omega}_{\kappa,M,p,X})$ which is equivalent to $g_{\kappa,M,p}(y, X) \neq 0$ by Lemma 3.9. We claim that $\mathrm{rank}(\hat{Z}_X(y)) < k$ must hold. Assuming this claim and using $\mathrm{rank}(R) = q = k$, $X \in \mathfrak{X}_0$ which implies $\mathrm{rank}(X) = k$, and $y \notin N(\hat{\Omega}_{\kappa,M,p,X})$ which implies $\mathrm{rank}(I_k - \sum_{l=1}^p \hat{A}_l^{(p)}(y)) = k$, it then follows from the definition of $B_{p,X}(y)$ in Equation (7) that $\mathrm{rank}(B_{p,X}(y)) < q$. As a consequence, Part 2 of Lemma 3.10 then shows that $\hat{\Omega}_{\kappa,M,p,X}(y)$ is singular, which implies $y \in N^*(\hat{\Omega}_{\kappa,M,p,X})$. To prove $\mathrm{rank}(\hat{Z}_X(y)) < k$ we note that

$$\hat{Z}_X(y) = \hat{V}_{p,X}(y)\left[I_{n-p} - \hat{V}'_{0,X}(y)\left(\hat{V}_{0,X}(y)\hat{V}'_{0,X}(y)\right)^{-1}\hat{V}_{0,X}(y)\right]$$
$$= \hat{V}_{p,X}(y)\Pi_{\mathrm{span}(\hat{V}'_{0,X}(y))^\perp}.$$

We see from the previous display that $\mathrm{rank}(\hat{Z}_X(y)) = k$, i.e., $\hat{Z}_X(y)$ having full row rank, is equivalent to $\mathrm{rank}(\hat{V}_{p,X}(y)) = k$ and $\mathrm{span}(\hat{V}'_{0,X}(y)) \cap \mathrm{span}(\hat{V}'_{p,X}(y)) = \{0\}$. Using $\mathrm{rank}(\hat{V}_{0,X}(y)) = kp$, a consequence of $y \notin N(\hat{\Omega}_{\kappa,M,p,X})$ (cf. Remark 3.1), this implies

$$\mathrm{rank}\left(\left(\hat{V}'_{0,X}(y) : \hat{V}'_{p,X}(y)\right)\right) = (p+1)k.$$

But this is impossible, because the matrix $(\hat{V}'_{0,X}(y) : \hat{V}'_{p,X}(y))$ is $(n-p) \times ((p+1)k)$ dimensional, which together with $n < (p+1)k + p$ implies $\mathrm{rank}(\hat{V}'_{0,X}(y) : \hat{V}'_{p,X}(y)) \leq n - p < (p+1)k$.

Next, we prove Part 2 of the lemma. Under the present assumptions it is shown in Part 1 of Proposition 4.5 that for $\lambda_{\mathbb{R}^{n \times k}}$ almost every $X \in \mathfrak{X}_0$ we have $g^*_{\kappa,M,p}(e_+, X, R) \neq 0$ and therefore in particular $g^*_{\kappa,M,p}(., X, R) \not\equiv 0$. This proves the first statement. To show the remaining statement we construct a y such that $g^*_{\kappa,M,p}(y, e_+, R) \neq 0$. Note first that $2p + 1 + \mathbf{1}_{\mathbb{M}_{AM}}(M) \leq n$ obviously implies $2p + 1 \leq n$. Let $y \in \mathbb{R}^n$ satisfy $y_1 = -1$, $y_{p+2} = 1$ and $y_i = 0$ else. [Note that this is feasible, i.e., $p + 2 \leq n$ holds, because of $2p + 1 \leq n$ and $p \geq 1$.] We intend to apply Part 2 of Lemma B.1 with $t = k = 1$ to the tuple (y, e_+). We first have to show that the tuple (y, e_+), which is clearly an element of $\mathbb{R}^n \times \mathfrak{X}_0$, satisfies Assumptions (A1), (A2) and (A3). For this we observe that $e_+ \perp y$, which implies $\hat{u}_{e_+}(y) = y$ and therefore

$$\hat{V}_{e_+}(y) = \hat{u}'_{e_+}(y) = y'.$$

Hence (A1) is satisfied, because $y_1 = -1 \neq 0$ and y has only two nonzero coordinates. The corresponding indices are $j_1 = 1$ and $j_{t+1} = p + 2$. The first part of Assumption (A2) is therefore obviously satisfied. The second part, i.e., $n - j_{t+1} = n - (p+2) \geq p-1$, follows immediately from $2p+1 \leq n$. Assumption (A3) follows from $y \neq 0$ together with the previous display and $t = k$. Therefore, $\hat{Z}_{e_+}(y) = \hat{V}_{p,e_+}(y) = (0,1,0,\ldots 0) \in \mathbb{R}^{n-p}$ follows as an application of Part 1 of Lemma B.1. Obviously (CKV) holds if $M \in \mathbb{M}_{KV}$. Since $\hat{Z}_{e_+}(y) = (0,1,0,\ldots 0)$ it is also obvious that (CNW) holds if $M \in \mathbb{M}_{NW}$. Suppose now that $M \in \mathbb{M}_{AM}$ holds. In this case $\mathbf{1}_{\mathbb{M}_{AM}}(M) = 1$ and therefore $2(p+1) \leq n$ holds. The latter implies $n - (p+2) = n - j_{t+1} > p - 1$. Consequently the statement in brackets in (CAM) shows that the condition is satisfied.

It remains to prove Part 3. Under the present assumptions it is shown in Part 2 of Proposition 4.5 that for $\lambda_{\mathbb{R}^{n \times (k-1)}}$ almost every $\tilde{X} \in \tilde{\mathfrak{X}}_0$ we have $g_{\kappa,M,p}^*(e_-, (e_+, \tilde{X}), R) \neq 0$ and therefore $g_{\kappa,M,p}^*(., (e_+, \tilde{X}), R) \not\equiv 0$. $\qquad \square$

Appendix C: Proofs of results in Section 4

For a definition of the group $G(\mathfrak{M}_0)$ appearing in the following lemma we refer the reader to Preinerstorfer and Pötscher (2016) Section 5.1.

Lemma C.1. *Assume that the triple κ, M, p satisfies Assumption 1. Assume further that $g_{\kappa,M,p}^*(., X, R) \not\equiv 0$. Then, $\hat{\beta}$ and $\hat{\Omega}_{\kappa,M,p}$ satisfy Assumptions 5, 6, and 7 in Preinerstorfer and Pötscher (2016) with $N = N(\hat{\Omega}_{\kappa,M,p})$. In fact, $\hat{\Omega}_{\kappa,M,p}(y)$ is nonnegative definite for every $y \in \mathbb{R}^n \backslash N(\hat{\Omega}_{\kappa,M,p})$, and is positive definite $\lambda_{\mathbb{R}^n}$-almost everywhere. The test statistic T defined in Equation (5), with $\hat{\Omega} = \hat{\Omega}_{\kappa,M,p}$, is invariant under the group $G(\mathfrak{M}_0)$ and the rejection probabilities $P_{\mu,\sigma^2\Sigma}(T \geq C)$ depend on $(\mu, \sigma^2, \Sigma) \in \mathfrak{M} \times (0,\infty) \times \mathfrak{C}$ only through $((R\beta - r)/\sigma, \Sigma)$ (in fact, only through $(\langle (R\beta - r)/\sigma \rangle, \Sigma)$), where β corresponds to μ via $\mu = X\beta$.*

Proof of Lemma C.1. The assumption $g_{\kappa,M,p}^*(., X, R) \not\equiv 0$ together with Part 5 of Lemma 3.10 implies that the algebraic set $N^*(\hat{\Omega}_{\kappa,M,p})$ is a closed $\lambda_{\mathbb{R}^n}$-null set. Therefore, by Lemma 3.9, it follows that the algebraic set $N(\hat{\Omega}_{\kappa,M,p}) \subseteq N^*(\hat{\Omega}_{\kappa,M,p})$ is a closed $\lambda_{\mathbb{R}^n}$-null set as well. We claim that $\hat{\Omega}_{\kappa,M,p}$ is continuous (and obviously well defined by definition) on $\mathbb{R}^n \backslash N(\hat{\Omega}_{\kappa,M,p})$, because it can be written as a composition of continuous functions on this set: We recall from Equation (13) the representation

$$\hat{\Omega}_{\kappa,M,p}(y) = \frac{n}{n-p} B_p(y) \mathcal{W}_{n-p}(y) B_p'(y) \text{ for every } y \in \mathbb{R}^n \backslash N(\hat{\Omega}_{\kappa,M,p}).$$

We first observe that $B_p(.)$ (which was defined in Equation (7)) is continuous on $\mathbb{R}^n \backslash N(\hat{\Omega}_{\kappa,M,p})$, because $\hat{V}(.)$ and hence $\hat{A}^{(p)}(.)$ and $\hat{Z}(.)$ are continuous on $\mathbb{R}^n \backslash N(\hat{\Omega}_{\kappa,M,p})$. By considering each of the cases $M \in \mathbb{M}_{KV}$, $M \in \mathbb{M}_{NW}$ and $M \in \mathbb{M}_{AM}$ separately, it is easy to see that $M(.)$ is continuous on $\mathbb{R}^n \backslash N(\hat{\Omega}_{\kappa,M,p})$. The main diagonal entries of $\mathcal{W}_n(.)$ are by definition constant on $\mathbb{R}^n \backslash N(\hat{\Omega}_{\kappa,M,p})$.

Therefore, it remains to show that all off-diagonal entries are continuous on $\mathbb{R}^n \backslash N(\hat{\Omega}_{\kappa,M,p})$. Each of them is of the form $\kappa(i/M(y))$ for some fixed $|i| = 1, \ldots, n{-}p{-}1$ if $M(y) \neq 0$, and 0 if $M(y) = 0$. Since κ is a continuous function by Assumption 1, and $M(.)$ is continuous on $\mathbb{R}^n \backslash N(\hat{\Omega}_{\kappa,M,p})$ and satisfies $M(y) \geq 0$, it remains to check that $\kappa(x) \to 0$ as $|x| \to \infty$, which is a part of Assumption 1. This proves the claim. Since $\hat{\beta}$ is well defined and continuous everywhere on $\mathbb{R}^n$, it follows that both $\hat{\beta}$ and $\hat{\Omega}_{\kappa,M,p}$ are well-defined and continuous on $\mathbb{R}^n \backslash N(\hat{\Omega}_{\kappa,M,p})$. Clearly, $\hat{\Omega}_{\kappa,M,p}$ is symmetric on $\mathbb{R}^n \backslash N(\hat{\Omega}_{\kappa,M,p})$. This proves Part (i) of Assumption 5 in Preinerstorfer and Pötscher (2016). To prove the second part let $y \in \mathbb{R}^n \backslash N(\hat{\Omega}_{\kappa,M,p})$, $\alpha \neq 0$ and $\gamma \in \mathbb{R}^k$. We have to show that $\alpha y + X\gamma \in \mathbb{R}^n \backslash N(\hat{\Omega}_{\kappa,M,p})$. Note that $\hat{V}(\alpha y + X\gamma) = X' \operatorname{diag}(\hat{u}(\alpha y + X\gamma)) = \alpha \hat{V}(\hat{u}(y))$, which implies $\hat{A}^{(p)}(\alpha y + X\gamma) = \hat{A}^{(p)}(\hat{u}(y))$ and $\hat{Z}(\alpha y + X\gamma) = \alpha \hat{Z}(y)$. The latter immediately leads (considering each of the cases $M \in \mathbb{M}_{KV}$, $M \in \mathbb{M}_{NW}$ and $M \in \mathbb{M}_{AM}$ separately) to $M(\alpha y + X\gamma) = M(y)$, which in turn implies $\mathcal{W}_{n-p}(\alpha y + X\gamma) = \mathcal{W}_{n-p}(y)$. It then follows from the previous display and the definition of $B_p(y)$ that $\hat{\Omega}_{\kappa,M,p}(\alpha y + X\gamma) = \alpha^2 \hat{\Omega}_{\kappa,M,p}(y)$. Therefore, we clearly have $\alpha y + X\gamma \in \mathbb{R}^n \backslash N(\hat{\Omega}_{\kappa,M,p})$, which proves Part (ii) of Assumption 5 in Preinerstorfer and Pötscher (2016), and where we have also established the equivariance property of $\hat{\Omega}_{\kappa,M,p}$ required in Part (iii) of Assumption 5 in Preinerstorfer and Pötscher (2016). That $\hat{\beta}$ satisfies the equivariance property in Part (iii) of Assumption 5 in Preinerstorfer and Pötscher (2016) is obvious. It remains to show that $\hat{\Omega}_{\kappa,M,p}$ is $\lambda_{\mathbb{R}^n}$-almost everywhere nonsingular on $\mathbb{R}^n \backslash N(\hat{\Omega}_{\kappa,M,p})$. This is equivalent to

$$\left\{ y \in \mathbb{R}^n \backslash N(\hat{\Omega}_{\kappa,M,p}) : \det(\hat{\Omega}_{\kappa,M,p}(y)) = 0 \right\} = N^*(\hat{\Omega}_{\kappa,M,p}) \backslash N(\hat{\Omega}_{\kappa,M,p})$$

being a $\lambda_{\mathbb{R}^n}$-null set. This is obvious, since we have already observed that $N^*(\hat{\Omega}_{\kappa,M,p})$ is a $\lambda_{\mathbb{R}^n}$-null set under the maintained assumptions. This proves the claim concerning Assumption 5. That $\hat{\Omega}_{\kappa,M,p}(y)$ is nonnegative definite for every $y \notin N(\hat{\Omega}_{\kappa,M,p})$ (which is equivalent to $g_{\kappa,M,p}(y,X) \neq 0$ by Lemma 3.9) has been shown in Part 1 of Lemma 3.10. It follows that $\hat{\Omega}_{\kappa,M,p}$ is positive definite on the complement of $N^*(\hat{\Omega}_{\kappa,M,p})$. Hence $\hat{\Omega}_{\kappa,M,p}$ is $\lambda_{\mathbb{R}^n}$- almost everywhere positive definite. This immediately shows that Assumptions 6 and 7 in Preinerstorfer and Pötscher (2016) are satisfied. The remaining two claims in the lemma now follow immediately from what has been established together with Lemma 5.15 Part 3 and Proposition 5.4 in Preinerstorfer and Pötscher (2016). $\square$

Proof of Theorem 4.2. In each part of the theorem we have $g^*_{\kappa,M,p}(.,X,R) \not\equiv 0$. In Part 4 this is an explicit assumption. In the other parts this is implied by the assumption that $g^*_{\kappa,M,p}(.,X,R)$ does not vanish at a specific point. As a consequence Lemma C.1 is applicable in all parts of the theorem. We shall now apply the first two parts of Corollary 5.17 in Preinerstorfer and Pötscher (2016) to prove the first two parts of the present theorem. Lemma C.1 shows that $\hat{\beta}$ and $\hat{\Omega}_{\kappa,M,p}$ satisfy Assumption 5 in Preinerstorfer and Pötscher (2016) with $N = N(\hat{\Omega}_{\kappa,M,p})$. Furthermore, note that the set N^* figuring Corollary 5.17 of

Preinerstorfer and Pötscher (2016) coincides with $N^*(\hat{\Omega}_{\kappa,M,p})$. By Assumption 2 the spaces $\mathcal{Z}_+ = \mathrm{span}(e_+)$ and $\mathcal{Z}_- = \mathrm{span}(e_-)$ are concentration spaces of $\mathfrak{C}$ (cf. Lemma G.1 in Preinerstorfer and Pötscher (2016)). Hence, Parts 1 and 2 of the present theorem now follow by applying the first two parts of Corollary 5.17 in Preinerstorfer and Pötscher (2016) and Remark 5.18(i) in Preinerstorfer and Pötscher (2016) to $\mathcal{Z}_+$ as well as to $\mathcal{Z}_-$, and by noting that Part 5 of Lemma 3.10 shows that the statement $e_+ \in \mathbb{R}^n \backslash N^*(\hat{\Omega}_{\kappa,M,p})$ translates into $g^*_{\kappa,M,p}(e_+, X, R) \neq 0$, with a similar translation if e_+ is replaced by e_-. That the test is biased in Part 2 of the theorem follows immediately from Part 5 of Lemma 5.15 in Preinerstorfer and Pötscher (2016) (note that Assumptions 5 and 6 in Preinerstorfer and Pötscher (2016) are satisfied by Lemma C.1) showing that $W(C)$ contains a non-empty open set. To prove Part 4 we apply Theorem 5.19 in Preinerstorfer and Pötscher (2016). Lemma C.1 shows that $\hat{\beta}$ and $\hat{\Omega}_{\kappa,M,p}$ also satisfy Assumption 7 in Preinerstorfer and Pötscher (2016). We consider the case where $e_+ \in \mathfrak{M}$ and $R\hat{\beta}(e_+) \neq 0$. The other case can be handled analogously. From Remark 5.20 in Preinerstorfer and Pötscher (2016) we see that all conditions on the covariance model in Theorem 5.19 in Preinerstorfer and Pötscher (2016) are satisfied with $\bar{\Sigma} = e_+ e'_+$, $\mathrm{span}(\bar{\Sigma}) = \mathrm{span}(e_+)$ and $Z = e_+$. Clearly, $\mathrm{span}(\bar{\Sigma}) = \mathrm{span}(e_+) \subseteq \mathfrak{M}$ and $R\hat{\beta}(z) \neq 0$ holds $\lambda_{\mathrm{span}(\bar{\Sigma})}$-a.e. This shows that Equation (33) in Preinerstorfer and Pötscher (2016) holds in the present setup. To conclude, it remains to observe that K_2 in this equation equals one. This follows from the discussion preceding Theorem 5.19 in Preinerstorfer and Pötscher (2016), because $\hat{\Omega}_{\kappa,M,p}$ is almost everywhere positive definite by Lemma C.1.

