Supplementary Material to "Valid Heteroskedasticity Robust Testing"*

Benedikt M. Pötscher and David Preinerstorfer
Department of Statistics, University of Vienna
SEPS-SEW, University of St. Gallen

Preliminary draft: October 2017
First version: April 2021
Second version: November 2022
This version: July 2023

Abstract

This document contains the Appendices A-G to the article "Valid Heteroskedasticity Robust Testing".

*Address correspondence to Benedikt Pötscher, Department of Statistics, University of Vienna, A-1090 Oskar-Morgenstern Platz 1. E-Mail: benedikt.poetscher@univie.ac.at.

Contents

A Appendix: Size control over other heteroskedasticity models

As already noted earlier, if size control is possible over $\mathfrak{C}_{Het}$, then the same is true over *any* conceivable class of heteroskedasticity structures, since these can (possible after normalization) be cast as a subset $\mathfrak{C}$ of $\mathfrak{C}_{Het}$; and, in fact, any critical value delivering size control over $\mathfrak{C}_{Het}$ also delivers size control over any such $\mathfrak{C}$, but even smaller critical values may already suffice for size control over $\mathfrak{C}$. Also, for some heteroskedasticity models $\mathfrak{C} \subseteq \mathfrak{C}_{Het}$, the sufficient conditions employed in Theorems 5.1 and 6.4 (which imply size control over $\mathfrak{C}_{Het}$) may be unnecessarily restrictive, if one wants to establish size control over $\mathfrak{C}$ only. For this reason, we show in the following how the general theory laid out in Section 5 of Pötscher and Preinerstorfer (2018) can be used to derive size control results tailored to various subsets $\mathfrak{C}$ by exemplarily treating the cases $\mathfrak{C} = \mathfrak{C}_{(n_1,\ldots,n_m)}$ and $\mathfrak{C} = \mathfrak{C}_{Het,\tau_*}$ defined below. Size control results over other choices of $\mathfrak{C}$ can be derived from the results in Section 5 of Pötscher and Preinerstorfer (2018) in a similar manner, see Subsection A.1.2 further below for some discussion. Here $\mathfrak{C}_{(n_1,\ldots,n_m)}$ is defined as follows: Let $m \in \mathbb{N}$, and let $n_j \in \mathbb{N}$ for $j = 1, \ldots, m$ satisfy $\sum_{j=1}^{m} n_j = n$. Set $n_j^+ = \sum_{l=1}^{j} n_l$ and define

$$\mathfrak{C}_{(n_1,\ldots,n_m)} = \left\{ \operatorname{diag}(\tau_1^2, \ldots, \tau_n^2) \in \mathfrak{C}_{Het} : \tau^2_{n_{j-1}^+ + 1} = \ldots = \tau^2_{n_j^+} \text{ for } j = 1, \ldots, m \right\}$$

with the convention that $n_0^+ = 0$. This may be a natural heteroskedasticity model when the observations come from m groups and when it is reasonable to assume homoskedasticity within groups.[63] Note that in case $n_j = 1$ for all j, then $m = n$ and $\mathfrak{C}_{(n_1,\ldots,n_m)} = \mathfrak{C}_{Het}$ hold; and in case $m = 1$ we have $\mathfrak{C}_{(n_1,\ldots,n_m)} = \{n^{-1} I_n\}$, i.e., we have homoskedasticity. Furthermore, $\mathfrak{C}_{Het,\tau_*}$ is given by

$$\mathfrak{C}_{Het,\tau_*} = \left\{ \operatorname{diag}(\tau_1^2, \ldots, \tau_n^2) \in \mathfrak{C}_{Het} : \tau_i^2 \geq \tau_*^2 \text{ for all } i \right\},$$

where the lower bound τ_*, $0 < \tau_* < n^{-1/2}$, is set by the user.

A.1 Size control results for T_{Het} and T_{uc}

A.1.1 Size control over $\mathfrak{C}_{(n_1,\ldots,n_m)}$

Proofs of the results in this subsection can be found in Appendix C.

Theorem A.1. *Let $m \in \mathbb{N}$, and let $n_j \in \mathbb{N}$ for $j = 1, \ldots, m$ satisfy $\sum_{j=1}^{m} n_j = n$. Then:*

(a) For every $0 < \alpha < 1$ there exists a real number $C(\alpha)$ such that

$$\sup_{\mu_0 \in \mathfrak{M}_0} \sup_{0 < \sigma^2 < \infty} \sup_{\Sigma \in \mathfrak{C}_{(n_1,\ldots,n_m)}} P_{\mu_0,\sigma^2\Sigma}(T_{uc} \geq C(\alpha)) \leq \alpha \tag{A.1}$$

[63] As long as we assume that the grouping is known, there is little loss of generality to assume that the elements belonging to the same group are numbered contiguously, since we otherwise only need to relabel the data.

holds, provided that

$$\operatorname{span}\left(\left\{e_i(n) : i \in (n_{j-1}^+, n_j^+]\right\}\right) \nsubseteq \operatorname{span}(X) \quad \text{for every } j = 1, \ldots, m \text{ with } (n_{j-1}^+, n_j^+] \cap I_1(\mathfrak{M}_0^{lin}) \neq \emptyset. \tag{A.2}$$

Furthermore, under condition (A.2), even equality can be achieved in (A.1) by a proper choice of $C(\alpha)$*, provided* $\alpha \in (0, \alpha^*] \cap (0, 1)$ *holds, where* $\alpha^* = \sup_{C \in (C^*, \infty)} \sup_{\Sigma \in \mathfrak{C}_{(n_1,\ldots,n_m)}} P_{\mu_0,\Sigma}(T_{uc} \geq C)$ *is positive and where* C^* *is defined as in Lemma 5.11 of Pötscher and Preinerstorfer (2018) with* $\mathfrak{C} = \mathfrak{C}_{(n_1,\ldots,n_m)}$*,* $T = T_{uc}$*,* $N^\dagger = \operatorname{span}(X)$*, and* $\mathcal{L} = \mathfrak{M}_0^{lin}$ *(with neither* α^* *nor* C^* *depending on the choice of* $\mu_0 \in \mathfrak{M}_0$*).*

(b) Suppose Assumption 1 is satisfied. Then for every $0 < \alpha < 1$ *there exists a real number* $C(\alpha)$ *such that*

$$\sup_{\mu_0 \in \mathfrak{M}_0} \sup_{0 < \sigma^2 < \infty} \sup_{\Sigma \in \mathfrak{C}_{(n_1,\ldots,n_m)}} P_{\mu_0,\sigma^2\Sigma}(T_{Het} \geq C(\alpha)) \leq \alpha \tag{A.3}$$

holds, provided that

$$\operatorname{span}\left(\left\{e_i(n) : i \in (n_{j-1}^+, n_j^+]\right\}\right) \nsubseteq \mathsf{B} \quad \text{for every } j = 1, \ldots, m \text{ with } (n_{j-1}^+, n_j^+] \cap I_1(\mathfrak{M}_0^{lin}) \neq \emptyset. \tag{A.4}$$

Furthermore, under condition (A.4), even equality can be achieved in (A.3) by a proper choice of $C(\alpha)$*, provided* $\alpha \in (0, \alpha^*] \cap (0, 1)$ *holds, where* $\alpha^* = \sup_{C \in (C^*, \infty)} \sup_{\Sigma \in \mathfrak{C}_{(n_1,\ldots,n_m)}} P_{\mu_0,\Sigma}(T_{Het} \geq C)$ *is positive and where* C^* *is defined as in Lemma 5.11 of Pötscher and Preinerstorfer (2018) with* $\mathfrak{C} = \mathfrak{C}_{(n_1,\ldots,n_m)}$*,* $T = T_{Het}$*,* $N^\dagger = \mathsf{B}$*, and* $\mathcal{L} = \mathfrak{M}_0^{lin}$ *(with neither* α^* *nor* C^* *depending on the choice of* $\mu_0 \in \mathfrak{M}_0$*).*