Now we consider Part 3 of the theorem. We prove the case where $g^*_{\kappa,M,p}(e_+, X, R) \neq 0$, $T(e_+ + \mu_0^*) = C$ and $\mathrm{grad}\, T(e_+ + \mu_0^*)$ exists for some $\mu_0^* \in \mathfrak{M}_0$. The other case works analogously. The statement in the theorem saying that if $\mathrm{grad}\, T(e_+ + \mu_0^*)$ exists and $T(e_+ + \mu_0^*) = C$ holds for *some* $\mu_0^* \in \mathfrak{M}_0$, then $\mathrm{grad}\, T(e_+ + \mu_0^*)$ exists and $T(e_+ + \mu_0^*) = C$ holds for *all* $\mu_0^* \in \mathfrak{M}_0$ follows at once from invariance of T w.r.t. $G(\mathfrak{M}_0)$, which holds as a consequence of Lemma C.1. In a first step we now show that the linear functional on $\mathbb{R}^n$ corresponding to the row vector $\mathrm{grad}\, T(\mu_0^* + e_+)$ does not vanish everywhere on $\mathrm{span}(e_+)^\perp$: Arguing by contradiction, assume that $\mathrm{grad}\, T(\mu_0^* + e_+)w = 0$ for every $w \in \mathrm{span}(e_+)^\perp$, which is equivalent to $\mathrm{grad}\, T(\mu_0^* + e_+)' \perp \mathrm{span}(e_+)^\perp$ and therefore $\mathrm{grad}\, T(\mu_0^* + e_+)' \in \mathrm{span}(e_+)$ holds, i.e., $\mathrm{grad}\, T(\mu_0^* + e_+) = ce'_+$ for some $c \in \mathbb{R}$. Since T is $G(\mathfrak{M}_0)$ invariant, it holds for every $\gamma \neq 0$ that

$$T(\gamma e_+ + \mu_0^*) = T(\gamma(e_+ + \mu_0^* - \mu_0^*) + \mu_0^*) = T(e_+ + \mu_0^*) = C.$$

Hence on the set $\mathbb{R} \backslash \{-1\}$ the mapping

$$\alpha \mapsto T(e_+ + \mu_0^* + \alpha e_+) = T((1+\alpha)e_+ + \mu_0^*) = C$$

is constant, thus showing that the directional derivative of T at the point $e_+ + \mu_0^*$ in direction e_+ is zero. The latter is equivalent to $\mathrm{grad}\, T(\mu_0^* + e_+)e_+ = c\|e_+\|^2 = 0$, and hence $c = 0$ holds which implies $\mathrm{grad}\, T(\mu_0^* + e_+) = 0$. To arrive at a contradiction it remains to show that there is a vector v such that

the directional derivative of T at $e_+ + \mu_0$ in direction v does not vanish. To this end, recall that Assumption 5 in Preinerstorfer and Pötscher (2016) is satisfied, hence the discussion following that Assumption in Preinerstorfer and Pötscher (2016) shows that $N^*(\hat{\Omega}_{\kappa,M,p})$ is invariant w.r.t. $G(\mathfrak{M})$. Therefore, $e_+ \notin N^*(\hat{\Omega}_{\kappa,M,p})$ implies $e_+ + \mu_0^* \notin N^*(\hat{\Omega}_{\kappa,M,p})$. Since $N^*(\hat{\Omega}_{\kappa,M,p})$ is closed by Lemma 3.10, there exists an open ball U_ε of radius $\varepsilon > 0$ centered at $e_+ + \mu_0^*$ such that $U_\varepsilon \subseteq \mathbb{R}^n \backslash N^*(\hat{\Omega}_{\kappa,M,p})$. Additionally, we note that $e_+ \notin \mathfrak{M}$, because $\mathfrak{M} \subseteq N^*(\hat{\Omega}_{\kappa,M,p})$ always holds (see the discussion in Preinerstorfer and Pötscher (2016) after Assumption 5). Therefore, $v = \Pi_{\mathfrak{M}^\perp} e_+ / \|\Pi_{\mathfrak{M}^\perp} e_+\|$ is well defined and for $0 \leq |\alpha| < \varepsilon$ we have $e_+ + \mu_0^* + \alpha v \in U_\varepsilon$. Assume that $0 \leq |\alpha| < \varepsilon$. The OLS estimator $\hat{\beta}$ clearly satisfies

$$R\hat{\beta}(e_+ + \mu_0^* + \alpha v) = R\hat{\beta}(e_+ + \mu_0^*) + \alpha R\hat{\beta}(v) = R\hat{\beta}(e_+ + \mu_0^*),$$

where the second equality follows from $v \perp \mathfrak{M}$. Since $\hat{\Omega}_{\kappa,M,p}$ satisfies the equivariance condition in Assumption 5 of Preinerstorfer and Pötscher (2016) we can furthermore write

$$
\begin{aligned}
\hat{\Omega}_{\kappa,M,p}(e_+ + \mu_0^* + \alpha v) &= \hat{\Omega}_{\kappa,M,p}(e_+ + \alpha v) \\
&= \hat{\Omega}_{\kappa,M,p}(e_+ - \Pi_{\mathfrak{M}} e_+ + \alpha v) \\
&= \hat{\Omega}_{\kappa,M,p}(\Pi_{\mathfrak{M}^\perp} e_+ + \alpha \Pi_{\mathfrak{M}^\perp} e_+ / \|\Pi_{\mathfrak{M}^\perp} e_+\|) \\
&= \hat{\Omega}_{\kappa,M,p}\left((1 + \frac{\alpha}{\|\Pi_{\mathfrak{M}^\perp} e_+\|}) \Pi_{\mathfrak{M}^\perp} e_+ \right) \\
&= (1 + \frac{\alpha}{\|\Pi_{\mathfrak{M}^\perp} e_+\|})^2 \hat{\Omega}_{\kappa,M,p}\left(\Pi_{\mathfrak{M}^\perp} e_+ \right) \\
&= (1 + \frac{\alpha}{\|\Pi_{\mathfrak{M}^\perp} e_+\|})^2 \hat{\Omega}_{\kappa,M,p}\left(e_+ + \mu_0^* \right).
\end{aligned}
$$

By definition of T (cf. Equation (5) and recall that $e_+ + \mu_0^* + \alpha v \in U_\varepsilon \subseteq \mathbb{R}^n \backslash N^*(\hat{\Omega}_{\kappa,M,p})$) and the assumed equality $T(e_+ + \mu_0^*) = C$, the relations derived above allow us to show that

$$
\begin{aligned}
&T(e_+ + \mu_0^* + \alpha v) \\
&= \left(R\hat{\beta}(e_+ + \mu_0^* + \alpha v) - r \right)' \hat{\Omega}_{\kappa,M,p}^{-1}(e_+ + \mu_0^* + \alpha v) \left(R\hat{\beta}(e_+ + \mu_0^* + \alpha v) - r \right) \\
&= (1 + \frac{\alpha}{\|\Pi_{\mathfrak{M}^\perp} e_+\|})^{-2} \left(R\hat{\beta}(e_+ + \mu_0^*) - r \right)' \hat{\Omega}_{\kappa,M,p}^{-1}(e_+ + \mu_0^*) \left(R\hat{\beta}(e_+ + \mu_0^*) - r \right) \\
\\
&= (1 + \frac{\alpha}{\|\Pi_{\mathfrak{M}^\perp} e_+\|})^{-2} T(e_+ + \mu_0^*) \\
&= (1 + \frac{\alpha}{\|\Pi_{\mathfrak{M}^\perp} e_+\|})^{-2} C
\end{aligned}
$$

holds for every $0 \leq |\alpha| < \varepsilon$. This implies that the directional derivative of T in direction v at the point $e_+ + \mu_0^*$ equals $-2C/\|\Pi_{\mathfrak{M}^\perp} e_+\|$, which is nonzero as a

consequence of $C > 0$. In a second step we shall now derive an expansion of T at points of the form $y + \mu_0^*$ for y satisfying $e_+'y \neq 0$: For every $h \in \mathbb{R}^n$ we have

$$T(e_+ + \mu_0^* + h) = T(e_+ + \mu_0^*) + \operatorname{grad} T(e_+ + \mu_0^*)h + Q(h) \qquad (18)$$

where $Q(h)/\|h\| \to 0$ as $h \to 0$ and $h \neq 0$. Recall that T is invariant under the group $G(\mathfrak{M}_0)$. In particular for every y such that $e_+'y \neq 0$ we have

$$T(y + \mu_0^*) = T(\frac{e_+'y}{n}e_+ + \mu_0^* + \Pi_{\operatorname{span}(e_+)^\perp}y) = T(e_+ + \mu_0^* + \frac{n}{e_+'y}\Pi_{\operatorname{span}(e_+)^\perp}y),$$

where the first equality holds because of $y = \Pi_{\operatorname{span}(e_+)}y + \Pi_{\operatorname{span}(e_+)^\perp}y$ and the second follows from invariance of T w.r.t. $G(\mathfrak{M}_0)$. This means that whenever $e_+'y \neq 0$ holds, we can combine the equation in the previous display and Equation (18) with $h = \frac{n}{e_+'y}\Pi_{\operatorname{span}(e_+)^\perp}y$ to see that

$$T(y+\mu_0^*) = T(e_+ +\mu_0^*)+\frac{n}{e_+'y}\operatorname{grad} T(e_+ +\mu_0^*)\Pi_{\operatorname{span}(e_+)^\perp}y+Q(\frac{n}{e_+'y}\Pi_{\operatorname{span}(e_+)^\perp}y) \tag{19}$$

holds and that

$$Q(\frac{n}{e_+'y_m}\Pi_{\operatorname{span}(e_+)^\perp}y_m)/\|\frac{n}{e_+'y_m}\Pi_{\operatorname{span}(e_+)^\perp}y_m\| \to 0, \qquad (20)$$

for any sequence y_m satisfying $e_+'y_m \neq 0$, $\frac{n}{e_+'y_m}\Pi_{\operatorname{span}(e_+)^\perp}y_m \to 0$ and $\frac{n}{e_+'y_m}\Pi_{\operatorname{span}(e_+)^\perp}y_m \neq 0$. Now, we choose a sequence $\rho_m \in (-1,1)$ such that $\rho_m \to 1$ and apply Assumption 2 to obtain $\Lambda(\rho_m) \in \mathfrak{C}$ for every m. We intend to show that $P_{\mu_0^*,\Lambda(\rho_m)}(W(C)) \to 1/2$ along a subsequence. The last statement in Lemma C.1 then implies $P_{\mu_0,\sigma^2\Lambda(\rho_m)}(W(C)) \to 1/2$ for every pair $\mu_0 \in \mathfrak{M}_0$ and $0 < \sigma^2 < \infty$ along this subsequence. In Part 3 of Lemma G.1 in Preinerstorfer and Pötscher (2016) it is shown that $\Lambda(\rho_m) \to e_+e_+'$, $D_m = \Pi_{\operatorname{span}(e_+)^\perp}\Lambda(\rho_m)\Pi_{\operatorname{span}(e_+)^\perp}/s_m \to D$, where s_m is a sequence of numbers such that $s_m > 0$, $s_m \to 0$ and D is regular on $\operatorname{span}(e_+)^\perp$. Furthermore, it is shown that $\Pi_{\operatorname{span}(e_+)^\perp}\Lambda(\rho_m)\Pi_{\operatorname{span}(e_+)}/s_m^{1/2} \to 0$. We can use these relations to derive three useful facts: (i) Observe that the matrix $s_m^{-1/2}\Pi_{\operatorname{span}(e_+)^\perp}\Lambda(\rho_m)^{1/2}$ is an $n \times n$-dimensional nonnegative square root of the symmetric matrix D_m. Therefore, we can find an orthogonal matrix U_m such that

$$s_m^{-1/2}\Pi_{\operatorname{span}(e_+)^\perp}\Lambda(\rho_m)^{1/2}U_m = D_m^{1/2}$$

holds. The sequence $D_m^{1/2}$ converges to $D^{1/2}$ as a consequence of $D_m \to D$ together with the continuity of taking the nonnegative definite symmetric matrix square root of a symmetric and nonnegative definite matrix. Since U_m is orthogonal we can choose a subsequence m' along which U_m converges to U, say. Without loss of generality we henceforth assume $m' \equiv m$. Using the relation in the previous display we find that $s_m^{-1/2}\Pi_{\operatorname{span}(e_+)^\perp}\Lambda(\rho_m)^{1/2}$ converges to

$$D^* = D^{1/2}U', \tag{21}$$

and we recall from above that $D^{1/2} \in \mathbb{R}^{n \times n}$ is regular on $\mathrm{span}(e_+)^\perp$. (ii) We note that $\Lambda(\rho_m) \to e_+ e_+'$ implies

$$\Lambda(\rho_m)^{1/2} \to n^{-1/2} e_+ e_+'.$$

(iii) We show that $D^* e_+ = 0$ holds: Note that $\Pi_{\mathrm{span}(e_+)^\perp} \Lambda(\rho_m) \Pi_{\mathrm{span}(e_+)} / s_m^{1/2} \to 0$ can be rewritten as

$$\left(s_m^{-1/2} \Pi_{\mathrm{span}(e_+)^\perp} \Lambda(\rho_m)^{1/2} \right) \Lambda(\rho_m)^{1/2} \Pi_{\mathrm{span}(e_+)} \to 0,$$

that the term in brackets converges to D^* by (i) and that the other term converges to $n^{-1/2} e_+ e_+'$ by (ii). Therefore,

$$D^* n^{-1/2} e_+ e_+' = 0,$$

or equivalently $D^* \Pi_{\mathrm{span}(e_+)} = 0$. But this implies $D^* e_+ = 0$. Now, we are ready to show that $P_{\mu_0^*, \Lambda(\rho_m)}(W(C)) \to 1/2$. Let $\mathbf{G}$ be a random n-vector defined on some underlying probability space such that the probability measure induced by $\mathbf{G}$ on $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ equals P_{0, I_n}. Consequently, the random vector $\Lambda(\rho_m)^{1/2} \mathbf{G} + \mu_0^*$ induces the distribution $P_{\mu_0^*, \Lambda(\rho_m)}$ on $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$. For notational convenience we write $\mathbf{G}_m = \Lambda(\rho_m)^{1/2} \mathbf{G}$. Consequently, we have

$$\begin{aligned}
P_{\mu_0^*, \Lambda(\rho_m)}\left(W(C)\right) &= \Pr(T(\mathbf{G}_m + \mu_0^*) \geq C) \qquad (22) \\
&= \Pr\left(s_m^{-1/2} \left[T(\mathbf{G}_m + \mu_0^*) - T(e_+ + \mu_0^*) \right] \geq 0 \right),
\end{aligned}$$

where we used $s_m^{-1/2} > 0$ and $T(e_+ + \mu_0^*) = C$ in deriving the second equality. Note that $e_+' \mathbf{G}_m \neq 0$ and $\|\frac{n}{e_+' \mathbf{G}_m} \Pi_{\mathrm{span}(e_+)^\perp} \mathbf{G}_m\| > 0$ on an event of probability one. Using the expansion developed in Equation (19) we see that with probability one $s_m^{-1/2} \left[T(\mathbf{G}_m + \mu_0^*) - T(e_+ + \mu_0^*) \right]$ can be written as

$$s_m^{-1/2} \left[\frac{n}{e_+' \mathbf{G}_m} \mathrm{grad}\, T(e_+ + \mu_0^*) \Pi_{\mathrm{span}(e_+)^\perp} \mathbf{G}_m + Q\left(\frac{n}{e_+' \mathbf{G}_m} \Pi_{\mathrm{span}(e_+)^\perp} \mathbf{G}_m \right) \right]$$

$$= \frac{n}{e_+' \mathbf{G}_m} \mathrm{grad}\, T(e_+ + \mu_0^*) s_m^{-1/2} \Pi_{\mathrm{span}(e_+)^\perp} \mathbf{G}_m + \| \frac{n}{e_+' \mathbf{G}_m} s_m^{-1/2} \Pi_{\mathrm{span}(e_+)^\perp} \mathbf{G}_m \|$$

$$\times \quad \| \left(\frac{n}{e_+' \mathbf{G}_m} \Pi_{\mathrm{span}(e_+)^\perp} \mathbf{G}_m \right) \|^{-1} Q\left(\frac{n}{e_+' \mathbf{G}_m} \Pi_{\mathrm{span}(e_+)^\perp} \mathbf{G}_m \right).$$

To derive the almost sure limit as $m \to \infty$ of the expression in the previous display we first observe that $\mathbf{G}_m$ converges point-wise to $n^{-1/2} e_+ e_+' \mathbf{G}$ because of (ii). From that it follows that $e_+' \mathbf{G}_m$ converges point-wise to $\sqrt{n} e_+' \mathbf{G}$ and that $\Pi_{\mathrm{span}(e_+)^\perp} \mathbf{G}_m$ converges point-wise to zero. An application of the continuous mapping theorem hence shows that

$$\frac{n}{e_+' \mathbf{G}_m} \Pi_{\mathrm{span}(e_+)^\perp} \mathbf{G}_m \to 0$$

almost surely as $m \to \infty$, which immediately implies

$$\|\left(\frac{n}{e_+'\mathbf{G}_m}\Pi_{\mathrm{span}(e_+)^\perp}\mathbf{G}_m\right)\|^{-1}Q\left(\frac{n}{e_+'\mathbf{G}_m}\Pi_{\mathrm{span}(e_+)^\perp}\mathbf{G}_m\right) \to 0$$

almost surely as $m \to \infty$ as a consequence of Equation (20) together with $Q(0) = 0$. We also observe that (i) above implies

$$\Pi_{\mathrm{span}(e_+)^\perp}s_m^{-1/2}\mathbf{G}_m \to D^*\mathbf{G}$$

point-wise and thus, using the continuous mapping theorem again, we see that

$$\frac{n}{e_+'\mathbf{G}_m}\Pi_{\mathrm{span}(e_+)^\perp}s_m^{-1/2}\mathbf{G}_m \to \frac{\sqrt{n}}{e_+'\mathbf{G}}D^*\mathbf{G}$$

almost surely as $m \to \infty$ (where the limiting random vector is well defined almost-surely). This finally shows that

$$s_m^{-1/2}\left[T(\mathbf{G}_m + \mu_0^*) - T(e_+ + \mu_0^*)\right] \to \frac{\sqrt{n}}{e_+'\mathbf{G}}\operatorname{grad}T(e_+ + \mu_0^*)D^*\mathbf{G},$$

almost surely. We already know from Equation (21) that $D^* = D^{1/2}U'$, where U is an orthogonal matrix. Furthermore $D^{1/2}$ maps $\mathbb{R}^n$ onto $\mathrm{span}(e_+)^\perp$, and $\operatorname{grad}T(e_+ + \mu_0^*)$ does not vanish everywhere on $\mathrm{span}(e_+)^\perp$. Hence, we see that the probability that the limiting random variable in the previous display takes on the value 0 vanishes because $\operatorname{grad}T(e_+ + \mu_0^*)D^*\mathbf{G}$ is a Gaussian random variable with mean zero and positive variance. Hence, Equation (22) together with Portmanteau theorem shows that

$$P_{\mu_0^*,\Lambda(\rho_m)}(W(C)) \to \Pr(\frac{\sqrt{n}}{e_+'\mathbf{G}}\operatorname{grad}T(e_+ + \mu_0^*)D^*\mathbf{G} \geq 0). \qquad (23)$$

The covariance between the Gaussian mean-zero random variables $\operatorname{grad}T(e_+ + \mu_0^*)D^*\mathbf{G}$ and $e_+'\mathbf{G}$ is given by

$$\operatorname{grad}T(e_+ + \mu_0^*)D^*e_+ = 0,$$

where the equality follows from (iii). Therefore, $e_+'\mathbf{G}$ and $\operatorname{grad}T(e_+ + \mu_0^*)D^*\mathbf{G}$ are independent. Since the probability to the right in Equation (23) equals the probability that the random variables $e_+'\mathbf{G}$ and $\operatorname{grad}T(e_+ + \mu_0^*)D^*\mathbf{G}$ have the same sign it is now obvious that the limit equals $1/2$. $\qquad\square$

Lemma C.2. *Assume that the triple κ, M, p satisfies Assumption 1 and let T be as in Equation (5) with $\hat{\Omega} = \hat{\Omega}_{\kappa,M,p}$. Let $\mu_0 \in \mathfrak{M}_0$. Then the following holds.*

1. *If $M \in \mathbb{M}_{KV}$ and $g_{\kappa,M,p}^*(y, X, R) \neq 0$, then $\operatorname{grad}T(\mu_0 + y)$ exists.*
2. *Suppose that $M \notin \mathbb{M}_{KV}$, that κ is continuously differentiable on the non-void complement of a closed set $\Delta(\kappa) \subseteq \mathbb{R}$, and that $g_{\kappa,M,p}^*(y, X, R) \neq 0$. Assume further that one of the following conditions is satisfied:*

(a) $M(y) \neq 0$ and $\frac{i}{M(y)} \notin \Delta(\kappa)$ for $|i| = 1, \ldots, n - p - 1$.

(b) $M(y) = 0$ and κ has compact support.

Then $\operatorname{grad} T(\mu_0 + y)$ exists.

3. If $M \notin \mathbb{M}_{KV}$, then for every $\delta \geq 0$ we have

$$\left\{ y \in \mathbb{R}^n : g_{\kappa,M,p}^*(y, X, R) \neq 0 \text{ and } M(y) = \delta \right\}$$
$$= \left\{ y \in \mathbb{R}^n : g_{\kappa,M,p}^*(y, X, R) \neq 0 \text{ and } g_{\kappa,M,p}^{(\delta)}(y, X) = 0 \right\},$$

where $g_{\kappa,M,p}^{(\delta)} : \mathbb{R}^n \times \mathbb{R}^{n \times k} \to \mathbb{R}$ is a multivariate polynomial (explicitly constructed in the proof) that does not depend on the hypothesis (R, r).

Proof. We first verify Parts 1 and 2. Let us start by deriving a convenient expression for $T(\mu_0 + y)$ under the assumption $g_{\kappa,M,p}^*(y, X, R) \neq 0$. By Lemma 3.10 the assumption $g_{\kappa,M,p}^*(y, X, R) \neq 0$ is equivalent to $y \notin N^*(\hat{\Omega}_{\kappa,M,p})$. An application of Lemma C.1 shows that $\hat{\Omega}_{\kappa,M,p}$ satisfies Assumption 5 in Preinerstorfer and Pötscher (2016). An application of Part (ii) of this Assumption shows that $\mu_0 + y \notin N^*(\hat{\Omega}_{\kappa,M,p})$. We can therefore use Equation (5) together with $R\hat{\beta}(y + \mu_0) - r = R\hat{\beta}(y)$ and $\hat{\Omega}_{\kappa,M,p}(\mu_0 + y) = \hat{\Omega}_{\kappa,M,p}(y)$ (both following from Assumption 5 in Preinerstorfer and Pötscher (2016)) to see that $y \notin N^*(\hat{\Omega}_{\kappa,M,p})$ implies

$$\begin{aligned} T(\mu_0 + y) = & \hat{\beta}(y)' R' \hat{\Omega}_{\kappa,M,p}^{-1}(y) R\hat{\beta}(y) \\ = & \hat{\beta}(y)' R' \left(\frac{n}{n-p} B_p(y) \mathcal{W}_{n-p}(y) B_p'(y) \right)^{-1} R\hat{\beta}(y), \end{aligned}$$

where in deriving the second equality we made use of the representation of $\hat{\Omega}_{\kappa,M,p}(y)$ developed in Equation (13). The function $\hat{\beta}(.)$ is linear and hence totally differentiable on $\mathbb{R}^n$. Furthermore, in the proof of Lemma 3.10 it is shown that the coordinates of the matrix $B_p(.)$ are multivariate rational functions (without singularities) on $\mathbb{R}^n \backslash N^*(\hat{\Omega}_{\kappa,M,p})$. In particular the coordinates of $B_p(.)$ are continuously partially differentiable on $\mathbb{R}^n \backslash N^*(\hat{\Omega}_{\kappa,M,p})$. To show that $\operatorname{grad} T(\mu_0 + y)$ exists at a given point $y \in \mathbb{R}^n \backslash N^*(\hat{\Omega}_{\kappa,M,p})$ it is therefore sufficient to show that each off-diagonal element (recall that the diagonal is constant) of $\mathcal{W}_{n-p}(.)$ is continuously partially differentiable on an open neighborhood of y in $\mathbb{R}^n \backslash N^*(\hat{\Omega}_{\kappa,M,p})$. Recall that the i-th off-diagonal element $(i \in \{1, \ldots, n - p - 1\})$ of $\mathcal{W}_{n-p}(.)$ evaluated at some $y \in \mathbb{R}^n \backslash N^*(\hat{\Omega}_{\kappa,M,p})$ is given by

$$f_i(y) := \begin{cases} \kappa(i/M(y)) & \text{if } M(y) \neq 0 \\ 0 & \text{else.} \end{cases}$$

If $M \in \mathbb{M}_{KV}$ the sufficient condition above is obviously satisfied, because in this case $M > 0$ is constant and therefore $f_i(.)$ is constant on $\mathbb{R}^n \backslash N^*(\hat{\Omega}_{\kappa,M,p})$. This proves Part 1 of the lemma. Consider now Part 2. By considering separately the cases $M \in \mathbb{M}_{AM}$ and $M \in \mathbb{M}_{NW}$, we observe that $M(.)$ is continuously partially

differentiable in an open neighborhood of any element y of $\mathbb{R}^n \backslash N^*(\hat{\Omega}_{\kappa,M,p})$ satisfying $M(y) \neq 0$. We start with Condition (a). Let y satisfy $y \notin N^*(\hat{\Omega}_{\kappa,M,p})$ and $M(y) \neq 0$. Fix an $i \in \{1,\dots,n-p-1\}$. By assumption κ is continuously differentiable on an open neighborhood of $i/M(y) \notin \Delta(\kappa)$. Hence there exists an open neighborhood U of y in $\mathbb{R}^n \backslash N^*(\hat{\Omega}_{\kappa,M,p})$ on which $M(.)$ is strictly greater than zero and such that $\kappa(i/M(.))$ is continuously partially differentiable on U. It hence follows that $f_i(.)$ is continuously partially differentiable on U because it coincides with $\kappa(i/M(.))$ on this set. To establish existence of the gradient under Condition (b) let y satisfy $y \notin N^*(\hat{\Omega}_{\kappa,M,p})$ and $M(y) = 0$. Let i be as before and recall that the support of κ is compact by assumption. Since M is continuous on $\mathbb{R}^n \backslash N^*(\hat{\Omega}_{\kappa,M,p})$, there exists an open neighborhood of y in $\mathbb{R}^n \backslash N^*(\hat{\Omega}_{\kappa,M,p})$ such that for every point y^* in this neighborhood we either have $M(y^*) = 0$ or that $i/M(y^*)$ is not contained in the support of κ. It follows that the function f_i is constant equal to 0, and thus is in particular continuously partially differentiable, on this neighborhood.

To prove the third part of the lemma consider first the case $M \equiv M_{AM,1,\omega} \in \mathbb{M}_{AM}$, where we dropped the index c because the argument and the resulting polynomial do not depend on it. Suppose y satisfies $g^*_{\kappa,M_{AM,1,\omega},p}(y,X,R) \neq 0$. Then $M_{AM,1,\omega}(y)$ is well defined and by definition $M_{AM,1,\omega}(y) = \delta$ if and only if $\hat{\alpha}_1(y) = n^{-1}(c_1^{-1}\delta)^{1/c_2} =: \delta^*$ holds, where c_1 and c_2 are positive constants. This can equivalently be written as

$$\sum_{i=1}^{k} \omega_i \frac{4\hat{\rho}_i^2(y)\hat{\sigma}_i^4(y)}{(1-\hat{\rho}_i(y))^6(1+\hat{\rho}_i(y))^2} = \delta^* \sum_{i=1}^{k} \omega_i \frac{\hat{\sigma}_i^4(y)}{(1-\hat{\rho}_i(y))^4},$$

which, after multiplying both sides of the equation by $\prod_{j=1}^{k}(1-\hat{\rho}_j(y))^6(1+\hat{\rho}_j(y))^2$ (which is nonzero), is seen to be equivalent to

$$\sum_{i=1}^{k} \omega_i \left[4\hat{\rho}_i^2(y)\hat{\sigma}_i^4(y) \prod_{j\neq i}^{k}(1-\hat{\rho}_j(y))^6(1+\hat{\rho}_j(y))^2 \right]$$
$$- \delta^* \sum_{i=1}^{k} \omega_i \left[\hat{\sigma}_i^4(y)(1+\hat{\rho}_i(y))^2(1-\hat{\rho}_i(y))^2 \prod_{j\neq i}^{k}(1-\hat{\rho}_j(y))^6(1+\hat{\rho}_j(y))^2 \right] = 0.$$

By multiplying both sides of this equation by a suitably large power of the products of the denominators of $\hat{\rho}_i(y)$ (which are nonzero), we can write the preceding equation equivalently as

$$\sum_{i=1}^{k} \omega_i \bar{p}_i^{(\delta)}(\hat{Z}(y)) = 0$$

where each $\bar{p}_i^{(\delta)} : \mathbb{R}^{k \times (n-p)} \to \mathbb{R}$ for $i = 1,\dots,k$ is a multivariate polynomial. In a final step we multiply both sides of the equation by a suitably large power of the non-vanishing factor in Equation (10) to obtain an equivalent equation of the form

$$g^{(\delta)}_{\kappa,M_{AM,1,\omega},p}(y,X) = \sum_{i=1}^{k} \omega_i p_i^{(\delta)}(y,X) = 0,$$

where each $p_i^{(\delta)} : \mathbb{R}^n \times \mathbb{R}^{n \times k} \to \mathbb{R}$ is a multivariate polynomial. Therefore, the condition

$$g^{*}_{\kappa,M_{AM,1,\omega},p}(y,X) \neq 0 \text{ and } M_{AM,1,\omega}(y) = \delta$$

can be equivalently stated as

$$g^{*}_{\kappa,M_{AM,1,\omega},p}(y,X,R) \neq 0 \text{ and } g^{(\delta)}_{\kappa,M_{AM,1,\omega},p}(y,X) = 0.$$

Finally, we note that the multivariate polynomial $g^{(\delta)}_{\kappa,M_{AM,1,\omega},p} : \mathbb{R}^n \times \mathbb{R}^{n \times k} \to \mathbb{R}$ does not depend on the hypothesis (R,r). This proves the last part of the lemma in case $M \equiv M_{AM,1,\omega} \in \mathbb{M}_{AM}$. The proof of the case $M \equiv M_{AM,2,\omega} \in \mathbb{M}_{AM}$ is almost identical and therefore we omit it. We finally note that similar arguments can be used to prove the statement in case $M \in \mathbb{M}_{NW}$, but we omit details. $\quad \square$

Proof of Proposition 4.5. We first prove that the sets $\mathfrak{X}_2(e_+)$, $\mathfrak{X}_2(e_-)$, $\tilde{\mathfrak{X}}_2(e_+)$ and $\tilde{\mathfrak{X}}_2(e_-)$ do not depend on the specific choice of $\mu^*_{0,X} \in \mathfrak{M}_{0,X}$. This follows from an invariance argument. Consider for example the set $\mathfrak{X}_2(e_+)$, which is by definition a subset of $\mathfrak{X}_0 \backslash \mathfrak{X}_1(e_+)$. Every element X of this superset satisfies $g_{\kappa,M,p}(.,X,R) \not\equiv 0$. Hence, for every such X the corresponding test statistic T_X is invariant w.r.t. $G(\mathfrak{M}_{0,X})$ by Lemma C.1. It now immediately follows that $\mathfrak{X}_1(e_+)$ does not depend on the specific choice of $\mu^*_{0,X} \in \mathfrak{M}_{0,X}$. The same argument shows that the statement in Part 2 Condition (d) is independent of the specific choice of $\mu^*_{0,(e_+,\tilde{X})}$. We shall now prove the three main parts of the proposition and start with the first:

1) We begin with the statement concerning $\mathfrak{X}_1(e_+)$. Under the maintained assumptions we know from Part 5 of Lemma 3.10 that $g^*_{\kappa,M,p}(.,.,.) : \mathbb{R}^n \times \mathbb{R}^{n \times k} \times \mathbb{R}^{q \times k} \to \mathbb{R}$ is a multivariate polynomial. This immediately implies that $g^*_{\kappa,M,p}(e_+,.,R) : \mathbb{R}^{n \times k} \to \mathbb{R}$ is a multivariate polynomial, showing that

$$\left\{ X \in \mathbb{R}^{n \times k} : g^*_{\kappa,M,p}(e_+,X,R) = 0 \right\}$$

is an algebraic superset of $\mathfrak{X}_1(e_+)$. It hence suffices to show that the set in the previous display is a $\lambda_{\mathbb{R}^{n \times k}}$- null set, or equivalently that $g^*_{\kappa,M,p}(e_+,.,R) \not\equiv 0$. To this end we shall use Lemma B.1 (with $t = k$) to construct a matrix $X \in \mathfrak{X}_0$ such that $g^*_{\kappa,M,p}(e_+,X,R) \neq 0$. Let $H \in \mathbb{R}^{(k+1) \times k}$ be an auxiliary matrix the column vectors of which span $\text{span}(\bar{e}_+)^{\perp}$, where $\bar{e}_+ = (1,\ldots,1)' \in \mathbb{R}^{k+1}$ is the vector obtained from e_+ by selecting the coordinates with indices $j_i = 1 + (i-1)(p+1)$ for $i = 1,\ldots,k+1$ (we shall need a similar construction for e_- later on). Note that this selection is feasible because $j_{k+1} = 1 + k(p+1) \leq n$ holds as a consequence of the assumption $n - [k(p+1)+p] - \mathbf{1}_{\mathbb{M}_{AM}}(M) \geq 0$ together with $p \geq 1$. We also note that H does not contain a row consisting of zeros only. If the construction of M involves a weights vector ω (which is assumed to be functionally independent of the design) we choose the columns of H in such

a way that $H\omega = (-1, 1, 0, \ldots, 0)'$ which is possible because this vector is an element of $\operatorname{span}(\bar{e}_+)^\perp$ and $\omega \neq 0$ holds. We now let $X \in \mathbb{R}^{n \times k}$ be the matrix the non-zero rows of which are precisely $X_{j_i \cdot} = H_{i \cdot}$ for $i = 1, \ldots, k + 1$, i.e.,

$$X = \left(H'_{1\cdot}, 0_{k,p}, H'_{2\cdot}, 0_{k,p}, H'_{3\cdot}, 0_{k,p}, \ldots, H'_{(k+1)\cdot}, 0_{k,n-k(p+1)-1}\right)' \in \mathbb{R}^{n \times k}, \quad (24)$$

where $0_{m_1, m_2}$ denotes the $m_1 \times m_2$-dimensional zero matrix (here $0_{k,n-k(p+1)-1}$ vanishes if $n - k(p+1) - 1 = 0$). Obviously $\operatorname{rank}(X) = \operatorname{rank}(H) = k$ holds, which implies $X \in \mathfrak{X}_0$. Furthermore, $e_+ \perp \operatorname{span}(X)$ holds, which follows immediately from $\bar{e}_+ \perp \operatorname{span}(H)$, because the non-zero columns of X have column indices j_i for $i = 1, \ldots, k+1$ by construction. Therefore, we see that $\hat{u}_X(e_+) = e_+$ showing that

$$\hat{V}_X(e_+) = X' \operatorname{diag}(\hat{u}_X(e_+)) = X' \operatorname{diag}(e_+).$$