(c) Under the assumptions of Part (a) (Part (b), respectively) implying existence of a critical value $C(\alpha)$ *satisfying (A.1) ((A.3), respectively), a smallest critical value, denoted by* $C_\Diamond(\alpha)$*, satisfying (A.1) ((A.3), respectively) exists for every* $0 < \alpha < 1$*. And* $C_\Diamond(\alpha)$ *corresponding to Part (a) (Part (b), respectively) is also the smallest among the critical values leading to equality in (A.1) ((A.3), respectively) whenever such critical values exist. [Although* $C_\Diamond(\alpha)$ *corresponding to Part (a) and (b), respectively, will typically be different, we use the same symbol.]*[64]

It is easy to see that the discussion in the first paragraph following Theorem 5.1 applies mutatis mutandis also to the above theorem. Similarly, Remarks 5.2, 5.3, 5.4, 5.6, 5.9, and Proposition 5.5 carry over. Furthermore, we have the following result corresponding to Proposition 5.7:

Proposition A.2. *(a) If (A.2) is violated, then* $\sup_{\Sigma \in \mathfrak{C}_{(n_1,\ldots,n_m)}} P_{\mu_0,\sigma^2\Sigma}(T_{uc} \geq C) = 1$ *for* every *choice of critical value* C*, every* $\mu_0 \in \mathfrak{M}_0$*, and every* $\sigma^2 \in (0, \infty)$ *(implying that size equals* 1 *for every* C*). As a consequence, the sufficient condition for size control (A.2) in Part (a) of Theorem A.1 is also necessary.*

(b) Suppose Assumption 1 is satisfied.[65] *If (A.2) is violated, then* $\sup_{\Sigma \in \mathfrak{C}_{(n_1,\ldots,n_m)}} P_{\mu_0,\sigma^2\Sigma}(T_{Het} \geq$

[64]Cf. also Appendix A.3.

[65]If this assumption is violated then T_{Het} is identically zero, an uninteresting trivial case.

C) = 1 for every *choice of critical value C, every $\mu_0 \in \mathfrak{M}_0$, and every $\sigma^2 \in (0,\infty)$ (implying that size equals* 1 *for every C). [In case X and R are such that $\mathsf{B} = \operatorname{span}(X)$, conditions (A.2) and (A.4) coincide; hence the sufficient condition for size control (A.4) in Part (b) of Theorem A.1 is then also necessary in this case.]*

Remark A.3. *(Homoskedasticity)* Theorem A.1 allows also for the case $m = 1$, in which case $\mathfrak{C}_{(n_1,\ldots,n_m)} = \{n^{-1}I_n\}$, i.e., errors are homoskedastic. In this case it is easy to see that the sufficient conditions for size control in the theorem are trivially satisfied and size control for T_{Het} (and T_{uc}) is possible.[66] Of course, this is in line with the fact that T_{Het} and T_{uc} are obviously pivotal under the null if the errors are homoskedastic.

Remark A.4. *(Behrens-Fisher problem)* Consider again the problem of testing the equality of the means of two independent normal populations as in Example 5.4 with the only difference that the variance within each of the two groups is now assumed to be constant, i.e., the heteroskedasticity model used is now given by $\mathfrak{C}_{(n_1,n_2)}$, where $n_1 \geq 2$ and $n_2 \geq 2$ are the group sizes. This is the celebrated Behrens-Fisher problem. The square of the two-sample t-statistic t_{FB}, say, often used in this context coincides with T_{Het} for the choice $d_i = (1 - h_{ii})^{-1}$. The size controllability of T_{Het} over $\mathfrak{C}_{Het}$ established in Example 5.4 therefore a fortiori implies size controllability of T_{Het} (and hence of t_{FB}^2) over $\mathfrak{C}_{(n_1,n_2)}$. Of course, this does not add anything new to the literature on the Behrens-Fisher problem, since it is known that under the null hypothesis $|t_{FB}|$ is stochastically not larger than a t-distributed random variable with $\min(n_1, n_2) - 1$ degrees of freedom when $\mathfrak{C}_{(n_1,n_2)}$ is the heteroskedasticity model, see Mickey and Brown (1966). For more on the Behrens-Fisher problem see Kim and Cohen (1998), Ruben (2002), Lehmann and Romano (2005), Belloni and Didier (2008), and the references cited therein.

A.1.2 Further size control results

In this subsection it is understood that Assumption 1 is maintained when discussing results relating to T_{Het}.

(i) Given a heteroskedasticity model $\mathfrak{C}$ (i.e., $\emptyset \neq \mathfrak{C} \subseteq \mathfrak{C}_{Het}$), with the property that $\mathbb{J}(\mathfrak{M}_0^{lin}, \mathfrak{C})$ is empty (where the collection $\mathbb{J}(\mathfrak{M}_0^{lin}, \mathfrak{C})$ is defined on p. 421 of Pötscher and Preinerstorfer (2018), see also Appendix B further below), the tests based on T_{uc} and T_{Het} are always size controllable over $\mathfrak{C}$. This follows from Corollary 5.6 and Remark 5.7 in Pötscher and Preinerstorfer (2018). In fact, exact size control is then possible for every $\alpha \in (0, 1)$ as a consequence of Proposition 5.12 in the same reference upon noting that then $C^* = -\infty$ and $\alpha^* = 1$ hold. [We note in passing that for such a heteroskedasticity model $\mathfrak{C}$ the size of the rejection region $\{T_{uc} \geq C\}$ ($\{T_{Het} \geq C\}$, respectively) is less than 1 for every $C > 0$ (this follows from Proposition 5.2 and Remark 5.4 in Pötscher and Preinerstorfer (2018) as well as Part 6 of Lemma 5.15

[66] A related but slightly different argument proceeds by directly noting from its definition that $\mathbb{J}(\mathfrak{M}_0^{lin}, \mathfrak{C}_{(n_1,\ldots n_m)})$ is empty in case $m = 1$ (cf. Apendix B), and then to appeal to Remark 5.7 (or Proposition 5.12) in Pötscher and Preinerstorfer (2018).

in Preinerstorfer and Pötscher (2016)).][67]

(ii) A particular instance of the situation described in (i) is provided by heteroskedasticity models $\mathfrak{C}$ that are subsets of a set of the form $\mathfrak{C}_{Het,\tau_*}$ $(0 < \tau_* < n^{-1/2})$, as in this case $\mathbb{J}(\mathfrak{M}_0^{lin}, \mathfrak{C})$ is easily seen to be empty.

(iii) More generally, the tests based on T_{uc} (on T_{Het}, respectively) are size controllable over a heteroskedasticity model $\mathfrak{C}$, provided any $\mathcal{S} \in \mathbb{J}(\mathfrak{M}_0^{lin}, \mathfrak{C})$ is not contained in $\operatorname{span}(X)$ (B, respectively). This follows easily from Corollary 5.6 and Proposition 5.12 in Pötscher and Preinerstorfer (2018), the latter proposition also providing an exact size result, which we refrain from spelling out in detail. Again there is a (partial) converse: If an $\mathcal{S} \in \mathbb{J}(\mathfrak{M}_0^{lin}, \mathfrak{C})$ exists with $\mathcal{S} \subseteq \operatorname{span}(X)$, then the size over $\mathfrak{C}$ of the rejection region $\{T_{uc} \geq C\}$ ($\{T_{Het} \geq C\}$, respectively) is equal to 1; see Theorem 3.1 in Pötscher and Preinerstorfer (2019). Furthermore, lower bounds for critical values that lead to size less than 1 (in particular, for size-controlling critical values) can be had with the help of Corollary 5.17 in Preinerstorfer and Pötscher (2016), Lemma 5.11 and Proposition 5.12 in Pötscher and Preinerstorfer (2018), or Lemma 4.1 in Pötscher and Preinerstorfer (2019).