Now we apply Lemma B.1 with $t = k$ to $(e_+, X) \in \mathbb{R}^n \times \mathfrak{X}_0$. That the tuple (e_+, X) satisfies (A1) of that lemma is obvious from the preceding display and Equation (24). We also see that (A2) is satisfied because $j_{i+1} - j_i = p + 1$ for $i = 1, \ldots k$, and $k(p + 1) + p + \mathbf{1}_{\mathbb{M}_{AM}}(M) \leq n$ implies $n - j_{k+1} = n - k(p+1) - 1 \geq p - 1$. That (A3) is satisfied follows from the preceding display together with $\operatorname{rank}(X) = \operatorname{rank}(H) = k$. To infer $g^*_{\kappa,M,p}(e_+, X, R) \neq 0$ from Part 2 of Lemma B.1, we consider three cases: First, if $M \in \mathbb{M}_{KV}$ (CKV) is obviously satisfied and we are done. Secondly, assume that $M \in \mathbb{M}_{AM}$. In this case we have by assumption $k(p + 1) + p + 1 \leq n$ which implies $n - j_{k+1} = n - k(p+1) - 1 > p - 1$. This shows that (CAM) is satisfied. Thirdly suppose that $M \in \mathbb{M}_{NW}$. Since $\hat{A}^{(p)}_X(e_+) = 0$ follows from Part 1 of Lemma B.1, we see that $\hat{Z}_X(e_+) = \hat{V}_{p,X}(e_+)$ and hence that the nonzero columns of $\hat{Z}_X(e_+)$ are precisely $H'_{i\cdot}$ for $i = 2, \ldots, k+1$. By construction we have $H_{2\cdot}\omega \neq 0$ and $H_{i\cdot}\omega = 0$ for $i = 3, \ldots, k+1$. This shows that exactly one coordinate of $\omega'\hat{Z}_X(e_+)$ is non-zero which implies that (CNW) holds. To show that $\mathfrak{X}_1(e_-)$ is a $\lambda_{\mathbb{R}^{n \times k}}$−null set, we can use a similar construction: we replace e_+ by e_- throughout. If p is even we then have $\bar{e}_- = (-1, 1, -1, \ldots, (-1)^{k+1})'$. If p is odd we then have $\bar{e}_- = (-1, -1, \ldots, -1)'$. Furthermore, if the construction of M involves a weights vector we choose H such that $H\omega = (1, 1, 0, \ldots, 0)'$ if p is even, and $H\omega = (-1, 1, 0, \ldots, 0)'$ if p is odd. The remaining arguments are identical.

Now consider $\mathfrak{X}_2(e_+)$. Using Part 1 of Lemma C.2 we see that in case $M \in \mathbb{M}_{KV}$ the set $\mathfrak{X}_2(e_+)$ is empty, because $(\operatorname{grad} T_X(.))|_{e_+ + \mu^*_{0,X}}$ exists whenever $X \in \mathfrak{X}_0 \backslash \mathfrak{X}_1(e_+)$. Next consider the cases where $M \notin \mathbb{M}_{KV}$ and where κ satisfies Assumption 3. From Part 2a of Lemma C.2 we know that for $X \in \mathfrak{X}_0 \backslash \mathfrak{X}_1(e_+)$ the non-existence of $(\operatorname{grad} T_X(.))|_{e_+ + \mu^*_{0,X}}$ implies either $M(e_+) = 0$ or $i/M(e_+) \in \Delta(\kappa)$ for some $|i| = 1, \ldots, n - p - 1$. The latter two cases can clearly be summarized as $M(e_+) \in \bar{\Delta}$, where $\bar{\Delta} = \{\delta_0, \delta_1, \ldots, \delta_m\}$ is a set consisting of finitely many elements. Therefore,

$$\mathfrak{X}_2(e_+) \subseteq \left\{X \in \mathfrak{X}_0 \backslash \mathfrak{X}_1(e_+) : (\operatorname{grad} T_X(.))|_{e_+ + \mu^*_0} \text{ does not exist}\right\}$$
$$\subseteq \left\{X \in \mathfrak{X}_0 \backslash \mathfrak{X}_1(e_+) : M(e_+) \in \bar{\Delta}\right\}$$

$$= \bigcup_{i=0}^{m} \{X \in \mathfrak{X}_0 \backslash \mathfrak{X}_1(e_+) : M(e_+) = \delta_i\}.$$

We use Part 3 of Lemma C.2 to rewrite the latter set as

$$\bigcup_{i=0}^{m} \left\{X \in \mathfrak{X}_0 : g^*_{\kappa,M,p}(e_+, X, R) \neq 0 \text{ and } g^{(\delta_i)}_{\kappa,M,p}(e_+, X) = 0\right\}$$

$$= \left\{X \in \mathfrak{X}_0 : g^*_{\kappa,M,p}(e_+, X, R) \neq 0 \text{ and } \prod_{i=0}^{m} g^{(\delta_i)}_{\kappa,M,p}(e_+, X) = 0\right\},$$

which is clearly a subset of

$$\left\{X \in \mathbb{R}^{n \times k} : \prod_{i=0}^{m} g^{(\delta_i)}_{\kappa,M,p}(e_+, X) = 0\right\}.$$

Part 3 of Lemma C.2 shows that $\prod_{i=0}^{m} g^{(\delta_i)}_{\kappa,M,p}(e_+, .) : \mathbb{R}^{n \times k} \to \mathbb{R}$ is a multivariate polynomial. We consider two cases: First assume that $\prod_{i=0}^{m} g^{(\delta_i)}_{\kappa,M,p}(e_+, .) \not\equiv 0$. Consequently, the set in the previous display is a $\lambda_{\mathbb{R}^{n \times k}}$-null set. It hence follows that $\mathfrak{X}_2(e_+)$ is a $\lambda_{\mathbb{R}^{n \times k}}$-null set and we are done. Next, assume that $\prod_{i=0}^{m} g^{(\delta_i)}_{\kappa,M,p}(e_+, .) \equiv 0$. It follows that there must exist a single index i such that $g^{(\delta_i)}_{\kappa,M,p}(e_+, .) \equiv 0$ holds [this is easily shown by contradiction]. Part 3 of Lemma C.2 hence shows that $X \in \mathfrak{X}_0$ and $g^*_{\kappa,M,p}(e_+, X, R) \neq 0$, i.e., $X \in \mathfrak{X}_0 \backslash \mathfrak{X}_1(e_+)$, implies $M(e_+) = \delta_i$. Clearly,

$$\mathfrak{X}_2(e_+) \subseteq \{X \in \mathfrak{X}_0 \backslash \mathfrak{X}_1(e_+) : T_X(e_+ + \mu_0^*) = C\}.$$

If $X \in \mathfrak{X}_0 \backslash \mathfrak{X}_1(e_+)$, then (cf. the argument in the beginning of the proof of Lemma C.2 applied to $y = e_+$)

$$T_X(e_+ + \mu_{0,X}^*)$$
$$= \hat{\beta}_X(e_+)' R' \left(\frac{n}{n-p} B_{p,X}(e_+) \mathcal{W}_{n-p}(e_+) B'_{p,X}((e_+))\right)^{-1} R\hat{\beta}_X(e_+). \tag{25}$$

Furthermore, since $X \in \mathfrak{X}_0 \backslash \mathfrak{X}_1(e_+)$ implies $M(e_+) = \delta_i$, the matrix $\mathcal{W}_{n-p}(e_+)$ is constant $\bar{\mathcal{W}}_{n-p}$, say, on $\mathfrak{X}_0 \backslash \mathfrak{X}_1(e_+)$. Hence, for $X \in \mathfrak{X}_0 \backslash \mathfrak{X}_1(e_+)$, the statement $T_X(e_+ + \mu_{0,X}^*) = C$ is equivalent to

$$F_p^2(e_+, X) \left[\det(X'X)\hat{\beta}_X(e_+)\right]' R'$$
$$\times \operatorname{adj}(\bar{B}_{p,X}(e_+)\bar{\mathcal{W}}_{n-p}\bar{B}'_{p,X}((e_+)))R \left[\det(X'X)\hat{\beta}_X(e_+)\right]$$
$$= \frac{n}{n-p} \det(X'X)^2 \det(\bar{B}_{p,X}(e_+)\bar{\mathcal{W}}_{n-p}\bar{B}'_{p,X}((e_+)))C$$

where $\bar{B}_{p,X}(e_+)$ and $F_p(e_+, X)$ have been defined in the proof of Lemma 3.10, where it is shown that $g_{\kappa,M,p}(e_+, X) \neq 0$ and $X \in \mathfrak{X}_0$ (which is weaker than

$X \in \mathfrak{X}_0 \backslash \mathfrak{X}_1(e_+))$ implies $F_p(e_+, X) \neq 0$. Furthermore, it is shown in the proof of Lemma 3.10 that $[\bar{B}_{p,.}(e_+)]_{ij} : \mathbb{R}^{n \times k} \to \mathbb{R}$ (for $1 \leq i \leq q$ and $1 \leq j \leq n - p$) is a multivariate polynomial, and that $F_p(e_+, .) : \mathbb{R}^{n \times k} \to \mathbb{R}$ is a multivariate polynomial as well. It is easily seen that the coordinates of $\det(X'X)\hat{\beta}_X(e_+)$ as a function of X are multivariate polynomials. Putting this together we have shown that

$$\mathfrak{X}_2(e_+) \subseteq \left\{ X \in \mathbb{R}^{n \times k} : p_{\kappa,M,p}(e_+, X, C) = 0 \right\},$$

where $p_{\kappa,M,p}(e_+, ., C) : \mathbb{R}^{n \times k} \to \mathbb{R}$ is a multivariate polynomial. Therefore, if we can show that $p_{\kappa,M,p}(e_+, ., C) \not\equiv 0$ we obtain that $\mathfrak{X}_2(e_+)$ is a $\lambda_{\mathbb{R}^{n \times k}}$-null set. In the proof of the part concerning $\mathfrak{X}_1(e_+)$ above we have already constructed an $X \in \mathfrak{X}_0 \backslash \mathfrak{X}_1$ that satisfies $e_+ \perp \mathrm{span}(X)$ which implies $\hat{\beta}_X(e_+) = 0$. Together with Equation (25) this shows that $T_X(e_+ + \mu_{0,X}^*) = 0 < C$ holds for this specific X. But this immediately shows that $p_{\kappa,M,p}(e_+, X, C) \neq 0$ and we are done. Clearly, we can use an almost identical argument to prove the statement concerning $\mathfrak{X}_2(e_-)$. Under Assumption 2 the set of matrices $X \in \mathfrak{X}_0$ for which the first three cases of Theorem 4.2 do not apply is obviously a subset of $(\mathfrak{X}_1(e_+) \cup \mathfrak{X}_2(e_+)) \cap (\mathfrak{X}_1(e_-) \cup \mathfrak{X}_2(e_-))$. Hence the first part of the proposition follows.

2) We start with the statement concerning $\tilde{\mathfrak{X}}_1(e_-)$. Under the maintained assumptions we know from Part 5 of Lemma 3.10 that $g_{\kappa,M,p}^*(., ., .) : \mathbb{R}^n \times \mathbb{R}^{n \times k} \times \mathbb{R}^{q \times k} \to \mathbb{R}$ is a multivariate polynomial. This immediately implies that $g_{\kappa,M,p}^*(e_-, (e_+, .), R) : \mathbb{R}^{n \times (k-1)} \to \mathbb{R}$ is a multivariate polynomial, which shows that

$$\left\{ \tilde{X} \in \mathbb{R}^{n \times (k-1)} : g_{\kappa,M,p}^*(e_-, (e_+, \tilde{X}), R) = 0 \right\}$$

is an algebraic superset of $\tilde{\mathfrak{X}}_1(e_-)$. It hence suffices to show that the set in the previous display is a $\lambda_{\mathbb{R}^{n \times (k-1)}}$-null set, or equivalently to show that $g_{\kappa,M,p}^*(e_-, (e_+, .), R) \not\equiv 0$. Again, we shall use Lemma B.1 with $t = k$ to construct a matrix $\tilde{X} \in \tilde{\mathfrak{X}}_0$ such that $g_{\kappa,M,p}^*(e_-, (e_+, \tilde{X}), R) \neq 0$. The situation here is more complicated than in the first part, because the first column of the design matrix we seek has to be the intercept. For our construction we need some additional ingredients: By definition $p^* = p + 1$ if p is odd, and $p^* = p$ if p is even. If p is odd set $v = \bar{e}_-^* = (-1, -1, \ldots, -1, 1)' \in \mathbb{R}^{k+1}$, where $\bar{e}_-^*$ is the vector obtained from e_- by selecting the coordinates $j_i^* = j_i$ for $i = 1, \ldots, k$ and $j_{k+1}^* = j_k^* + p^* + 1$, where $j_i = 1 + (i-1)(p+1)$ for $i = 1, \ldots, k+1$ was defined in Part 1 above. This selection is feasible, because by assumption we have $k(p+1) + p^* \leq n$, which, since p is odd, gives $k(p+1) + p + 1 \leq n$, implying that $j_{k+1}^* = (k-1)(p+1) + p^* + 2 = k(p+1) + 2 \leq n$, because of $p \geq 1$. If p is even set $v = \bar{e}_- = (-1, 1, -1, \ldots, (-1)^{k+1})' \in \mathbb{R}^{k+1}$, the vector obtained from e_- by selecting the coordinates $j_i^* = j_i$ for $i = 1, \ldots, k+1$. Next, define

$$z = (-1, k^{-1}, \ldots, k^{-1})' \in \mathbb{R}^{k+1}.$$

We claim that v and z satisfy $u := \Pi_{\mathrm{span}(z)} v \neq 0$, $x := \Pi_{\mathrm{span}(z)^\perp} v = v - u$ is linearly independent of $e := (1, \ldots, 1)' \in \mathbb{R}^{k+1}$ and z is orthogonal to e. The

latter property is clearly always satisfied, regardless of whether p is even or odd. We thus only have to verify the first two conditions. We start with the case p odd. Here we have $z'v = 2k^{-1} \neq 0$ and therefore $\Pi_{\text{span}(z)}v \neq 0$. Furthermore $\Pi_{\text{span}(z)^\perp}v = v - \|z\|^{-2}z'vz$ can not equal ce for some $c \in \mathbb{R}$, because the last and the last but one coordinate of v are unequal. For p even $z'v = (1 + k^{-1}s)$, where s equals either 0 (if k is even) or 1 (if k is odd), therefore $z'v \neq 0$ holds and thus $\Pi_{\text{span}(z)}v \neq 0$. Furthermore $\Pi_{\text{span}(z)^\perp}v = v - \|z\|^{-2}z'vz$ can not equal ce for some $c \in \mathbb{R}$, because the second and third coordinate of v are unequal. This proves the claim. Using these properties, we see that $u \in \text{span}(z) \setminus \{0\}$ and

$$u = v - x = v - \Pi_{\text{span}(e,x)}(v - \Pi_{\text{span}(z)}v) = v - \Pi_{\text{span}(e,x)}v = \Pi_{\text{span}(e,x)^\perp}v,$$

where we have used that z is orthogonal to both e and x to derive the third equality. We shall now define an auxiliary matrix. Let L denote a $(k+1) \times k$-dimensional matrix such that $L_{\cdot 1} = e$, $L_{\cdot 2} = x$ and such that the remaining $k - 2$ columns $L_{\cdot j}$ for $j = 3, \ldots, k$ are linearly independent and orthogonal to $\text{span}(e, x, v)$. Since e and x are linearly independent, we have $\text{rank}(L) = k$. For later use we observe that

$$\Pi_{\text{span}(L)}v = \Pi_{\text{span}((e,x))}v = v - \Pi_{\text{span}(e,x)^\perp}v = v - u = x,$$

where the first equality follows immediately from $L_{\cdot j}$ for $j = 3, \ldots, k$ being linearly independent and orthogonal to $\text{span}(e, x, v)$. This immediately shows

$$\hat{\beta}_L(v) = (0, 1, 0, \ldots, 0)' \in \mathbb{R}^k.$$

Define the two k-vectors $r_- = (1, -1, 0, \ldots, 0)'$ and $r_+ = (1, 1, 0, \ldots, 0)'$. Let $X \in \mathbb{R}^{n \times k}$ be such that $X_{j_i^* \cdot} = L_i$ for $i = 1, \ldots, k+1$, and if the index $j \notin \{j_1^*, \ldots, j_{k+1}^*\}$, then let $X_{j \cdot} = r_+$ if $[e_-]_j = 1$, and let $X_{j \cdot} = r_-$ if $[e_-]_j = -1$. By construction the matrix X is of the form $X = (e_+, \tilde{X})$. We claim that $\hat{\beta}_X(e_-) = \hat{\beta}_L(v)$. To see this denote the set of indices $j \in \{1, \ldots, n\} \setminus \{j_1^*, \ldots, j_{k+1}^*\}$ such that $[e_-]_j = -1$ by $\mathcal{I}_-$, and the set of indices $j \in \{1, \ldots, n\} \setminus \{j_1^*, \ldots, j_{k+1}^*\}$ such that $[e_-]_j = 1$ by $\mathcal{I}_+$. The sum of squares $S(\beta) = \|e_- - X\beta\|^2$ can be written as

$$S(\beta) = \sum_{i=1}^{k+1}([e_-]_{j_i^*} - X_{j_i^* \cdot}\beta)^2 + \sum_{j \in \mathcal{I}_-}(-1 - r_-\beta)^2 + \sum_{j \in \mathcal{I}_+}(1 - r_+\beta)^2$$

$$= \sum_{i=1}^{k+1}(v_i - L_{i\cdot}\beta)^2 + \sum_{j \in \mathcal{I}_-}(-1 - r'_-\beta)^2 + \sum_{j \in \mathcal{I}_+}(1 - r'_+\beta)^2$$

$$= \|v - L\beta\|^2 + \sum_{j \in \mathcal{I}_-}(-1 - r'_-\beta)^2 + \sum_{j \in \mathcal{I}_+}(1 - r'_+\beta)^2$$