A.2 Size control results for $\tilde{T}_{Het}$ and $\tilde{T}_{uc}$

The proof of the subsequent theorem is given in Appendix D. We note that the first statement in Part (a) of the subsequent theorem is actually trivial, since $\tilde{T}_{uc}$ is bounded as has been shown in Section 6.2.1.

Theorem A.5. *Let $m \in \mathbb{N}$, and let $n_j \in \mathbb{N}$ for $j = 1, \ldots, m$ satisfy $\sum_{j=1}^{m} n_j = n$. Then:*

(a) For every $0 < \alpha < 1$ there exists a real number $C(\alpha)$ such that

$$\sup_{\mu_0 \in \mathfrak{M}_0} \sup_{0<\sigma^2<\infty} \sup_{\Sigma \in \mathfrak{C}_{(n_1,\ldots,n_m)}} P_{\mu_0,\sigma^2\Sigma}(\tilde{T}_{uc} \geq C(\alpha)) \leq \alpha \tag{A.5}$$

holds. Furthermore, even equality can be achieved in (A.5) by a proper choice of $C(\alpha)$, provided $\alpha \in (0, \alpha^*] \cap (0, 1)$ holds, where $\alpha^* = \sup_{C \in (C^*,\infty)} \sup_{\Sigma \in \mathfrak{C}_{(n_1,\ldots,n_m)}} P_{\mu_0,\Sigma}(\tilde{T}_{uc} \geq C)$ and where C^* is defined as in Lemma 5.11 of Pötscher and Preinerstorfer (2018) with $\mathfrak{C} = \mathfrak{C}_{(n_1,\ldots,n_m)}$, $T = \tilde{T}_{uc}$, $N^\dagger = \mathfrak{M}_0$, and $\mathcal{L} = \mathfrak{M}_0^{lin}$ (with neither α^* nor C^* depending on the choice of $\mu_0 \in \mathfrak{M}_0$).

(b) Suppose Assumption 2 is satisfied. Suppose further that $\tilde{T}_{Het}$ is not constant on $\mathbb{R}^n \backslash \tilde{\mathsf{B}}$.[68] *Then for every $0 < \alpha < 1$ there exists a real number $C(\alpha)$ such that*

$$\sup_{\mu_0 \in \mathfrak{M}_0} \sup_{0<\sigma^2<\infty} \sup_{\Sigma \in \mathfrak{C}_{(n_1,\ldots,n_m)}} P_{\mu_0,\sigma^2\Sigma}(\tilde{T}_{Het} \geq C(\alpha)) \leq \alpha \tag{A.6}$$

[67] The verification of the assumptions in Corollary 5.6 and Propositions 5.2 and 5.12 of Pötscher and Preinerstorfer (2018) proceeds as in the proofs of Theorems 5.1 and A.1.

[68] Cf. Footnote 39.

holds, provided that for some $\mu_0 \in \mathfrak{M}_0$ *(and hence for all* $\mu_0 \in \mathfrak{M}_0$*)*

$$\mu_0 + \operatorname{span}\left(\left\{e_i(n) : i \in (n_{j-1}^+, n_j^+]\right\}\right) \nsubseteq \tilde{\mathsf{B}} \quad \text{for every } j = 1, \ldots, m \text{ with } (n_{j-1}^+, n_j^+] \cap I_1(\mathfrak{M}_0^{lin}) \neq \emptyset. \tag{A.7}$$

Furthermore, under condition (A.7), even equality can be achieved in (A.6) by a proper choice of $C(\alpha)$*, provided* $\alpha \in (0, \alpha^*] \cap (0,1)$ *holds, where* $\alpha^* = \sup_{C \in (C^*, \infty)} \sup_{\Sigma \in \mathfrak{C}_{(n_1,\ldots,n_m)}} P_{\mu_0,\Sigma}(\tilde{T}_{Het} \geq C)$ *and where* C^* *is defined as in Lemma 5.11 of Pötscher and Preinerstorfer (2018) with* $\mathfrak{C} = \mathfrak{C}_{(n_1,\ldots,n_m)}$*,* $T = \tilde{T}_{Het}$*,* $N^\dagger = \tilde{\mathsf{B}}$*, and* $\mathcal{L} = \mathfrak{M}_0^{lin}$ *(with neither* α^* *nor* C^* *depending on the choice of* $\mu_0 \in \mathfrak{M}_0$*).*

(c) Under the assumptions of Part (a) (Part (b), respectively) implying existence of a critical value $C(\alpha)$ *satisfying (A.5) ((A.6), respectively), a smallest critical value, denoted by* $C_\Diamond(\alpha)$*, satisfying (A.5) ((A.6), respectively) exists for every* $0 < \alpha < 1$.[69] *And* $C_\Diamond(\alpha)$ *corresponding to Part (a) (Part (b), respectively) is also the smallest among the critical values leading to equality in (A.5) ((A.6), respectively) whenever such critical values exist. [Although* $C_\Diamond(\alpha)$ *corresponding to Part (a) and (b), respectively, will typically be different, we use the same symbol.]*[70]

It is easy to see that the discussion in the first paragraph following Theorem 6.4 applies mutatis mutandis also to the above theorem. Similarly, Remarks 6.5, 6.6, 6.8, 6.11, and Proposition 6.7 carry over.

A discussion of size control results for $\tilde{T}_{uc}$ and $\tilde{T}_{Het}$ over other choices of $\mathfrak{C}$ based on the results in Section 5 of Pötscher and Preinerstorfer (2018) can also be given (cf. the discussion in Subsection A.1.2), but we refrain from spelling out the details. We only note that the test based on $\tilde{T}_{Het}$ is always size controllable over $\mathfrak{C}_{Het,\tau_*}$ $(0 < \tau_* < n^{-1/2})$, and the same is trivially true for $\tilde{T}_{uc}$.

A.3 A useful observation

Let $\mathfrak{C}$ be an arbitrary heteroskedasticity model (i.e., $\emptyset \neq \mathfrak{C} \subseteq \mathfrak{C}_{Het}$), let $0 < \alpha < 1$, and let T stand for T_{uc} or T_{Het}, respectively, where in case of $T = T_{Het}$ we assume that Assumption 1 is satisfied. Suppose that T is size-controllable at significance level α (i.e., that $\sup_{\mu_0 \in \mathfrak{M}_0} \sup_{0<\sigma^2<\infty} \sup_{\Sigma \in \mathfrak{C}} P_{\mu_0,\sigma^2\Sigma}(T \geq C) \leq \alpha$ holds for some real C). Then a smallest size-controlling critical value $C_\Diamond(\alpha)$ always exists.[71] And if a critical value $C \in \mathbb{R}$ exists such that $\sup_{\mu_0 \in \mathfrak{M}_0} \sup_{0<\sigma^2<\infty} \sup_{\Sigma \in \mathfrak{C}} P_{\mu_0,\sigma^2\Sigma}(T \geq C) = \alpha$ holds, then $C_\Diamond(\alpha)$ is also the smallest among these critical values. This follows from Remark 5.10 and Lemma 5.16 in Pötscher and Preinerstorfer (2018) combined with Remark C.1 in Appendix C. The same is true for $T = \tilde{T}_{uc}$ and $T = \tilde{T}_{Het}$, where in case of $T = \tilde{T}_{Het}$ we assume that Assumption 2 is satisfied and that $\tilde{T}_{Het}$

[69] Note that there are in fact no assumptions for Part (a). We have chosen this formulation for reasons of brevity.

[70] Cf. also Appendix A.3.