If we now plug in $\beta = \hat{\beta}_L(v)$ and note that $r'_+\hat{\beta}_L(v) = 1$ and $r'_-\hat{\beta}_L(v) = -1$ we see that

$$S(\hat{\beta}_L(v)) = \sum_{i=1}^{k+1}(v_i - L_{i\cdot}\hat{\beta}_L(v))^2 = \min_{\beta \in \mathbb{R}^k}\|v - L\beta\|^2.$$

This immediately proves the claim $\hat{\beta}_X(e_-) = \hat{\beta}_L(v)$. Hence, the residual vector satisfies

$$[\hat{u}_X(e_-)]_j = \begin{cases} u_i & \text{if } j = j_i^* \text{ for some } i = 1, \ldots, k+1 \\ 0 & \text{else.} \end{cases}$$

This immediately entails that $\hat{V}_X(e_-) = X' \operatorname{diag}(\hat{u}_X(e_-))$ equals

$$(u_1 L'_{1\cdot}, 0_{k,p}, u_2 L'_{2\cdot}, 0_{k,p}, \ldots \\ \ldots, u_k L'_{k\cdot}, 0_{k,p^*}, u_{k+1} L'_{(k+1)\cdot}, 0_{k,n-[k(p+1)+p^*-p+1]}), \tag{26}$$

where the indices of the nonzero columns of this matrix are precisely j_i^* for $i = 1, \ldots, k+1$, because the first column of L is e and $u_i \neq 0$ for $i = 1, \ldots, k+1$, the latter following since $u \in \operatorname{span}(z) \setminus \{0\}$ and $z_i \neq 0$ for $i = 1, \ldots, k+1$ by definition. In deriving the dimension of $0_{k,n-[k(p+1)+p^*-p+1]}$ we used

$$j_{k+1}^* = k(p+1) + 1 + p^* - p = \begin{cases} k(p+1) + 2 & \text{if } p \text{ odd} \\ k(p+1) + 1 & \text{if } p \text{ even.} \end{cases}$$

Now we apply Lemma B.1 (with $t = k$). Clearly, $\operatorname{rank}(X) = \operatorname{rank}(L) = k$. From Equation (26), and the discussion following it, we see that the tuple $(e_-, X) \in \mathbb{R}^n \times \mathfrak{X}_0$ satisfies Assumption (A1). Assumption (A2) is satisfied, because $j_{i+1}^* - j_i^* \geq p + 1$ for $i = 1, \ldots, k$, and because we see from the previous display that $n - j_{k+1}^* = n - [k(p+1) + 1 + p^* - p]$, which together with the assumption $k(p+1) + p^* + \mathbf{1}_{\mathbb{M}_{AM}}(M) \leq n$ implies $n - j_{k+1}^* \geq p - 1$. Assumption (A3) is satisfied because $\operatorname{rank}(\hat{V}_X(e_-)) = \operatorname{rank}(L) = k$. If $M \in \mathbb{M}_{KV}$ we are done. Consider the case $M \in \mathbb{M}_{AM}$. We show that Condition (CAM) is satisfied. But this is obvious, because the assumption $k(p+1) + p^* + 1 \leq n$ immediately implies $n - j_{k+1}^* > p - 1$. Suppose $M \in \mathbb{M}_{NW}$. We apply Part 4 of Lemma B.1. For this we claim that either $[\hat{V}_X(e_-)]_{1j_i^*} > 0$ for $i = 2, \ldots, k+1$ or $[\hat{V}_X(e_-)]_{1j_i^*} < 0$ for $i = 2, \ldots, k+1$. Assuming that this claim is true, the lemma shows that there exists a regular matrix $\bar{Q} \in \mathbb{R}^{k \times k}$ such that $X\bar{Q} \in \mathfrak{X}_0$, the first column of $X\bar{Q}$ is e_+ and $g^*_{\kappa,M,p}(e_+, X\bar{Q}, R) \neq 0$, and we are done. To prove the claim recall that by construction $u \in \operatorname{span}(z) \setminus \{0\}$ holds, which shows that either $u_i < 0$ for $i = 2, \ldots, k+1$ or $u_i > 0$ for $i = 2, \ldots, k+1$. Furthermore, the first column of L is the vector $e = (1, \ldots, 1)$. Equation (26) now shows that $[\hat{V}_X(e_-)]_{1j_i^*} = u_i$ for $i = 2, \ldots, k+1$. This proves the claim.

The part of the statement concerning $\tilde{X}_2(e_-)$ is established by exploiting an argument similar to the one given in Part 1 of the proof. Firstly, if $M \in \mathbb{M}_{KV}$, then we know from Part 1 of Lemma C.2 that $g^*_{\kappa,M,p}(e_-, (e_+, \tilde{X}), R) \neq 0$ implies existence of $\operatorname{grad}(T_{(e_+,\tilde{X})}(.))|_{e_- + \mu^*_{0,(e_+,\tilde{X})}}$. Therefore, $\tilde{X}_2(e_-)$ is empty in Case (a). It remains to prove the remaining three cases, in all of which Assumption 3 holds. Clearly we can also assume that $M \notin \mathbb{M}_{KV}$. We start with the following observation: Combining Assumption 3 with Part 2a of Lemma C.2 as in Part 1 of the proof, we see that there exists an integer $m \geq 0$ and real numbers

$\delta_0, \ldots, \delta_m$, such that

$$\tilde{X}_2(e_-) \subseteq \left\{ \tilde{X} \in \mathbb{R}^{n \times (k-1)} : \prod_{i=0}^{m} g_{\kappa,M,p}^{(\delta_i)}(e_-, (e_+, \tilde{X})) = 0 \right\}. \tag{27}$$

It follows with the same argument as in Part 1 that either $\tilde{X}_2(e_-)$ is a $\lambda_{\mathbb{R}^{n \times (k-1)}}$-null set, or there exists an index i such that $g_{\kappa,M,p}^{(\delta_i)}(e_-, (e_+, .)) \equiv 0$. In the former case we are done. In the latter case one can show, with a similar argument as in Part 1 of the proof, that

$$\tilde{\mathfrak{X}}_2(e_-) \subseteq \left\{ \tilde{X} \in \mathbb{R}^{n \times (k-1)} : p_{\kappa,M,p}(e_-, (e_+, \tilde{X}), C) = 0 \right\}, \tag{28}$$

where $p_{\kappa,M,p}(e_-, (e_+, .), C) : \mathbb{R}^{n \times (k-1)} \to \mathbb{R}$ is a multivariate polynomial. Either we have that $p_{\kappa,M,p}(e_-, (e_+, .), C) \not\equiv 0$ and $\tilde{\mathfrak{X}}_2(e_-)$ is a $\lambda_{\mathbb{R}^{n \times (k-1)}}$-null set, or $p_{\kappa,M,p}(e_-, (e_+, .), C) \equiv 0$ and the superset in the previous display coincides with $\tilde{\mathfrak{X}}_0 \backslash \tilde{\mathfrak{X}}_1(e_-)$. Consider Condition (d). If the function $\tilde{X} \mapsto T_{(e_+, \tilde{X})}(\mu_{0,(e_+, \tilde{X})}^* + e_-)$ is not constant C on $\tilde{\mathfrak{X}}_0 \backslash \tilde{\mathfrak{X}}_1(e_-)$, then $p_{\kappa,M,p}(e_-, (e_+, .), C) \not\equiv 0$, and hence the superset in Equation (28) is a $\lambda_{\mathbb{R}^{n \times (k-1)}}$-null set. This shows that $\tilde{\mathfrak{X}}_2(e_-)$ is a null set under Condition (d).

For Condition (b) we consider again the inclusion in Equation (27). Either the superset is a null set and we are done, or there must exist a real number δ_i such that $g_{\kappa,M,p}^{(\delta_i)}(e_-, (e_+, .)) \equiv 0$. Assume the latter. We exploit a property of the matrix $X = (e_+, \tilde{X})$ with $\tilde{X} \in \tilde{\mathfrak{X}}_0 \backslash \tilde{\mathfrak{X}}_1(e_-)$ constructed above. In the proof of Part 2 (CAM) of Lemma B.1 it is shown that for this specific X we have $M(e_-) = 0$. Therefore, $g_{\kappa,M,p}^{(0)}(e_-, (e_+, .)) \equiv 0$, or equivalently $M(e_-) = 0$ for every design matrix $X = (e_+, \tilde{X})$ with $\tilde{X} \in \tilde{\mathfrak{X}}_0 \backslash \tilde{\mathfrak{X}}_1(e_-)$. But since the kernel satisfies Assumption 4, the existence of $(\operatorname{grad} T_{(e_+, \tilde{X})}(.))|_{\mu_{0,(e_+, \tilde{X})}^* + e_-}$ for every $\tilde{X} \in \tilde{\mathfrak{X}}_0 \backslash \tilde{\mathfrak{X}}_1(e_-)$ then follows from Part 2b of Lemma C.2. Hence, $\tilde{\mathfrak{X}}_2(e_-) = \emptyset$.

Consider Condition (c). We use a similar argument as under Condition (b): We establish the existence of a sequence of matrices $(e_+, \tilde{X}_m)$ with $\tilde{X}_m$ eventually in $\tilde{\mathfrak{X}}_0 \backslash \tilde{\mathfrak{X}}_1(e_-)$, such that $M(e_-) \to 0$ as $m \to \infty$. From an argument as in the proof under Condition (b) this then implies that either $\tilde{\mathfrak{X}}_2(e_-)$ is a null set, or that $M(e_-) \equiv 0$ on $\tilde{\mathfrak{X}}_0 \backslash \tilde{\mathfrak{X}}_1(e_-)$. But since κ satisfies Assumption 4, it then follows from Part 2b of Lemma C.2 that in the latter case $\tilde{\mathfrak{X}}_2(e_-)$ is empty. This then proves the claim. It remains to construct a sequence $\tilde{X}_m$ as claimed. By assumption $\omega_i > 0$ for some $i > 1$. Assume without loss of generality that $i = 2$ (otherwise we have to interchange the columns of the $\tilde{X}_m$ sequence to be constructed accordingly). Let $X = (e_+, \tilde{X})$ be as constructed above. Let $\gamma_m > 0$ be a sequence diverging to ∞. Recall that by construction $\tilde{X}_{j_i^*1} = x_i$ for $i = 1, \ldots, k+1$. Since p is odd, a simple calculation shows that $x = \Pi_{\operatorname{span}(z)^\perp} v$ equals

$$x = \left(-1 + \frac{2}{k+1}, -1 - \frac{2k^{-1}}{k+1}, \ldots, -1 - \frac{2k^{-1}}{k+1}, 1 - \frac{2k^{-1}}{k+1} \right)'.$$

Let $c = 1 + \frac{2k^{-1}}{k+1}$ and note that

$$x + ce = 2(\frac{k^{-1}+1}{k+1}, 0, \ldots, 0, 1)'.$$

Define the $k \times k$ dimensional regular matrix

$$Q_m = QD_m = \begin{pmatrix} 1 & c & \ldots & \ldots & \ldots & 0 \\ 0 & 1 & 0 & \ldots & \ldots & 0 \\ 0 & 0 & 1 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & 0 & \ldots & 1 \end{pmatrix} \mathrm{diag}(1, \gamma_m, 1, \ldots, 1).$$

Clearly, post-multiplying a matrix with k columns by Q_m has the same effect as adding c times the first column to the second column, then multiplying the column so obtained by γ_m and leaving all other columns unchanged. Since $L_{\cdot 1} = e$ and $L_{\cdot 2} = x$, the expression for $x + ce$ above shows that the second column of LQ, has precisely two nonzero elements with indices 1 and $k + 1$, respectively. Since $X = (e_+, \tilde{X}) \in \mathfrak{X}_0$ and (e_-, X) satisfies (A1)-(A3) as established above, Part 3 of Lemma B.1 shows that

$$X_m = (e_+, \tilde{X})Q_m = (e_+, \tilde{X}_m) \in \mathfrak{X}_0,$$

and that the tuple $(e_-, X_m) \in \mathbb{R}^n \times \mathfrak{X}_0$ satisfies (A1), (A2) and (A3). An application of Lemma B.1 to that tuple then shows that $\mathrm{rank}(\hat{Z}_{X_m})(e_-)) = k$, which together with $\omega \neq 0$ immediately implies $\bar{\sigma}_{0,X_m}(e_-) \neq 0$. Since $\hat{Z}_X(e_-)$ is obtained from $\hat{V}_X(e_-)$ by deleting its first p columns, we observe, using the remark concerning the second column of LQ above together with Equation (26), that the second row of $Q'\hat{Z}_X(e_-)$ has exactly one non-zero coordinate, namely $2u_{k+1}$. Consider

$$\begin{aligned} \frac{\bar{\sigma}_{i,X_m}(e_-)}{\bar{\sigma}_{0,X_m}(e_-)} &= \frac{\sum_{j=|i|+1}^{n-p} \omega'[\hat{Z}_{X_m}(y)]_{\cdot j}[\hat{Z}_{X_m}(y)]'_{\cdot(j-|i|)}\omega}{\sum_{j=1}^{n-p} \omega'[\hat{Z}_{X_m}(y)]_{\cdot j}[\hat{Z}_{X_m}(y)]'_{\cdot j}\omega} \\ &= \frac{\sum_{j=|i|+1}^{n-p} \bar{\omega}'_m[Q'\hat{Z}_X(y)]_{\cdot j}[Q'\hat{Z}_X(y)]'_{\cdot(j-|i|)}\bar{\omega}_m}{\sum_{j=1}^{n-p} \bar{\omega}'_m[Q'\hat{Z}_X(y)]_{\cdot j}[Q'\hat{Z}_X(y)]'_{\cdot j}\bar{\omega}_m}, \end{aligned} \tag{29}$$

for $|i| = 1, \ldots, n-p-1$, where $\bar{\omega}_m = D_m\omega/\|D_m\omega\|$. Clearly $\bar{\omega}_m \to (0, 1, 0, \ldots, 0)$. Since the second row of $Q'\hat{Z}_{(e_+, \tilde{X})}$ contains by construction precisely one nonzero entry, it follows that the limit of (29) must be 0 for $i = 1, \ldots, n - p - 1$. Furthermore we have $\bar{\sigma}_{0,X_m}(e_-) \to \infty$. It immediately follows from $w(0) = 1$ and the definition of M that $M(e_-)$ is well defined for m large and that it converges to 0 as $m \to \infty$. The remaining part of the proposition is obvious.

3) Let $\tilde{X} \in \tilde{\mathfrak{X}}_0$ and assume that $X = (e_+, \tilde{X})$ satisfies $g^*_{\kappa,M,p}(., X, R) \not\equiv 0$. Obviously, $e_+ \in \mathrm{span}(X)$. Note that $\hat{\beta}_X(e_+) = e_1(k)$. The first column of R is non-zero. Therefore $R\hat{\beta}_X(e_+) \neq 0$. Thus we can (since Assumption 2 holds) apply Part 4 of Theorem 4.2. □

Appendix D: Proofs of results in Section 5

Proof of Theorem 5.2. We verify the assumptions of Theorem 5.21 in Preinerstorfer and Pötscher (2016) with $\hat{\Omega}_{\kappa,M,p} = \check{\Omega}$ and $\hat{\beta} = \check{\beta}$. Because $g^*_{\kappa,M,p}(., X, R) \not\equiv 0$ by assumption, and since the triple κ, M, p satisfies Assumption 1, we can use Lemma C.1 to conclude that $\hat{\beta}$ and $\hat{\Omega}_{\kappa,M,p}$ satisfy Assumptions 5, 6 and 7 in Preinerstorfer and Pötscher (2016), that $\hat{\Omega}_{\kappa,M,p}$ is almost everywhere positive definite and that T is invariant w.r.t. $G(\mathfrak{M}_0)$. Assumption 5, Remark 5.1 (Part (iii)) together with Remark 5.14 (ii) in Preinerstorfer and Pötscher (2016) now shows that $J(\mathfrak{C}) = \text{span}(e_+) \cup \text{span}(e_-)$ and that all assumptions on $\mathfrak{C}$ appearing in Theorem 5.21 in Preinerstorfer and Pötscher (2016) are satisfied. Because $e_+, e_- \in \mathfrak{M}$ is assumed we have $J(\mathfrak{C}) \subseteq \mathfrak{M}$. The assumption $R\hat{\beta}(e_+) = R\hat{\beta}(e_-) = 0$ even implies $J(\mathfrak{C}) \subseteq \mathfrak{M}_0 - \mu_0$ (for some arbitrary $\mu_0 \in \mathfrak{M}_0$). Invariance of T w.r.t. $G(\mathfrak{M}_0)$ then shows that Equation (34) in Theorem 5.21 of Preinerstorfer and Pötscher (2016) is satisfied. The assumptions on $\check{\Omega}$ appearing in Parts 2 and 3 of that theorem are satisfied, because $\hat{\Omega}_{\kappa,M,p}$ is positive definite almost everywhere. The theorem now follows from Theorem 5.21 in Preinerstorfer and Pötscher (2016), using a standard subsequence argument, positive definiteness of every element of $\mathfrak{C}$ and compactness of $\mathfrak{C}^*$, to obtain the second statement in Part 3 from the corresponding Part of Theorem 5.21 in Preinerstorfer and Pötscher (2016). The claim in parenthesis in Part 3 follows from the corresponding claim in parenthesis in Theorem 5.21 of Preinerstorfer and Pötscher (2016), together with the observation that the conditions on e_+ and e_- have only been used to verify the condition in Equation (34) of Theorem 5.21 of Preinerstorfer and Pötscher (2016) (cf. the proof of Theorem 3.7 in Preinerstorfer and Pötscher (2016)). □

Proof of Theorem 5.4. We apply Theorem 5.21 of Preinerstorfer and Pötscher (2016) with the estimators $\hat{\Omega}_{\kappa,\bar{M},p,\bar{X}} = \check{\Omega}$ and $(I_k, 0)\hat{\beta}_{\bar{X}} = \check{\beta}$. Obviously, the test statistic defined in Equation (28) of Preinerstorfer and Pötscher (2016) based on these estimators coincides with the test statistic $\bar{T}$ as defined in the statement of the present theorem. Since the assumptions concerning $\mathfrak{C}$ in the present theorem are the same as in Proposition 5.2, we see from the proof of this proposition that it suffices to verify that $\hat{\Omega}_{\kappa,\bar{M},p,\bar{X}}$ and $(I_k, 0)\hat{\beta}_{\bar{X}}$ satisfy Assumption 5 in Preinerstorfer and Pötscher (2016), that $\hat{\Omega}_{\kappa,\bar{M},p,\bar{X}}$ is almost everywhere positive definite (implying that Assumptions 6 and 7 in Preinerstorfer and Pötscher (2016) are satisfied), and that the invariance condition in Equation (34) of Preinerstorfer and Pötscher (2016) is satisfied by $\bar{T}$. By definition, $\hat{\Omega}_{\kappa,\bar{M},p,\bar{X}}$ is the estimator one would obtain following Steps 1-3 of the construction in Section 3 based on κ, $\bar{M}$ and p, if $\bar{X}$ was the underlying design matrix (observe that $\bar{X}$ is of full column rank) and $(\bar{R}, r)$ was the hypothesis to be tested. By assumption, the triple κ, M, p satisfies Assumption 1 w.r.t. the (dimensions k and n of the) design matrix X and additionally $1 \le p \le n/(k+3)$ holds. From $1 \le \bar{k} - k \le 2$, and the definition of $\bar{M}$ it follows that the triple κ, $\bar{M}$, p satisfies Assumption 1 w.r.t. (the dimensions $\bar{k}$ and n of) $\bar{X}$. Furthermore, it is assumed that $g^*_{\kappa,\bar{M},p}(., \bar{X}, \bar{R}) \not\equiv 0$. Therefore, we can apply Lemma C.1, acting as if $\bar{X}$ was the

underlying design matrix, to conclude that $\hat{\beta}_{\bar{X}}$ and $\hat{\Omega}_{\kappa,\bar{M},p,\bar{X}}$ satisfy Assumption 5 in Preinerstorfer and Pötscher (2016) with $N = N(\hat{\Omega}_{\kappa,\bar{M},p,\bar{X}})$ and k replaced by $\bar{k}$, X replaced by $\bar{X}$ and $\mathfrak{M}$ replaced by $\bar{\mathfrak{M}} = \mathrm{span}(\bar{X})$. Furthermore, Lemma C.1 shows that $\hat{\Omega}_{\kappa,\bar{M},p,\bar{X}}$ is almost everywhere positive definite. We now apply Part 1 of Proposition 5.23 in Preinerstorfer and Pötscher (2016) to obtain that $\hat{\Omega}_{\kappa,\bar{M},p,\bar{X}}$ and $(I_k, 0)\hat{\beta}_{\bar{X}}$ satisfy (the original) Assumption 5 in Preinerstorfer and Pötscher (2016), and that the invariance condition is satisfied. To this end, it suffices to verify that in each of the four cases we have $\mathrm{span}(J(\mathfrak{C})) \cap \mathfrak{M} \subseteq \mathfrak{M}_0 - \mu_0$ (for some arbitrary $\mu_0 \in \mathfrak{M}_0$). This is obvious in the first three cases. For Case 4 we can use exactly the same argument as in the proof of Part 4 in Theorem 3.8 in Preinerstorfer and Pötscher (2016). □

Proof of Proposition 5.5. We begin with the proof of the first statement. We note that for $\lambda_{\mathbb{R}^{n \times k}}$- almost every $X \in \mathfrak{X}_0$ Case 3 of Theorem 5.4 applies: Since by assumption $(k+3)(p^*+2) + p - 1 \leq n$ and by definition $p^* \geq 1$, we have $k + 2 < n$. Therefore, the set of matrices X in $\mathbb{R}^{n \times k}$ such that $\det((X, e_+, e_-)(X, e_+, e_-)') = 0$ holds is a $\lambda_{\mathbb{R}^{n \times k}}$- null set. Hence, $e_+, e_- \notin \mathfrak{M}_X$ and $\mathrm{rank}((X, e_+, e_-)) = k + 2$ holds for $\lambda_{\mathbb{R}^{n \times k}}$- almost every $X \in \mathfrak{X}_0$. It remains to verify that $g^*_{\kappa,\bar{M},p}(., \bar{X}, \bar{R}) \not\equiv 0$ for almost every $X \in \mathfrak{X}_0$, where $\bar{X} = \bar{X}(X) = (X, e_+, e_-)$, $\bar{R} = (R, 0, 0)$ and $\bar{M}$ is constructed as outlined in Theorem 5.4. For that it suffices to find a matrix $X \in \mathbb{R}^{n \times k}$ and a vector $y \in \mathbb{R}^n$ such that $g^*_{\kappa,\bar{M},p}(y, \bar{X}, \bar{R}) \neq 0$. To see this note that the triple κ, $\bar{M}$, p satisfies Assumption 1 w.r.t. (the dimensions of) $\bar{X}$ (cf. the proof of Theorem 5.4). Therefore, Lemma 3.10 shows that $(y, X) \mapsto g^*_{\kappa,\bar{M},p}(y, \bar{X}, \bar{R})$ is a multivariate polynomial. If we can find a matrix X and a vector y as above, this implies that the multivariate polynomial $(y, X) \mapsto g^*_{\kappa,\bar{M},p}(y, \bar{X}, \bar{R})$ is not the zero polynomial, and therefore the zero set of this multivariate polynomial is a $\lambda_{\mathbb{R}^n \times \mathbb{R}^{n \times k}}$- null set. It follows that for $\lambda_{\mathbb{R}^{n \times k}}$- almost every X we must have $g^*_{\kappa,\bar{M},p}(., \bar{X}, \bar{R}) \not\equiv 0$ [Assuming the opposite, there exists a set $A \in \mathcal{B}(\mathbb{R}^{n \times k})$ of positive $\lambda_{\mathbb{R}^{n \times k}}$- measure such that $g^*_{\kappa,\bar{M},p}(., \bar{X}, \bar{R}) \equiv 0$ for every $X \in A$, which implies that $\mathbb{R}^n \times A \subseteq \mathbb{R}^n \times \mathbb{R}^{n \times k}$ is a subset of the zero set of $(y, X) \mapsto g^*_M(y, \bar{X}, \bar{R})$. But clearly $\mathbb{R}^n \times A$ has positive Lebesgue measure, a contradiction.]. In the following we shall construct such a pair (y, X) as above:

Let $\delta \neq 0$ and define $w_1, w_2, v(\delta) \in \mathbb{R}^{k+4}$ as $w_1 = (1, 1, \ldots, 1)'$, $w_2 = (-1, -1, 1, \ldots, 1)'$ and $v(\delta) = (-\delta, \delta, -(k+1), 1, \ldots, 1, 1)'$. By construction $v(\delta)$ is orthogonal to w_1 and w_2. Noting that $[w_1]_i = [w_2]_i$ for $i \geq 3$ a dimensionality argument implies existence of k normalized vectors $w_3, \ldots, w_{k+2} \in \mathbb{R}^{k+4}$, that are functionally independent of δ, linearly independent and orthogonal to $e_1(k+4)$, $e_2(k+4)$, w_1, w_2 and $v(\delta)$ (for every $\delta \neq 0$). Recall that $e_1(k+4)$ and $e_2(k+4)$ are the first two elements of the canonical basis of $\mathbb{R}^{k+4}$. Hence, the first two coordinates of w_i for $i = 3, \ldots, k+2$ are zero. These orthogonality properties readily imply

$$\Pi_{\mathrm{span}(w_3, \ldots, w_{k+2}, w_1, w_2)^\perp} v(\delta) = \Pi_{\mathrm{span}(w_3+w_1, \ldots, w_{k+2}+w_1, w_1, w_2)^\perp} v(\delta) = v(\delta) \quad (30)$$

and $\mathrm{rank}(\bar{W}) = k + 2$ for $\bar{W} = (w_3 + w_1, \ldots, w_{k+2} + w_1, w_1, w_2)$. Inserting zero

coordinates and rows, respectively, we shall now suitably embed $v(\delta) \in \mathbb{R}^{k+4}$ and $W = (w_3 + w_1, \ldots, w_{k+2} + w_1) \in \mathbb{R}^{k+4 \times k}$ into $\mathbb{R}^n$ and $\mathbb{R}^{n \times k}$. Define $y(\delta)$ as

$$(v_1(\delta), 0_{1,p^*+1}, v_2(\delta), 0_{1,p^*}, v_3(\delta), 0_{1,p^*+1}, v_4(\delta), 0_{1,p^*+1}, \ldots$$
$$\ldots, v_{k+4}(\delta), 0_{1,p-1}, 0_{1,n-n^*})'$$

and X as

$$(W'_{1.}, 0_{k,p^*+1}, W'_{2.}, 0_{k,p^*}, W'_{3.}, 0_{k,p^*+1}, W'_{4.}, 0_{k,p^*+1}, \ldots, W'_{(k+4).}, 0_{1,p-1}, 0_{1,n-n^*})',$$

where $n^* = (k+3)(p^*+2)+p-1$, a number that does not exceed n by assumption. We emphasize that by construction X does *not* depend on δ. Furthermore, if we delete from e_+ and e_- those coordinates that correspond to the zero coordinates that have been inserted to obtain $y(\delta)$ from $v(\delta)$, we obtain the vectors w_1 and w_2. Therefore, it follows from Equation (30) that $y(\delta)$ is orthogonal to $\operatorname{span}(\bar{X}) = \operatorname{span}((X, e_+, e_-))$, that $\operatorname{rank}(X, e_+, e_-) = \operatorname{rank}(\bar{W}) = k + 2$ and that for every $\delta \neq 0$ we have

$$\hat{u}_{\bar{X}}(y(\delta)) = y(\delta).$$

As an immediate consequence we obtain

$$\begin{aligned}
\hat{V}_{\bar{X}}(y(\delta)) &= \bar{X}' \operatorname{diag}(y(\delta)) \\
&= (v_1(\delta)\bar{W}'_{1.}, 0_{k+2,p^*+1}, v_2(\delta)\bar{W}'_{2.}, 0_{k+2,p^*}, v_3(\delta)\bar{W}'_{3.}, 0_{k+2,p^*+1}, \ldots \\
&\qquad \ldots, v_{k+4}(\delta)\bar{W}'_{k+4.}, 0_{k+2,p-1}, 0_{k+2,n-n^*}),
\end{aligned} \tag{31}$$

where we recall that all coordinates of $v(\delta)$ are nonzero and therefore $\hat{V}_{\bar{X}}(y(\delta))$ has precisely $k + 4$ nonzero columns. We now intend to apply Lemma B.1 with $t = k + 3$, acting as if $\bar{X} \in \mathbb{R}^{n \times (k+2)}$ was the underlying design, $(\bar{R}, r)$ was the hypothesis to be tested and with the triple κ, $\bar{M}$, p which obviously satisfies Assumption 1 with respect to $\bar{X}$ (a matrix with $k+2$ columns), since by assumption we have $1 \leq p \leq \frac{n}{k+3}$ (Note that due to interpreting $\bar{X}$ as the underlying design, $k + 2$ corresponds to the 'k' in Lemma B.1). We note first that removing the first or last row of $\bar{W}$ does not reduce its rank, because $v(\delta)$ (a vector all coordinates of which are nonzero) is orthogonal to every column of this matrix (cf. the argument in the beginning of the proof of Lemma B.1). Using $p^* \geq p$ we hence see that Assumptions (A1)-(A3) in Lemma B.1 are satisfied by construction. Now consider the case $M \in \mathbb{M}_{KV}$. By definition $\bar{M}$ is an element of $\mathbb{M}_{KV}$ (acting as if $\bar{X}$ was the underlying design matrix). Therefore, Condition (CKV) is satisfied for $\delta \neq 0$ arbitrary, and $g^*_{\kappa, \bar{M}, p}(y(\delta), \bar{X}, \bar{R}) \neq 0$ follows. Consider the case where $M \in \mathbb{M}_{AM}$. Since $n - n^* \geq \mathbf{1}_{\mathbb{M}_{AM}}(M) = 1$ and because of $j_{k+4} = n^*-p+1$ it follows that $n-j_{k+4} > p-1$. Therefore, Condition (CAM) in Lemma B.1 is satisfied and therefore $g^*_{\kappa, \bar{M}, p}(y(\delta), \bar{X}, \bar{R}) \neq 0$ for $\delta \neq 0$ arbitrary. It remains to consider the case where $M \in \mathbb{M}_{NW}$. It suffices to find a $\delta^* \neq 0$ such that $\bar{M}(y(\delta^*))$ is well defined (see the proof of Part 2 of Lemma B.1). The latter statement is equivalent to the denominator in the fraction appearing in

the definition of $\bar{M}(y(\delta^*))$ being nonzero, i.e.,

$$\sum_{i=-(n-p-1)}^{n-p-1} w(i)\bar{\sigma}_i(y(\delta^*)) \neq 0,$$

where for $|i| = 0, \ldots, n-p-1$

$$\bar{\sigma}_i(y(\delta^*)) = (n-p)^{-1} \sum_{j=|i|+1}^{n-p} \bar{\omega}'[\hat{Z}_{\bar{X}}(y(\delta^*))]_{\cdot j}[\hat{Z}_{\bar{X}}(y(\delta^*))]'_{\cdot(j-|i|)}\bar{\omega}.$$

By definition $\bar{\omega} = (\omega', 0, 0)'$ and we recall that $\hat{Z}_{\bar{X}} = \hat{V}_{p,\bar{X}}(y(\delta))$ which implies via Equation (31) that $(\omega', 0, 0)\hat{Z}_{\bar{X}}(y(\delta))$ equals

$$(0_{k,p^*-p+2}, v_2(\delta)\omega'W'_{2\cdot}, 0_{k,p^*}, v_3(\delta)\omega'W'_{3\cdot}, 0_{k,p^*+1}, \ldots$$
$$\ldots, v_{k+4}(\delta)\omega'W'_{k+4\cdot}, 0_{k,p-1}, 0_{k,n-n^*}).$$

The only coordinate of this vector that depends on δ is $v_2(\delta)\omega'W'_{2\cdot} = \delta\sum_{i=1}^{k}\omega_i$, the latter equation following from $v_2(\delta) = \delta$ and $W_{2\cdot} = (1, \ldots, 1)$. Since $\sum_{j=1}^{k}\omega_i > 0$, the denominator appearing in the definition of $M(y(\delta))$ interpreted as a function of δ is now seen to be a polynomial of degree 2 in δ. Hence, there must exist a $\delta^* \neq 0$ such that the denominator does not vanish. It follows that $g^*_{\kappa,\bar{M},p}(y(\delta^*), \bar{X}, \bar{R}) \neq 0$.