[71] Note that this, e.g., covers the case discussed in Example 5.5, where size-control can be established for T_{Het} despite the fact that the sufficient conditions in Theorem 5.1 are not satisfied (and hence Part (c) of that theorem can not be used).

is not constant on $\mathbb{R}^n \backslash \tilde{\mathsf{B}}$. This follows again from Remark 5.10 in Pötscher and Preinerstorfer (2018) now together with Lemma D.1 in Appendix D.

B Appendix: Characterization of $\mathbb{J}(\mathcal{L}, \mathfrak{C})$ for $\mathfrak{C} = \mathfrak{C}_{Het}$ and $\mathfrak{C} = \mathfrak{C}_{(n_1,\ldots,n_m)}$

A key ingredient in the proof of size control results such as Theorem 5.1 or 6.4 is a certain collection $\mathbb{J}(\mathcal{L}, \mathfrak{C})$ of linear subspaces of $\mathbb{R}^n$ introduced in Pötscher and Preinerstorfer (2018). For the convenience of the reader we reproduce this definition, specialized to the present setting, below. The leading case in the applications will be the case $\mathcal{L} = \mathfrak{M}_0^{lin}$.

Definition B.1. Let $\mathfrak{C}$ be a heteroskedasticity model (i.e., $\emptyset \neq \mathfrak{C} \subseteq \mathfrak{C}_{Het}$). Given a linear subspace $\mathcal{L}$ of $\mathbb{R}^n$ with $\dim(\mathcal{L}) < n$ and an element $\Sigma \in \mathfrak{C}$, we let

$$\mathcal{L}(\Sigma) = \frac{\Pi_{\mathcal{L}^\perp} \Sigma \Pi_{\mathcal{L}^\perp}}{\|\Pi_{\mathcal{L}^\perp} \Sigma \Pi_{\mathcal{L}^\perp}\|}$$

and $\mathcal{L}(\mathfrak{C}) = \{\mathcal{L}(\Sigma) : \Sigma \in \mathfrak{C}\}$. Furthermore, we define

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

where the closure $\operatorname{cl}(\cdot)$ is to be understood w.r.t. $\mathbb{R}^{n\times n}$.

Recalling the definition of $I_0(\mathcal{L})$, it is easy to see that $I_0(\mathcal{L}) = \left\{i : 1 \leq i \leq n, \pi_{\mathcal{L}^\perp,i} = 0\right\}$ holds, where $\pi_{\mathcal{L}^\perp,i}$ denotes the i-th column of $\Pi_{\mathcal{L}^\perp}$. Also recall that $I_1(\mathcal{L})$ is nonempty in case $\dim(\mathcal{L}) < n$ holds. The characterization of $\mathbb{J}(\mathcal{L}, \mathfrak{C}_{Het})$ is now given in the next proposition.

Proposition B.1. *Suppose* $\dim(\mathcal{L}) < n$ *holds. Then the set* $\mathbb{J}(\mathcal{L}, \mathfrak{C}_{Het})$ *is given by*

$$\left\{\operatorname{span}\left(\left\{\pi_{\mathcal{L}^\perp,i} : i \in I\right\}\right) : \emptyset \neq I \subseteq I_1(\mathcal{L}),\ \dim\left(\operatorname{span}\left(\left\{\pi_{\mathcal{L}^\perp,i} : i \in I\right\}\right)\right) < n - \dim(\mathcal{L})\right\}. \quad \text{(B.1)}$$

This proposition is a special case of Proposition B.2 given below since $\mathfrak{C}_{Het}$ coincides with $\mathfrak{C}_{(n_1,\ldots,n_m)}$ in case $m = n$ and $n_j = 1$ for all $j = 1, \ldots, m$.

We next turn to the characterization of $\mathbb{J}(\mathcal{L}, \mathfrak{C}_{(n_1,\ldots,n_m)})$, where $\mathfrak{C}_{(n_1,\ldots,n_m)}$ has been defined in Appendix A. Here $m \in \mathbb{N}$, and $n_j \in \mathbb{N}$ for $j = 1, \ldots, m$ satisfy $\sum_{j=1}^m n_j = n$. Consider the partition of the set $\{1, \ldots, n\}$ into the intervals $(n_0^+, n_1^+]$, $(n_1^+, n_2^+]$,..., $(n_{m-1}^+, n_m^+]$ where n_j^+ has been defined in Appendix A. Let $\boldsymbol{I}_{(n_1,\ldots,n_m)}$ consist of all non-empty subsets I of $\{1, \ldots, n\}$ that can be represented as a union of intervals of the form $(n_{j-1}^+, n_j^+]$.

Proposition B.2. *Suppose* $\dim(\mathcal{L}) < n$ *holds. Let* $m \in \mathbb{N}$, *and let* $n_j \in \mathbb{N}$ *for* $j = 1, \ldots, m$

satisfy $\sum_{j=1}^{m} n_j = n$*. Then the set* $\mathbb{J}(\mathcal{L}, \mathfrak{C}_{(n_1,\ldots,n_m)})$ *is given by*

$$\left\{\operatorname{span}\left(\left\{\pi_{\mathcal{L}^\perp,i} : i \in I\right\}\right) : I \in \boldsymbol{I}_{(n_1,\ldots,n_m)},\ \emptyset \neq I \cap I_1(\mathcal{L}),\ \dim\left(\operatorname{span}\left(\left\{\pi_{\mathcal{L}^\perp,i} : i \in I\right\}\right)\right) < n - \dim(\mathcal{L})\right\}. \tag{B.2}$$

Note that in (B.2) we have $\operatorname{span}\left(\left\{\pi_{\mathcal{L}^\perp,i} : i \in I\right\}\right) = \operatorname{span}\left(\left\{\pi_{\mathcal{L}^\perp,i} : i \in I \cap I_1(\mathcal{L})\right\}\right)$.

Proof: Suppose $\mathcal{S}$ is an element of $\mathbb{J}(\mathcal{L}, \mathfrak{C}_{(n_1,\ldots,n_m)})$. Then there exist a sequence $\Sigma_m \in \mathfrak{C}_{(n_1,\ldots,n_m)}$ such that $\Pi_{\mathcal{L}^\perp}\Sigma_m\Pi_{\mathcal{L}^\perp} / \left\|\Pi_{\mathcal{L}^\perp}\Sigma_m\Pi_{\mathcal{L}^\perp}\right\|$ converges to a limit $\bar{\Sigma}$, say, in $\mathbb{R}^{n\times n}$ with $\operatorname{span}(\bar{\Sigma}) = \mathcal{S}$. Now,

$$\begin{aligned}
\Pi_{\mathcal{L}^\perp}\Sigma_m\Pi_{\mathcal{L}^\perp} / \left\|\Pi_{\mathcal{L}^\perp}\Sigma_m\Pi_{\mathcal{L}^\perp}\right\| &= \left\|\Pi_{\mathcal{L}^\perp}\Sigma_m\Pi_{\mathcal{L}^\perp}\right\|^{-1} \sum_{i=1}^{n} \tau_i^2(m)\pi_{\mathcal{L}^\perp,i}\pi_{\mathcal{L}^\perp,i}' \\
&= \left\|\Pi_{\mathcal{L}^\perp}\Sigma_m\Pi_{\mathcal{L}^\perp}\right\|^{-1} \sum_{j=1}^{m} \sum_{i\in(n_{j-1}^+,n_j^+]} \tau_i^2(m)\pi_{\mathcal{L}^\perp,i}\pi_{\mathcal{L}^\perp,i}' \\
&= \sum_{j:(n_{j-1}^+,n_j^+]\cap I_1(\mathcal{L})\neq\emptyset} \left\|\Pi_{\mathcal{L}^\perp}\Sigma_m\Pi_{\mathcal{L}^\perp}\right\|^{-1} \tau_{n_j^+}^2(m) \sum_{i\in(n_{j-1}^+,n_j^+]} \pi_{\mathcal{L}^\perp,i}\pi_{\mathcal{L}^\perp,i}',
\end{aligned}$$