Concerning the second statement we observe that $(k+2)(p^*+2)+p-1 \leq n$ implies $k+1 < n$, and therefore we have $\operatorname{rank}(\tilde{X}, e_+, e_-) = k+1$ for $\lambda_{\mathbb{R}^{n\times(k-1)}}$-almost every $\tilde{X} \in \tilde{\mathfrak{X}}_0$. By assumption the first column of R is zero. Therefore, for $\lambda_{\mathbb{R}^{n\times(k-1)}}$-almost every $X = (e_+, \tilde{X}) \in \mathbb{R}^{n\times k}$ we have $e_+ \in \mathfrak{M}_X$, $R\hat{\beta}_X(e_+) = 0$ and $e_- \notin \mathfrak{M}_X$, i.e., for $\lambda_{\mathbb{R}^{n\times(k-1)}}$-almost every $(e_+, \tilde{X}) \in \mathbb{R}^{n\times k}$ Scenario (1) in Theorem 5.4 applies. As above, it suffices to construct a pair $y \in \mathbb{R}^n$ and $\tilde{X} \in \mathbb{R}^{n\times(k-1)}$ (recall that $k \geq 2$), such that $g^*_{\kappa,\bar{M},p}(y, \bar{X}, \bar{R}) \neq 0$, where $\bar{X} = (e_+, \tilde{X}, e_-) \in \mathbb{R}^{n\times(k+1)}$ and $\bar{R} = (R, 0)$. Here, the matrix $\tilde{X}$ is $n \times (k-1)$ dimensional. By assumption $k^\sharp = k-1$ obviously satisfies $(k^\sharp + 3)(p^* + 2) + p - 1 + 1_{\mathbb{M}_{AM}}(M) \leq n$. To construct the matrix $\tilde{X}$ we can thus use the same argument as was used to construct X in the proof of the first statement ($k^\sharp$ replacing k). The matrix $\bar{X}$ so obtained has (after a permutation of its columns) the same structure as has the matrix $\bar{X}$ constructed in the proof of the first statement. We can therefore use almost the same arguments to conclude that $g^*_{\kappa,\bar{M},p}(y(\delta^*), \bar{X}, \bar{R}) \neq 0$ for some $\delta^* \neq 0$ and $y(\delta)$ as constructed in the first part of the proof. $\square$

Appendix E: Tables

In Tables 1, 2, and 3 rows correspond to ρ and columns correspond to β_1, whereas in Tables 4, 5, 6, 7, 8, 9, rows correspond to ρ and columns correspond to β_2.

TABLE 1
Rejection probabilities for test (i) in Example 1

	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1	1.5	2
-0.9999	0.18	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
-0.999	0.16	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
-0.99	0.13	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
-0.95	0.08	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
-0.9	0.07	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
-0.8	0.06	0.85	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
-0.7	0.06	0.66	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
-0.6	0.05	0.53	0.98	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
-0.5	0.05	0.43	0.93	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
-0.4	0.05	0.33	0.86	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
-0.3	0.06	0.29	0.79	0.98	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
-0.2	0.06	0.25	0.69	0.95	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
-0.1	0.06	0.22	0.60	0.92	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0	0.06	0.19	0.54	0.85	0.98	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.1	0.06	0.16	0.45	0.79	0.95	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.2	0.06	0.15	0.40	0.69	0.91	0.98	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.3	0.06	0.13	0.32	0.60	0.83	0.95	0.99	1.00	1.00	1.00	1.00	1.00	1.00
0.4	0.07	0.12	0.28	0.52	0.74	0.90	0.97	1.00	1.00	1.00	1.00	1.00	1.00
0.5	0.07	0.12	0.25	0.44	0.65	0.82	0.93	0.98	0.99	1.00	1.00	1.00	1.00
0.6	0.07	0.11	0.21	0.37	0.54	0.72	0.86	0.93	0.98	0.99	1.00	1.00	1.00
0.7	0.09	0.10	0.18	0.30	0.41	0.59	0.72	0.85	0.91	0.96	0.98	1.00	1.00
0.8	0.09	0.13	0.18	0.23	0.34	0.46	0.58	0.68	0.78	0.85	0.91	1.00	1.00
0.9	0.15	0.16	0.18	0.23	0.29	0.37	0.43	0.53	0.59	0.67	0.72	0.92	0.98
0.95	0.23	0.22	0.25	0.29	0.32	0.37	0.40	0.45	0.49	0.54	0.61	0.81	0.93
0.99	0.48	0.47	0.48	0.50	0.49	0.52	0.55	0.58	0.59	0.60	0.65	0.77	0.84
0.999	0.75	0.76	0.76	0.78	0.76	0.78	0.78	0.78	0.80	0.81	0.83	0.87	0.92
0.9999	0.89	0.89	0.90	0.89	0.89	0.91	0.91	0.91	0.91	0.93	0.92	0.95	0.97

TABLE 2
Rejection probabilities for test (ii) in Example 1

	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1	1.5	2
-0.9999	0.19	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
-0.999	0.17	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
-0.99	0.14	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
-0.95	0.09	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
-0.9	0.09	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
-0.8	0.07	0.86	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
-0.7	0.08	0.68	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
-0.6	0.07	0.55	0.98	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
-0.5	0.07	0.46	0.94	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
-0.4	0.07	0.37	0.88	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
-0.3	0.08	0.32	0.80	0.98	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
-0.2	0.09	0.28	0.71	0.96	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
-0.1	0.08	0.24	0.63	0.93	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0	0.07	0.21	0.56	0.86	0.98	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.1	0.08	0.17	0.47	0.80	0.95	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.2	0.07	0.16	0.42	0.70	0.92	0.98	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.3	0.06	0.15	0.34	0.63	0.84	0.95	0.99	1.00	1.00	1.00	1.00	1.00	1.00
0.4	0.07	0.13	0.29	0.54	0.75	0.91	0.98	1.00	1.00	1.00	1.00	1.00	1.00
0.5	0.07	0.13	0.25	0.45	0.66	0.83	0.94	0.98	0.99	1.00	1.00	1.00	1.00
0.6	0.07	0.11	0.22	0.38	0.55	0.73	0.87	0.94	0.98	0.99	1.00	1.00	1.00
0.7	0.10	0.11	0.19	0.30	0.41	0.59	0.72	0.85	0.91	0.96	0.99	1.00	1.00
0.8	0.09	0.13	0.18	0.24	0.34	0.46	0.59	0.69	0.78	0.85	0.91	1.00	1.00
0.9	0.15	0.16	0.18	0.23	0.29	0.38	0.43	0.53	0.59	0.67	0.72	0.92	0.98
0.95	0.23	0.22	0.25	0.29	0.32	0.36	0.40	0.46	0.49	0.54	0.61	0.81	0.93
0.99	0.47	0.47	0.48	0.50	0.49	0.51	0.55	0.58	0.59	0.60	0.64	0.77	0.84
0.999	0.75	0.76	0.76	0.78	0.76	0.78	0.78	0.78	0.80	0.81	0.83	0.87	0.92
0.9999	0.89	0.89	0.90	0.89	0.89	0.91	0.91	0.91	0.91	0.93	0.92	0.95	0.97

TABLE 3
Rejection probabilities for test (iii) in Example 1

	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1	1.5	2
-0.9999	0.02	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
-0.999	0.02	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
-0.99	0.02	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
-0.95	0.03	0.97	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
-0.9	0.04	0.86	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
-0.8	0.04	0.62	0.98	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
-0.7	0.05	0.44	0.92	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
-0.6	0.04	0.34	0.84	0.98	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
-0.5	0.04	0.29	0.75	0.96	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
-0.4	0.04	0.23	0.67	0.90	0.98	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
-0.3	0.05	0.20	0.57	0.85	0.97	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00
-0.2	0.05	0.17	0.48	0.79	0.94	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00
-0.1	0.04	0.15	0.42	0.73	0.90	0.96	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0	0.04	0.13	0.36	0.65	0.85	0.95	0.98	0.99	1.00	1.00	1.00	1.00	1.00
0.1	0.05	0.11	0.30	0.56	0.78	0.90	0.97	0.99	1.00	1.00	1.00	1.00	1.00
0.2	0.05	0.10	0.25	0.49	0.70	0.84	0.93	0.98	0.99	1.00	1.00	1.00	1.00
0.3	0.04	0.10	0.22	0.42	0.62	0.79	0.91	0.95	0.98	0.99	1.00	1.00	1.00
0.4	0.05	0.09	0.18	0.35	0.53	0.70	0.84	0.92	0.96	0.98	0.99	1.00	1.00
0.5	0.04	0.07	0.16	0.28	0.44	0.62	0.75	0.86	0.92	0.96	0.97	1.00	1.00
0.6	0.05	0.07	0.14	0.24	0.35	0.51	0.64	0.76	0.84	0.91	0.94	1.00	1.00
0.7	0.05	0.05	0.10	0.18	0.26	0.37	0.50	0.63	0.73	0.81	0.87	0.99	1.00
0.8	0.05	0.07	0.10	0.14	0.20	0.27	0.36	0.46	0.57	0.64	0.71	0.95	0.99
0.9	0.08	0.08	0.09	0.11	0.15	0.20	0.24	0.32	0.36	0.42	0.48	0.74	0.88
0.95	0.11	0.11	0.13	0.14	0.17	0.20	0.22	0.26	0.27	0.32	0.36	0.58	0.74
0.99	0.27	0.25	0.28	0.30	0.29	0.31	0.33	0.36	0.38	0.39	0.43	0.56	0.64
0.999	0.56	0.56	0.58	0.58	0.56	0.59	0.61	0.60	0.62	0.63	0.66	0.74	0.81
0.9999	0.78	0.77	0.77	0.77	0.78	0.80	0.81	0.81	0.82	0.83	0.83	0.89	0.92

TABLE 4
Rejection probabilities for test (i) in Example 2

	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1	1.5	2
-0.9999	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.87	0.88	0.88	0.88	0.88	0.89
-0.999	0.68	0.68	0.69	0.69	0.69	0.70	0.71	0.71	0.72	0.73	0.74	0.77	0.81
-0.99	0.50	0.51	0.51	0.53	0.55	0.58	0.61	0.63	0.66	0.69	0.71	0.83	0.89
-0.95	0.24	0.25	0.29	0.34	0.41	0.47	0.54	0.61	0.68	0.73	0.78	0.93	0.98
-0.9	0.17	0.19	0.25	0.33	0.42	0.53	0.64	0.73	0.80	0.86	0.90	0.99	1.00
-0.8	0.12	0.15	0.25	0.39	0.55	0.70	0.82	0.90	0.95	0.97	0.99	1.00	1.00
-0.7	0.10	0.15	0.30	0.50	0.70	0.84	0.93	0.98	0.99	1.00	1.00	1.00	1.00
-0.6	0.09	0.16	0.35	0.59	0.79	0.92	0.98	0.99	1.00	1.00	1.00	1.00	1.00
-0.5	0.08	0.16	0.39	0.67	0.87	0.96	0.99	1.00	1.00	1.00	1.00	1.00	1.00
-0.4	0.07	0.17	0.44	0.75	0.92	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00
-0.3	0.07	0.18	0.46	0.77	0.94	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00
-0.2	0.07	0.19	0.50	0.80	0.95	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00
-0.1	0.07	0.20	0.54	0.85	0.97	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0	0.07	0.19	0.53	0.84	0.97	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.1	0.07	0.20	0.52	0.83	0.97	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.2	0.07	0.19	0.51	0.82	0.96	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.3	0.07	0.18	0.46	0.76	0.93	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.4	0.08	0.18	0.45	0.74	0.92	0.98	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.5	0.08	0.16	0.39	0.67	0.87	0.96	0.99	1.00	1.00	1.00	1.00	1.00	1.00
0.6	0.09	0.16	0.35	0.60	0.80	0.93	0.98	0.99	1.00	1.00	1.00	1.00	1.00
0.7	0.10	0.15	0.29	0.50	0.70	0.84	0.93	0.97	0.99	1.00	1.00	1.00	1.00
0.8	0.12	0.15	0.25	0.40	0.56	0.71	0.83	0.91	0.95	0.98	0.99	1.00	1.00
0.9	0.16	0.19	0.24	0.33	0.43	0.54	0.64	0.73	0.80	0.86	0.91	0.98	1.00
0.95	0.22	0.24	0.26	0.32	0.38	0.45	0.53	0.60	0.67	0.73	0.78	0.93	0.98
0.99	0.28	0.28	0.31	0.34	0.38	0.43	0.48	0.53	0.58	0.63	0.68	0.84	0.91
0.999	0.35	0.35	0.37	0.39	0.42	0.46	0.50	0.54	0.58	0.62	0.65	0.79	0.87
0.9999	0.37	0.37	0.39	0.41	0.44	0.48	0.52	0.56	0.60	0.63	0.67	0.80	0.88

TABLE 5
Rejection probabilities for the adjusted version of test (i) in Example 2

	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1	1.5	2
-0.9999	0.06	0.06	0.07	0.09	0.12	0.15	0.18	0.22	0.26	0.30	0.34	0.52	0.65
-0.999	0.06	0.06	0.08	0.09	0.12	0.16	0.19	0.23	0.27	0.31	0.35	0.53	0.66
-0.99	0.05	0.05	0.07	0.09	0.12	0.15	0.20	0.24	0.28	0.33	0.38	0.59	0.73
-0.95	0.04	0.04	0.06	0.09	0.13	0.18	0.24	0.30	0.37	0.43	0.49	0.74	0.87
-0.9	0.03	0.04	0.06	0.10	0.16	0.22	0.30	0.38	0.47	0.55	0.62	0.85	0.94
-0.8	0.03	0.05	0.10	0.18	0.30	0.43	0.57	0.69	0.78	0.85	0.90	0.99	1.00
-0.7	0.04	0.06	0.15	0.30	0.49	0.67	0.81	0.90	0.95	0.97	0.99	1.00	1.00
-0.6	0.04	0.07	0.20	0.41	0.63	0.81	0.91	0.97	0.99	1.00	1.00	1.00	1.00
-0.5	0.04	0.09	0.25	0.52	0.76	0.91	0.97	0.99	1.00	1.00	1.00	1.00	1.00
-0.4	0.04	0.10	0.31	0.61	0.84	0.95	0.99	1.00	1.00	1.00	1.00	1.00	1.00
-0.3	0.03	0.10	0.32	0.63	0.86	0.96	0.99	1.00	1.00	1.00	1.00	1.00	1.00
-0.2	0.04	0.12	0.38	0.71	0.91	0.98	1.00	1.00	1.00	1.00	1.00	1.00	1.00
-0.1	0.03	0.12	0.41	0.75	0.94	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0	0.04	0.12	0.40	0.74	0.93	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.1	0.04	0.12	0.39	0.73	0.93	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.2	0.03	0.11	0.36	0.69	0.91	0.98	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.3	0.03	0.10	0.31	0.61	0.84	0.95	0.99	1.00	1.00	1.00	1.00	1.00	1.00
0.4	0.03	0.10	0.30	0.60	0.84	0.95	0.99	1.00	1.00	1.00	1.00	1.00	1.00
0.5	0.03	0.09	0.25	0.50	0.74	0.89	0.96	0.99	1.00	1.00	1.00	1.00	1.00
0.6	0.03	0.06	0.18	0.37	0.58	0.77	0.88	0.95	0.98	0.99	1.00	1.00	1.00
0.7	0.03	0.06	0.15	0.30	0.49	0.67	0.81	0.90	0.95	0.98	0.99	1.00	1.00
0.8	0.03	0.05	0.10	0.20	0.32	0.46	0.60	0.72	0.81	0.88	0.93	0.99	1.00
0.9	0.03	0.05	0.07	0.12	0.18	0.25	0.34	0.42	0.51	0.59	0.67	0.89	0.97
0.95	0.04	0.04	0.06	0.09	0.12	0.17	0.22	0.28	0.34	0.40	0.46	0.70	0.85
0.99	0.05	0.05	0.06	0.09	0.11	0.14	0.18	0.22	0.26	0.31	0.35	0.56	0.70
0.999	0.06	0.06	0.07	0.09	0.11	0.14	0.17	0.20	0.24	0.27	0.31	0.48	0.62
0.9999	0.06	0.06	0.07	0.09	0.11	0.14	0.17	0.21	0.24	0.28	0.32	0.49	0.62