where $\tau_i^2(m)$ denotes the i-th diagonal element of Σ_m. Here we have used the fact that variances are constant within groups, as well as that $\pi_{\mathcal{L}^\perp,i} = 0$ for all $i \in (n_{j-1}^+, n_j^+]$ if $(n_{j-1}^+, n_j^+]$ is disjoint from $I_1(\mathcal{L})$. Also note that the outer sum extends over a nonempty index set since $\operatorname{card}(I_1(\mathcal{L})) \geq 1$ must hold in view of $\dim(\mathcal{L}) < n$. Since the l.h.s. converges to the limit $\bar{\Sigma} \in \mathbb{R}^{n\times n}$, since the r.h.s. is bounded from below in the Loewner order by

$$\left\|\Pi_{\mathcal{L}^\perp}\Sigma_m\Pi_{\mathcal{L}^\perp}\right\|^{-1} \tau_{n_j^+}^2(m) \sum_{i\in(n_{j-1}^+,n_j^+]} \pi_{\mathcal{L}^\perp,i}\pi_{\mathcal{L}^\perp,i}',$$

for every j appearing in the range of the outer sum, and since $\pi_{\mathcal{L}^\perp,i} \neq 0$ for at least one $i \in (n_{j-1}^+, n_j^+]$ holds when $(n_{j-1}^+, n_j^+] \cap I_1(\mathcal{L}) \neq \emptyset$, it follows that the sequence

$$\left(\left\|\Pi_{\mathcal{L}^\perp}\Sigma_m\Pi_{\mathcal{L}^\perp}\right\|^{-1} \tau_{n_j^+}^2(m) : m \in \mathbb{N}\right)$$

is bounded for every j satisfying $(n_{j-1}^+, n_j^+] \cap I_1(\mathcal{L}) \neq \emptyset$. Possibly after passing to a subsequence, we may thus assume that these sequences converge to nonnegative real numbers γ_j for such j. It follows that

$$\begin{aligned}
\bar{\Sigma} &= \sum_{j:(n_{j-1}^+,n_j^+]\cap I_1(\mathcal{L})\neq\emptyset} \gamma_j \sum_{i\in(n_{j-1}^+,n_j^+]} \pi_{\mathcal{L}^\perp,i}\pi_{\mathcal{L}^\perp,i}' \\
&= \sum_{j:(n_{j-1}^+,n_j^+]\cap I_1(\mathcal{L})\neq\emptyset} \sum_{i\in(n_{j-1}^+,n_j^+]} \gamma_j^{1/2}\pi_{\mathcal{L}^\perp,i}\left(\gamma_j^{1/2}\pi_{\mathcal{L}^\perp,i}\right)'.
\end{aligned}$$

Let I be the union of those intervals $(n_{j-1}^+, n_j^+]$ satisfying (i) $(n_{j-1}^+, n_j^+] \cap I_1(\mathcal{L}) \neq \emptyset$ and (ii)

$\gamma_j > 0$. Note that I cannot be the empty set as this would imply $\bar{\Sigma} = 0$, which is impossible since it is the limit of a sequence of matrices residing in the unit sphere of $\mathbb{R}^{n\times n}$. Furthermore, by construction, $I \in \boldsymbol{I}_{(n_1,\ldots,n_m)}$ and $I \cap I_1(\mathcal{L}) \neq \emptyset$ hold. Using the fact that $\operatorname{span}(\sum_{l=1}^{L} A_l A_l') = \operatorname{span}(A_1, \ldots, A_L)$ holds for arbitrary real matrices of the same row-dimension, we obtain $\mathcal{S} = \operatorname{span}(\bar{\Sigma}) = \operatorname{span}\left(\left\{\pi_{\mathcal{L}^\perp,i} : i \in I\right\}\right)$ for the before constructed set I. [Note that $\pi_{\mathcal{L}^\perp,i} = 0$ if $i \in (n_{j-1}^+, n_j^+]$ but $i \notin I_1(\mathcal{L})$.] Since $\mathcal{S}$, being an element of $\mathbb{J}(\mathcal{L}, \mathfrak{C}_{(n_1,\ldots,n_m)})$, satisfies $\dim(\mathcal{S}) < n - \dim(\mathcal{L})$, we have established that $\mathcal{S}$ is also an element of (B.2).

To prove the converse, suppose that $\mathcal{S}$ is an element of (B.2), i.e., that $\mathcal{S} = \operatorname{span}\left(\left\{\pi_{\mathcal{L}^\perp,i} : i \in I\right\}\right)$ for some $I \in \boldsymbol{I}_{(n_1,\ldots,n_m)}$ with $\emptyset \neq I \cap I_1(\mathcal{L})$ and that $\dim(\mathcal{S}) < n - \dim(\mathcal{L})$ holds. Note that $\operatorname{card}(I) < n$ holds, since otherwise $\mathcal{S} = \mathcal{L}^\perp$ would follow, contradicting $\dim(\mathcal{S}) < n - \dim(\mathcal{L})$. Also note that $\operatorname{card}(I) \geq 1$ as $\emptyset \neq I \cap I_1(\mathcal{L})$. Define diagonal $n \times n$ matrices Σ_m via their diagonal elements

$$\tau_i^2(m) = \begin{cases} (\operatorname{card}(I))^{-1} - \delta_m & \text{if } i \in I \\ (\operatorname{card}(I)/(n - \operatorname{card}(I)))\, \delta_m & \text{if } i \notin I \end{cases}$$

where $0 < \delta_m < 1/\operatorname{card}(I)$ with $\delta_m \to 0$ for $m \to \infty$. Then $\tau_i^2(m) > 0$ as well as $\sum_{i=1}^{n} \tau_i^2(m) = 1$ hold, and $\tau_{n_{j-1}^+ + 1}^2(m) = \ldots = \tau_{n_j^+}^2(m)$ holds for $j = 1, \ldots, m$ since $I \in \boldsymbol{I}_{(n_1,\ldots,n_m)}$. That is, Σ_m belongs to $\mathfrak{C}_{(n_1,\ldots,n_m)}$. Obviously, Σ_m converges to a diagonal matrix Σ^* with diagonal elements given by

$$\tau_i^{*2} = \begin{cases} (\operatorname{card}(I))^{-1} & \text{if } i \in I \\ 0 & \text{if } i \notin I \end{cases}.$$

Consequently, $\Pi_{\mathcal{L}^\perp} \Sigma_m \Pi_{\mathcal{L}^\perp} / \left\|\Pi_{\mathcal{L}^\perp} \Sigma_m \Pi_{\mathcal{L}^\perp}\right\|$ converges to $\bar{\Sigma} := \Pi_{\mathcal{L}^\perp} \Sigma^* \Pi_{\mathcal{L}^\perp} / \left\|\Pi_{\mathcal{L}^\perp} \Sigma^* \Pi_{\mathcal{L}^\perp}\right\|$, since $\Pi_{\mathcal{L}^\perp} \Sigma^* \Pi_{\mathcal{L}^\perp} \neq 0$ in view of

$$\Pi_{\mathcal{L}^\perp} \Sigma^* \Pi_{\mathcal{L}^\perp} = \sum_{i=1}^{n} \tau_i^{*2} \pi_{\mathcal{L}^\perp,i} \pi_{\mathcal{L}^\perp,i}' = (\operatorname{card}(I))^{-1} \sum_{i \in I} \pi_{\mathcal{L}^\perp,i} \pi_{\mathcal{L}^\perp,i}'$$

and the fact that $\emptyset \neq I \cap I_1(\mathcal{L})$ holds and thus $\pi_{\mathcal{L}^\perp,i} \neq 0$ must hold at least for one $i \in I$. Again using $\operatorname{span}(\sum_{l=1}^{L} A_l A_l') = \operatorname{span}(A_1, \ldots, A_L)$ we arrive at

$$\operatorname{span}(\bar{\Sigma}) = \operatorname{span}(\Pi_{\mathcal{L}^\perp} \Sigma^* \Pi_{\mathcal{L}^\perp}) = \operatorname{span}\left((\operatorname{card}(I))^{-1} \sum_{i \in I} \pi_{\mathcal{L}^\perp,i} \pi_{\mathcal{L}^\perp,i}'\right) = \operatorname{span}\left(\left\{\pi_{\mathcal{L}^\perp,i} : i \in I\right\}\right) = \mathcal{S}.$$

Because we have assumed that $\dim(\mathcal{S}) < n - \dim(\mathcal{L})$ holds, the preceding display shows that $\mathcal{S} \in \mathbb{J}(\mathcal{L}, \mathfrak{C}_{(n_1,\ldots,n_m)})$. ■

Remark B.3. Note that $\mathbb{J}(\mathcal{L}, \mathfrak{C}_{(n_1,\ldots,n_m)})$ is empty if $m = 1$ (as can be seen directly from the definition of $\mathbb{J}(\mathcal{L}, \mathfrak{C}_{(n_1,\ldots,n_m)})$ or from (B.2)).

Remark B.4. It is easy to see that the concentration spaces of $\mathfrak{C}_{Het}$ in the sense of Preinerstorfer and Pötscher (2016) are precisely given by all spaces of the form $\operatorname{span}\left(\{e_i(n) : i \in I\}\right)$ where I varies through all subsets of $\{1, \ldots, n\}$ that satisfy $0 < \operatorname{card}(I) < n$. More generally, the concentration spaces of $\mathfrak{C}_{(n_1,\ldots,n_m)}$ are precisely given by all spaces of the form span $(\{e_i(n) : i \in I\})$ where $I \in \boldsymbol{I}_{(n_1,\ldots,n_m)}$ satisfies $0 < \text{card}(I) < n$. [In view of Remark 5.1(i) in Pötscher and Preinerstorfer (2018) these results correspond to the case $\dim(\mathcal{L}) = 0$ in the preceding propositions.]

C Appendix: Proofs for Section 5 and Appendix A.1

The facts collected in the subsequent remark will be used in the proofs further below.

Remark C.1. (i) Suppose Assumption 1 holds. Then the test statistic T_{Het} is a non-sphericity corrected F-type test statistic in the sense of Section 5.4 in Preinerstorfer and Pötscher (2016). More precisely, T_{Het} is of the form (28) in Preinerstorfer and Pötscher (2016) and Assumption 5 in the same reference is satisfied with $\check{\beta} = \hat{\beta}$, $\check{\Omega} = \hat{\Omega}_{Het}$, and $N = \emptyset$. Furthermore, the set N^* defined in (27) of Preinerstorfer and Pötscher (2016) satisfies $N^* = \mathsf{B}$. And also Assumptions 6 and 7 of Preinerstorfer and Pötscher (2016) are satisfied. All these claims follow easily in view of Lemma 4.1 in Preinerstorfer and Pötscher (2016), see also the proof of Theorem 4.2 in that reference.

(ii) The test statistic T_{uc} is also a non-sphericity corrected F-type test statistic in the sense of Section 5.4 in Preinerstorfer and Pötscher (2016) (terminology being somewhat unfortunate here as no correction for the non-sphericity is being attempted). More precisely, T_{uc} is of the form (28) in Preinerstorfer and Pötscher (2016) and Assumption 5 in the same reference is satisfied with $\check{\beta} = \hat{\beta}$, $\check{\Omega} = \hat{\sigma}^2 R\left(X'X\right)^{-1} R'$, and $N = \emptyset$. Furthermore, the set N^* defined in (27) of Preinerstorfer and Pötscher (2016) satisfies $N^* = \text{span}(X)$. And also Assumptions 6 and 7 of Preinerstorfer and Pötscher (2016) are satisfied. All these claims are evident (and obviously do not rely on Assumption 1).

(iii) We note that any non-sphericity corrected F-type test statistic (for testing (3)) in the sense of Section 5.4 in Preinerstorfer and Pötscher (2016), i.e., any test statistic T of the form (28) in Preinerstorfer and Pötscher (2016) that also satisfies Assumption 5 in that reference, is invariant under the group $G(\mathfrak{M}_0)$. Furthermore, the associated set N^* defined in (27) of Preinerstorfer and Pötscher (2016) is even invariant under the larger group $G(\mathfrak{M})$. See Sections 5.1 and 5.4 of Preinerstorfer and Pötscher (2016) as well as Lemma 5.16 in Pötscher and Preinerstorfer (2018) for more information.

Proof of Theorem 5.1: We first prove Part (b). We apply Part (b) of Theorem A.1 with $n_j = 1$ for $j = 1, \ldots, n = m$ observing that then $\mathfrak{C}_{(n_1,\ldots,n_m)} = \mathfrak{C}_{Het}$ and that condition (A.4) reduces to (10) (exploiting that B is a finite union of proper linear subspaces as discussed in Lemma 3.1). This establishes (9). The final claim in Part (b) of the theorem follows from Part (b) of Theorem A.1, if we can show that α^* and C^* given there can be written as claimed in Theorem 5.1: To this end we proceed as follows:[72] Choose an element μ_0 of $\mathfrak{M}_0$. Observe that $I_1(\mathfrak{M}_0^{lin}) \neq \emptyset$ (since $\dim(\mathfrak{M}_0^{lin}) = k - q < n$), and that for every $i \in I_1(\mathfrak{M}_0^{lin})$ the linear space

[72]Alternatively, one could base a proof on Lemma C.1 in Pötscher and Preinerstorfer (2019).

$\mathcal{S}_i = \text{span}(\Pi_{(\mathfrak{M}_0^{lin})^\perp} e_i(n))$ is 1-dimensional (since $\mathcal{S}_i = \{0\}$ is impossible in view of $i \in I_1(\mathfrak{M}_0^{lin})$), and belongs to $\mathbb{J}(\mathfrak{M}_0^{lin}, \mathfrak{C}_{Het})$ (since $n - k + q > 1 = \dim(\mathcal{S}_i)$ holds) in view of Proposition B.1 in Section B. Since T_{Het} is $G(\mathfrak{M}_0)$-invariant (Remark C.1 above), it follows that T_{Het} is constant on $(\mu_0 + \mathcal{S}_i) \backslash \{\mu_0\}$, cf. the beginning of the proof of Lemma 5.11 in Pötscher and Preinerstorfer (2018). Hence, $\mathcal{S}_i$ belongs to $\mathbb{H}$ (defined in Lemma 5.11 in Pötscher and Preinerstorfer (2018)) and consequently for C^* as defined in that lemma

$$C^* \geq \max \left\{ T_{Het}(\mu_0 + \Pi_{(\mathfrak{M}_0^{lin})^\perp} e_i(n)) : i \in I_1(\mathfrak{M}_0^{lin}) \right\}$$