TABLE 6
Rejection probabilities for test (ii) in Example 2

	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1	1.5	2
-0.9999	0.87	0.88	0.88	0.87	0.88	0.87	0.87	0.87	0.87	0.88	0.88	0.88	0.89
-0.999	0.68	0.68	0.68	0.68	0.69	0.70	0.70	0.71	0.72	0.73	0.73	0.77	0.80
-0.99	0.50	0.50	0.51	0.53	0.55	0.58	0.60	0.63	0.66	0.69	0.71	0.82	0.89
-0.95	0.25	0.26	0.29	0.34	0.41	0.47	0.55	0.62	0.68	0.73	0.79	0.93	0.98
-0.9	0.17	0.20	0.25	0.33	0.43	0.54	0.64	0.73	0.80	0.86	0.90	0.99	1.00
-0.8	0.13	0.16	0.26	0.40	0.56	0.71	0.82	0.90	0.95	0.98	0.99	1.00	1.00
-0.7	0.10	0.16	0.31	0.51	0.71	0.85	0.94	0.98	0.99	1.00	1.00	1.00	1.00
-0.6	0.10	0.17	0.37	0.61	0.81	0.93	0.98	1.00	1.00	1.00	1.00	1.00	1.00
-0.5	0.09	0.18	0.41	0.69	0.88	0.97	0.99	1.00	1.00	1.00	1.00	1.00	1.00
-0.4	0.09	0.19	0.46	0.76	0.93	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00
-0.3	0.08	0.20	0.49	0.78	0.94	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00
-0.2	0.08	0.21	0.52	0.82	0.96	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00
-0.1	0.08	0.22	0.56	0.86	0.98	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0	0.08	0.21	0.55	0.85	0.97	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.1	0.08	0.21	0.54	0.85	0.97	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.2	0.08	0.21	0.53	0.83	0.97	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.3	0.08	0.19	0.48	0.78	0.94	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.4	0.09	0.20	0.47	0.76	0.93	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.5	0.09	0.18	0.42	0.69	0.88	0.97	0.99	1.00	1.00	1.00	1.00	1.00	1.00
0.6	0.10	0.17	0.37	0.62	0.82	0.93	0.98	0.99	1.00	1.00	1.00	1.00	1.00
0.7	0.11	0.16	0.31	0.52	0.71	0.85	0.94	0.98	0.99	1.00	1.00	1.00	1.00
0.8	0.12	0.16	0.26	0.41	0.57	0.72	0.84	0.91	0.95	0.98	0.99	1.00	1.00
0.9	0.16	0.19	0.24	0.33	0.43	0.54	0.65	0.73	0.81	0.86	0.91	0.98	1.00
0.95	0.22	0.24	0.26	0.32	0.38	0.45	0.53	0.60	0.67	0.73	0.78	0.93	0.97
0.99	0.27	0.28	0.30	0.34	0.38	0.43	0.48	0.53	0.58	0.63	0.67	0.84	0.91
0.999	0.35	0.35	0.37	0.39	0.42	0.46	0.50	0.54	0.58	0.62	0.65	0.79	0.87
0.9999	0.37	0.37	0.39	0.41	0.44	0.48	0.52	0.55	0.60	0.63	0.67	0.80	0.87

TABLE 7
Rejection probabilities for the adjusted version of test (ii) in Example 2

	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1	1.5	2
-0.9999	0.06	0.06	0.08	0.09	0.12	0.15	0.18	0.22	0.26	0.29	0.34	0.52	0.64
-0.999	0.05	0.06	0.07	0.09	0.11	0.15	0.18	0.22	0.26	0.30	0.34	0.51	0.65
-0.99	0.05	0.05	0.07	0.09	0.12	0.15	0.19	0.24	0.28	0.33	0.37	0.58	0.73
-0.95	0.04	0.04	0.06	0.09	0.13	0.18	0.24	0.30	0.36	0.42	0.49	0.73	0.87
-0.9	0.03	0.04	0.06	0.10	0.16	0.22	0.30	0.38	0.46	0.54	0.61	0.85	0.94
-0.8	0.03	0.05	0.10	0.18	0.30	0.44	0.57	0.69	0.79	0.85	0.90	0.99	1.00
-0.7	0.03	0.06	0.15	0.29	0.47	0.66	0.80	0.89	0.95	0.97	0.99	1.00	1.00
-0.6	0.04	0.07	0.20	0.40	0.62	0.80	0.91	0.97	0.99	1.00	1.00	1.00	1.00
-0.5	0.03	0.08	0.24	0.49	0.74	0.90	0.97	0.99	1.00	1.00	1.00	1.00	1.00
-0.4	0.04	0.10	0.30	0.59	0.83	0.95	0.99	1.00	1.00	1.00	1.00	1.00	1.00
-0.3	0.04	0.10	0.32	0.62	0.85	0.96	0.99	1.00	1.00	1.00	1.00	1.00	1.00
-0.2	0.04	0.11	0.35	0.67	0.89	0.98	1.00	1.00	1.00	1.00	1.00	1.00	1.00
-0.1	0.03	0.11	0.37	0.71	0.92	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0	0.04	0.11	0.38	0.72	0.92	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.1	0.04	0.12	0.37	0.70	0.92	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.2	0.03	0.10	0.33	0.66	0.89	0.98	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.3	0.03	0.09	0.28	0.57	0.81	0.94	0.99	1.00	1.00	1.00	1.00	1.00	1.00
0.4	0.04	0.10	0.30	0.59	0.83	0.95	0.99	1.00	1.00	1.00	1.00	1.00	1.00
0.5	0.03	0.08	0.24	0.49	0.73	0.88	0.96	0.99	1.00	1.00	1.00	1.00	1.00
0.6	0.03	0.06	0.17	0.36	0.57	0.76	0.88	0.94	0.97	0.99	0.99	1.00	1.00
0.7	0.04	0.06	0.15	0.30	0.49	0.67	0.81	0.90	0.95	0.98	0.99	1.00	1.00
0.8	0.03	0.05	0.10	0.19	0.31	0.45	0.59	0.71	0.81	0.88	0.92	0.99	1.00
0.9	0.03	0.05	0.07	0.12	0.18	0.25	0.34	0.42	0.51	0.59	0.67	0.89	0.97
0.95	0.04	0.04	0.06	0.09	0.12	0.17	0.22	0.28	0.34	0.40	0.45	0.70	0.85
0.99	0.05	0.05	0.06	0.09	0.11	0.14	0.18	0.22	0.26	0.31	0.35	0.56	0.70
0.999	0.06	0.06	0.07	0.09	0.11	0.14	0.17	0.20	0.24	0.27	0.31	0.48	0.62
0.9999	0.06	0.06	0.07	0.09	0.11	0.13	0.16	0.20	0.24	0.27	0.31	0.48	0.61

TABLE 8
Rejection probabilities for test (iii) in Example 2

	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1	1.5	2
-0.9999	0.65	0.64	0.65	0.65	0.66	0.66	0.66	0.67	0.68	0.68	0.69	0.72	0.75
-0.999	0.47	0.47	0.47	0.48	0.49	0.50	0.51	0.52	0.53	0.55	0.56	0.62	0.67
-0.99	0.32	0.32	0.33	0.35	0.37	0.40	0.42	0.46	0.48	0.51	0.54	0.67	0.76
-0.95	0.14	0.15	0.18	0.22	0.27	0.33	0.40	0.46	0.52	0.58	0.63	0.83	0.92
-0.9	0.09	0.10	0.14	0.21	0.28	0.37	0.46	0.54	0.62	0.70	0.75	0.92	0.97
-0.8	0.07	0.09	0.17	0.27	0.40	0.53	0.65	0.75	0.83	0.89	0.93	0.99	1.00
-0.7	0.07	0.10	0.20	0.35	0.53	0.68	0.80	0.88	0.93	0.96	0.98	1.00	1.00
-0.6	0.06	0.10	0.24	0.42	0.61	0.77	0.87	0.93	0.97	0.98	0.99	1.00	1.00
-0.5	0.06	0.11	0.27	0.49	0.70	0.84	0.93	0.97	0.99	0.99	1.00	1.00	1.00
-0.4	0.06	0.12	0.31	0.55	0.76	0.89	0.95	0.98	0.99	1.00	1.00	1.00	1.00
-0.3	0.05	0.12	0.32	0.58	0.78	0.90	0.96	0.99	1.00	1.00	1.00	1.00	1.00
-0.2	0.05	0.13	0.34	0.60	0.80	0.92	0.97	0.99	1.00	1.00	1.00	1.00	1.00
-0.1	0.05	0.14	0.38	0.66	0.85	0.94	0.98	1.00	1.00	1.00	1.00	1.00	1.00
0	0.05	0.13	0.37	0.64	0.84	0.94	0.98	0.99	1.00	1.00	1.00	1.00	1.00
0.1	0.05	0.14	0.36	0.64	0.83	0.93	0.98	0.99	1.00	1.00	1.00	1.00	1.00
0.2	0.05	0.13	0.35	0.62	0.82	0.93	0.98	0.99	1.00	1.00	1.00	1.00	1.00
0.3	0.05	0.12	0.31	0.56	0.77	0.90	0.96	0.98	0.99	1.00	1.00	1.00	1.00
0.4	0.06	0.12	0.32	0.56	0.76	0.89	0.95	0.98	0.99	1.00	1.00	1.00	1.00
0.5	0.06	0.12	0.28	0.50	0.70	0.84	0.92	0.96	0.98	0.99	1.00	1.00	1.00
0.6	0.06	0.11	0.24	0.44	0.63	0.78	0.88	0.94	0.97	0.98	0.99	1.00	1.00
0.7	0.07	0.10	0.21	0.36	0.53	0.68	0.80	0.88	0.93	0.96	0.98	1.00	1.00
0.8	0.08	0.10	0.17	0.28	0.42	0.55	0.67	0.77	0.85	0.90	0.94	0.99	1.00
0.9	0.11	0.12	0.17	0.23	0.32	0.41	0.51	0.59	0.67	0.73	0.79	0.93	0.98
0.95	0.13	0.14	0.16	0.20	0.26	0.31	0.38	0.45	0.51	0.57	0.63	0.82	0.91
0.99	0.16	0.17	0.18	0.21	0.24	0.28	0.32	0.37	0.41	0.45	0.50	0.68	0.77
0.999	0.22	0.22	0.24	0.25	0.28	0.32	0.35	0.39	0.42	0.46	0.49	0.64	0.73
0.9999	0.23	0.23	0.25	0.27	0.30	0.33	0.36	0.40	0.44	0.46	0.50	0.64	0.74

TABLE 9
Rejection probabilities for the adjusted version of test (iii) in Example 2

	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1	1.5	2
-0.9999	0.06	0.06	0.07	0.09	0.11	0.15	0.17	0.21	0.25	0.28	0.32	0.50	0.62
-0.999	0.05	0.06	0.07	0.09	0.11	0.14	0.17	0.21	0.25	0.29	0.32	0.50	0.63
-0.99	0.05	0.05	0.07	0.09	0.12	0.15	0.19	0.24	0.28	0.32	0.37	0.57	0.71
-0.95	0.04	0.04	0.06	0.09	0.13	0.18	0.24	0.30	0.36	0.43	0.49	0.73	0.87
-0.9	0.03	0.04	0.07	0.11	0.16	0.23	0.31	0.39	0.47	0.54	0.61	0.85	0.94
-0.8	0.03	0.05	0.10	0.19	0.31	0.44	0.58	0.70	0.79	0.86	0.90	0.99	1.00
-0.7	0.04	0.06	0.15	0.30	0.49	0.67	0.81	0.90	0.95	0.98	0.99	1.00	1.00
-0.6	0.04	0.07	0.20	0.39	0.61	0.80	0.91	0.97	0.99	1.00	1.00	1.00	1.00
-0.5	0.04	0.09	0.25	0.50	0.75	0.90	0.97	0.99	1.00	1.00	1.00	1.00	1.00
-0.4	0.04	0.10	0.30	0.60	0.83	0.95	0.99	1.00	1.00	1.00	1.00	1.00	1.00
-0.3	0.04	0.10	0.32	0.61	0.85	0.95	0.99	1.00	1.00	1.00	1.00	1.00	1.00
-0.2	0.04	0.12	0.35	0.67	0.89	0.98	1.00	1.00	1.00	1.00	1.00	1.00	1.00
-0.1	0.04	0.12	0.39	0.72	0.92	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0	0.04	0.12	0.38	0.71	0.92	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.1	0.04	0.12	0.37	0.70	0.91	0.98	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.2	0.03	0.10	0.34	0.66	0.88	0.97	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.3	0.03	0.10	0.29	0.58	0.82	0.95	0.99	1.00	1.00	1.00	1.00	1.00	1.00
0.4	0.04	0.10	0.29	0.57	0.81	0.94	0.99	1.00	1.00	1.00	1.00	1.00	1.00
0.5	0.04	0.09	0.24	0.48	0.72	0.88	0.96	0.99	1.00	1.00	1.00	1.00	1.00
0.6	0.03	0.07	0.19	0.38	0.59	0.77	0.89	0.95	0.98	0.99	0.99	1.00	1.00
0.7	0.04	0.07	0.15	0.31	0.49	0.67	0.81	0.89	0.95	0.97	0.99	1.00	1.00
0.8	0.03	0.05	0.10	0.20	0.32	0.46	0.60	0.72	0.82	0.88	0.93	0.99	1.00
0.9	0.04	0.05	0.07	0.12	0.18	0.25	0.34	0.42	0.50	0.58	0.66	0.88	0.96
0.95	0.04	0.05	0.06	0.09	0.13	0.17	0.22	0.28	0.34	0.40	0.45	0.70	0.84
0.99	0.05	0.05	0.06	0.08	0.11	0.14	0.17	0.21	0.25	0.30	0.34	0.54	0.68
0.999	0.06	0.06	0.07	0.09	0.11	0.14	0.17	0.20	0.24	0.27	0.30	0.47	0.60
0.9999	0.06	0.06	0.07	0.09	0.11	0.14	0.17	0.20	0.23	0.26	0.31	0.47	0.60

Acknowledgements

References

ANDREWS, D. W. K. (1991). Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation. *Econometrica* **59** pp. 817–858. MR1106513

ANDREWS, D. W. K. and MONAHAN, J. C. (1992). An Improved Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimator. *Econometrica* **60** pp. 953–966. MR1168742

BARTLETT, M. S. (1950). Periodogram Analysis and Continuous Spectra. *Biometrika* **37** 1–16. MR0035934

DEN HAAN, W. J. and LEVIN, A. T. (1997). A practitioner's guide to robust covariance matrix estimation. In *Robust Inference*, (G. S. Maddala and C. R. Rao, eds.). *Handbook of Statistics* **15** 299–342. Elsevier. MR1492717

GRENANDER, U. and ROSENBLATT, M. (1957). *Statistical analysis of stationary time series*. John Wiley & Sons, New York. MR0084975 (18,959b)

HANNAN, E. J. (1957). The Variance of the Mean of a Stationary Process. *Journal of the Royal Statistical Society. Series B* **19** 282–285. MR0092333

HANSEN, B. E. (1992). Consistent Covariance Matrix Estimation for Dependent Heterogeneous Processes. *Econometrica* **60** pp. 967–972. MR1168743

JANSSON, M. (2002). Consistent covariance matrix estimation for linear processes. *Econometric Theory* **18** 1449–1459. MR1945420

JANSSON, M. (2004). The Error in Rejection Probability of Simple Autocorrelation Robust Tests. *Econometrica* **72** 937–946. MR2051441

JOWETT, G. H. (1955). The Comparison of Means of Sets of Observations from Sections of Independent Stochastic Series. *Journal of the Royal Statistical Society. Series B* **17** 208–227. MR0076248

KIEFER, N. M., VOGELSANG, T. J. and BUNZEL, H. (2000). Simple robust testing of regression hypotheses. *Econometrica* **68** 695–714. MR1769382

KIEFER, N. M. and VOGELSANG, T. J. (2002). Heteroskedasticity-autocorrelation robust testing using bandwidth equal to sample size. *Econometric Theory* **18** 1350–1366. MR1945416 (2003k:62230)

KIEFER, N. M. and VOGELSANG, T. J. (2005). A new asymptotic theory for heteroskedasticity-autocorrelation robust tests. *Econometric Theory* **21** 1130–1164. MR2200988

NEWEY, W. K. and WEST, K. D. (1987). A simple, positive semidefinite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* **55** 703–708. MR890864 (88g:62111)

NEWEY, W. K. and WEST, K. D. (1994). Automatic Lag Selection in Covariance Matrix Estimation. *The Review of Economic Studies* **61** pp. 631–653. MR1299308

OKAMOTO, M. (1973). Distinctness of the eigenvalues of a quadratic form in a multivariate sample. *The Annals of Statistics* 763–765. MR0331643

PREINERSTORFER, D. and PÖTSCHER, B. M. (2016). On Size and Power of Heteroskedasticity and Autocorrelation Robust Tests. *Econometric Theory* **32** 261–358. MR3471120

PREINERSTORFER, D. and PÖTSCHER, B. M. (2017). On the power of invariant tests for hypotheses on a covariance matrix. *Econometric Theory* **33** 1–68. MR3574860

RHO, Y. and SHAO, X. (2013). Improving the bandwidth-free inference methods by prewhitening. *Journal of Statistical Planning and Inference* **143** 1912–1922. MR3095081

ROBINSON, G. (1979). Conditional properties of statistical procedures. *The Annals of Statistics* 742–755. MR0532239

SCHWERT, G. (2009). EViews 7 User's Guide II. *Quantitative Micro Software, LLC, Irvine, California*.

SUN, Y., PHILLIPS, P. C. B. and JIN, S. (2008). Optimal Bandwidth Selection in Heteroskedasticity-Autocorrelation Robust Testing. *Econometrica* **76** 175–194. MR2374985

Sun, Y., Phillips, P. C. B. and Jin, S. (2011). Power maximization and size control in heteroskedasticity and autocorrelation robust tests with exponentiated kernels. *Econometric Theory* **27** 1320–1368. MR2868841

R Core Team (2016). R: A Language and Environment for Statistical Computing R Foundation for Statistical Computing, Vienna, Austria.

Velasco, C. and Robinson, P. M. (2001). Edgeworth Expansions for Spectral Density Estimates and Studentized Sample Mean. *Econometric Theory* **17** 497–539. MR1841819

Zeileis, A. (2004). Econometric Computing with HC and HAC Covariance Matrix Estimators. *Journal of Statistical Software* **11** 1–17.

Zeileis, A. and Hothorn, T. (2002). Diagnostic Checking in Regression Relationships. *R News* **2** 7–10.

Zhang, X. and Shao, X. (2013). Fixed-smoothing asymptotics for time series. *The Annals of Statistics* **41** 1329–1349. MR3113813