must hold (recall that $\Pi_{(\mathfrak{M}_0^{lin})^\perp} e_i(n) \neq 0$). To prove the opposite inequality, let $\mathcal{S}$ be an arbitrary element of $\mathbb{H}$, i.e., $\mathcal{S} \in \mathbb{J}(\mathfrak{M}_0^{lin}, \mathfrak{C}_{Het})$ and T_{Het} is $\lambda_{\mu_0 + \mathcal{S}}$-almost everywhere equal to a constant $C(\mathcal{S})$, say. Then Proposition B.1 in Section B shows that $\mathcal{S}_i \subseteq \mathcal{S}$ holds for some $i \in I_1(\mathfrak{M}_0^{lin})$. Because of Condition (10) we have $\mathcal{S}_i \nsubseteq \mathsf{B}$ since $\Pi_{(\mathfrak{M}_0^{lin})^\perp} e_i(n)$ and $e_i(n)$ differ only by an element of $\mathfrak{M}_0^{lin} \subseteq \text{span}(X)$ and since $\mathsf{B} + \text{span}(X) = \mathsf{B}$. Thus $\mu_0 + \mathcal{S}_i \nsubseteq \mathsf{B}$ by the same argument as $\mu_0 \in \mathfrak{M}_0 \subseteq \text{span}(X)$. We thus can find $s \in \mathcal{S}_i$ such that $\mu_0 + s \notin \mathsf{B}$. Note that $s \neq 0$ must hold, since $\mu_0 \in \mathfrak{M}_0 \subseteq \text{span}(X) \subseteq \mathsf{B}$. In particular, T_{Het} is continuous at $\mu_0 + s$, since $\mu_0 + s \notin \mathsf{B}$. Now, for every open ball A_ε in $\mathbb{R}^n$ with center s and radius $\varepsilon > 0$ we can find an element $a_\varepsilon \in A_\varepsilon \cap \mathcal{S}$ such that $T_{Het}(\mu_0 + a_\varepsilon) = C(\mathcal{S})$. Since $a_\varepsilon \to s$ for $\varepsilon \to 0$, it follows that $C(\mathcal{S}) = T_{Het}(\mu_0 + s)$. Since $s \neq 0$ and since T_{Het} is constant on $(\mu_0 + \mathcal{S}_i) \backslash \{\mu_0\}$ as shown before, we can conclude that $C(\mathcal{S}) = T_{Het}(\mu_0 + s) = T_{Het}(\mu_0 + \Pi_{(\mathfrak{M}_0^{lin})^\perp} e_i(n))$, where we recall that $\Pi_{(\mathfrak{M}_0^{lin})^\perp} e_i(n) \neq 0$. But this now implies

$$C^* = \max \left\{ T_{Het}(\mu_0 + \Pi_{(\mathfrak{M}_0^{lin})^\perp} e_i(n)) : i \in I_1(\mathfrak{M}_0^{lin}) \right\}.$$

Using $G(\mathfrak{M}_0)$-invariance of T_{Het} we conclude that

$$C^* = \max \left\{ T_{Het}(\mu_0 + e_i(n)) : i \in I_1(\mathfrak{M}_0^{lin}) \right\}.$$

The expression for α^* given in the theorem now follows immediately from the expression for α^* given in Part (b) of Theorem A.1.

We next prove Part (a): Apply Part (a) of Theorem A.1 with $n_j = 1$ for $j = 1, \ldots, n = m$ observing that then $\mathfrak{C}_{(n_1, \ldots, n_m)} = \mathfrak{C}_{Het}$ and that condition (A.2) reduces to (8) (exploiting that $\text{span}(X)$ is a linear space). This establishes (7). The final claim in Part (a) of the theorem follows similarly as the corresponding claim of Part (b) upon replacing the set B by $\text{span}(X)$ in the argument and by noting that T_{uc} is $G(\mathfrak{M}_0)$-invariant.

Part (c) follows from Part (c) of Theorem A.1 upon setting $n_j = 1$ for $j = 1, \ldots, n = m$ (and upon noting that then the conditions in Theorem A.1 reduce to the conditions of the present theorem). ■

Proof of Proposition 5.5: Follows from Part A.1 of Proposition 5.12 of Pötscher and

Preinerstorfer (2018) and the sentence following this proposition. Note that the assumptions of this proposition have been verified in the proof of Theorem 5.1 (see also the proof of Theorem A.1, on which the proof of Theorem 5.1 is based), where it is also shown that the quantity C^* used in Proposition 5.12 of Pötscher and Preinerstorfer (2018) coincides with C^* defined in Theorem 5.1. ■

We note that the result for T_{Het} in Proposition 5.5 can also be obtained from Theorem 4.2 in Preinerstorfer and Pötscher (2016).

Proof of Proposition 5.7: (a) This can be seen as follows (cf. also the discussion on p.302 of Preinerstorfer and Pötscher (2016)): By Remark C.1 above, T_{uc} satisfies the assumptions in Corollary 5.17 in Preinerstorfer and Pötscher (2016) (with $\check{\beta} = \hat{\beta}$, $\check{\Omega}(y) = \hat{\sigma}^2(y)R\left(X'X\right)^{-1}R'$, $N = \emptyset$, and $N^* = \operatorname{span}(X)$). Let $e_i(n)$ be one of the standard basis vectors with $i \in I_1(\mathfrak{M}_0^{lin})$ that does belong to $\operatorname{span}(X)$. Set $\mathcal{Z} = \operatorname{span}(e_i(n))$ and note that this is a concentration space of $\mathfrak{C}_{Het}$, cf. Remark B.4 in Appendix B. The nonnegative definiteness assumption on $\check{\Omega}$ in Part 3 of Corollary 5.17 in Preinerstorfer and Pötscher (2016) is clearly satisfied. We also have $\check{\Omega}(\lambda e_i(n)) = 0$ (since $e_i(n) \in \operatorname{span}(X)$) for every $\lambda \in \mathbb{R}$ and $R\hat{\beta}(\lambda e_i(n)) \neq 0$ for every $\lambda \in \mathbb{R}\backslash\{0\}$ (since $e_i(n) \in \operatorname{span}(X)$ but $e_i(n) \notin \mathfrak{M}_0^{lin}$ in view of $i \in I_1(\mathfrak{M}_0^{lin})$). Part 3 of Corollary 5.17 in Preinerstorfer and Pötscher (2016) then proves the claim for $C > 0$. A fortiori it then also holds for all real C.

(b) This follows for $C > 0$ from Part 3 of Theorem 4.2 in Preinerstorfer and Pötscher (2016) upon observing that a vector $e_i(n)$ satisfying $e_i(n) \in \operatorname{span}(X)$ for some $i \in I_1(\mathfrak{M}_0^{lin})$ clearly satisfies $B(e_i(n)) = 0$ (as $e_i(n) \in \operatorname{span}(X)$) and $R\hat{\beta}(e_i(n)) \neq 0$ (since $e_i(n) \in \operatorname{span}(X)$ but $e_i(n) \notin \mathfrak{M}_0^{lin}$ in view of $i \in I_1(\mathfrak{M}_0^{lin})$). A fortiori it then also holds for all real C. ■

Proof of Theorem A.1: We first prove Part (b). We wish to apply Part A of Proposition 5.12 of Pötscher and Preinerstorfer (2018) with $\mathfrak{C} = \mathfrak{C}_{(n_1,\ldots,n_m)}$, $T = T_{Het}$, $\mathcal{L} = \mathfrak{M}_0^{lin}$, and $\mathcal{V} = \{0\}$. First, note that $\dim(\mathfrak{M}_0^{lin}) = k - q < n$. Second, under Assumption 1, T_{Het} is a non-sphericity corrected F-type test with $N^* = \mathsf{B}$, which is a closed $\lambda_{\mathbb{R}^n}$-null set (see Remarks 3.2 and C.1 as well as Lemma 3.1). Hence, the general assumptions on $T = T_{Het}$, on $N^\dagger = N^* = \mathsf{B}$, on $\mathcal{L} = \mathfrak{M}_0^{lin}$, as well as on $\mathcal{V}$ in Proposition 5.12 of Pötscher and Preinerstorfer (2018) are satisfied in view of Part 1 of Lemma 5.16 in the same reference. [Alternatively, this can be gleaned from Lemma 3.1 and the attending discussion.] Next, observe that condition (A.4) is equivalent to

$$\operatorname{span}\left(\left\{\Pi_{\left(\mathfrak{M}_0^{lin}\right)^\perp} e_i(n) : i \in (n_{j-1}^+, n_j^+]\right\}\right) \nsubseteq \mathsf{B}$$

for every $j = 1, \ldots, m$, such that $(n_{j-1}^+, n_j^+] \cap I_1(\mathfrak{M}_0^{lin}) \neq \emptyset$, since $\Pi_{\left(\mathfrak{M}_0^{lin}\right)^\perp} e_i(n)$ and $e_i(n)$ differ only by an element of $\mathfrak{M}_0^{lin} \subseteq \operatorname{span}(X)$ and since $\mathsf{B} + \operatorname{span}(X) = \mathsf{B}$ (as noted in Lemma 3.1). In view of Proposition B.2 in Appendix B, this implies that any $\mathcal{S} \in \mathbb{J}(\mathfrak{M}_0^{lin}, \mathfrak{C}_{(n_1,\ldots,n_m)})$ is not contained in B, and thus not in $N^\dagger$. Using $\mathfrak{M}_0 \subseteq \operatorname{span}(X)$ and $\mathsf{B} + \operatorname{span}(X) = \mathsf{B}$, it follows that $\mu_0 + \mathcal{S} \nsubseteq \mathsf{B} = N^\dagger$ for every $\mu_0 \in \mathfrak{M}_0$. Since $\mu_0 + \mathcal{S}$ is an affine space and $N^\dagger = \mathsf{B}$ is a finite union of proper affine (even linear) spaces under Assumption 1 as discussed in Lemma

3.1, we may conclude (cf. Corollary 5.6 in Pötscher and Preinerstorfer (2018) and its proof) that $\lambda_{\mu_0+\mathcal{S}}(N^\dagger) = 0$ for every $\mathcal{S} \in \mathbb{J}(\mathfrak{M}_0^{lin}, \mathfrak{C}_{(n_1,\ldots,n_m)})$ and every $\mu_0 \in \mathfrak{M}_0$. This completes the verification of the assumptions of Proposition 5.12 in Pötscher and Preinerstorfer (2018) that are not specific to Part A (or Part B) of this proposition. We next verify the assumptions specific to Part A of this proposition: Assumption (a) is satisfied (even for every $C \in \mathbb{R}$) as a consequence of Part 2 of Lemma 5.16 in Pötscher and Preinerstorfer (2018) and of Remark C.1(i) above. And Assumption (b) in Part A follows from Lemma 5.19 of Pötscher and Preinerstorfer (2018), since T_{Het} results as a special case of the test statistics T_{GQ} defined in Section 3.4 of Pötscher and Preinerstorfer (2018) upon choosing $\mathcal{W}_n^* = n^{-1}\operatorname{diag}(d_i)$. Part A of Proposition 5.12 of Pötscher and Preinerstorfer (2018) now immediately delivers claim (A.3), since $C^* < \infty$ as noted in that proposition. That C^* and α^* do not depend on the choice of $\mu_0 \in \mathfrak{M}_0$ is an immediate consequence of $G(\mathfrak{M}_0)$-invariance of T_{Het}. Also note that α^* as defined in the theorem coincides with α^* as defined in Proposition 5.12 of Pötscher and Preinerstorfer (2018) in view of $G(\mathfrak{M}_0)$-invariance of T_{Het}. Positivity of α^* then follows from Part 5 of Lemma 5.15 in Preinerstorfer and Pötscher (2016) in view of Remark C.1(i), noting that $\lambda_{\mathbb{R}^n}$ and $P_{\mu_0,\Sigma}$ are equivalent measures (since $\Sigma \in \mathfrak{C}_{Het}$ is positive definite); cf. Remark 5.13(vi) in Pötscher and Preinerstorfer (2018). In case $\alpha < \alpha^*$, the remaining claim in Part (b) of the theorem, namely that equality can be achieved in (A.3), follows from the definition of C^* in Lemma 5.11 of Pötscher and Preinerstorfer (2018) and from Part A.2 of Proposition 5.12 of Pötscher and Preinerstorfer (2018) (and the observation immediately following that proposition allowing one to drop the suprema w.r.t. μ_0 and σ^2, and to set $\sigma^2 = 1$); in case $\alpha = \alpha^* < 1$, it follows from Remarks 5.13(i),(ii) in Pötscher and Preinerstorfer (2018) using Lemma 5.16 in the same reference.

The proof of Part (a) proceeds along the same lines with some minor differences: Observe that T_{uc} is a non-sphericity corrected F-type test with $N^\dagger = N^* = \operatorname{span}(X)$, which obviously is a closed $\lambda_{\mathbb{R}^n}$-null set (see Remark C.1(ii)), showing similarly that the general assumptions on $T = T_{uc}$, on $N^\dagger = N^* = \operatorname{span}(X)$, as well as on $\mathcal{L} = \mathfrak{M}_0^{lin}$ in Proposition 5.12 of Pötscher and Preinerstorfer (2018) are again satisfied (with $\mathfrak{C} = \mathfrak{C}_{(n_1,\ldots,n_m)}$). A similar, even simpler argument as in the proof of Part (b), again shows that condition (A.2) implies $\lambda_{\mu_0+\mathcal{S}}(N^\dagger) = 0$ for every $\mathcal{S} \in \mathbb{J}(\mathfrak{M}_0^{lin}, \mathfrak{C}_{(n_1,\ldots,n_m)})$ and every $\mu_0 \in \mathfrak{M}_0$, thus completing the verification of the assumptions of Proposition 5.12 of Pötscher and Preinerstorfer (2018) that are not specific to Part A (or Part B) of this proposition. Verification of Assumption (a) in Part A of Proposition 5.12 of Pötscher and Preinerstorfer (2018) proceeds exactly as before. For Assumption (b) we now use Lemma 5.19(iii) of Pötscher and Preinerstorfer (2018), since T_{uc} results as a special case of the test statistics $T_{E,\mathsf{W}}$ defined in Section 3 of Pötscher and Preinerstorfer (2018) upon choosing W as $n(n-k)^{-1}I_n$. Part A of Proposition 5.12 of Pötscher and Preinerstorfer (2018) then delivers the claim (A.1), again since $C^* < \infty$ as noted in that proposition. Again, $G(\mathfrak{M}_0)$-invariance of T_{uc} implies that C^* and α^* do not depend on the choice of $\mu_0 \in \mathfrak{M}_0$, and that α^* as defined in the theorem coincides with α^* as defined in Proposition 5.12 of Pötscher and Preinerstorfer (2018).

Positivity of α^* follows exactly as before making now use of Remark C.1(ii). The remaining claim in Part (a) is proved completely analogous as the corresponding claim in Part (b).

We finally prove Part (c): The claims follow from Remark 5.10 and Lemma 5.16 in Pötscher and Preinerstorfer (2018) combined with Remark C.1 above; cf. also Appendix A.3. ■

Proof of Proposition A.2: (a) This follows from Part 3 of Corollary 5.17 in Preinerstorfer and Pötscher (2016): As shown in the proof of Proposition 5.7(a) T_{uc} satisfies the assumptions of this corollary (with $\check{\beta} = \hat{\beta}$, $\check{\Omega}(y) = \hat{\sigma}^2(y) R (X'X)^{-1} R'$, $N = \emptyset$, and $N^* = \operatorname{span}(X)$). Set now $\mathcal{Z} = \operatorname{span}(\{e_i(n) : i \in (n_{j-1}^+, n_j^+]\})$, where j is such that $(n_{j-1}^+, n_j^+] \cap I_1(\mathfrak{M}_0^{lin}) \neq \emptyset$ and $\mathcal{Z} \subseteq \operatorname{span}(X)$ hold. Note that $\mathcal{Z}$ is not contained in $\mathfrak{M}_0^{lin}$ by construction. Observe that $\mathcal{Z}$ is a concentration space of $\mathfrak{C}_{(n_1,\ldots,n_m)}$ in view of Remark B.4 in Appendix B (note that $\operatorname{card}((n_{j-1}^+, n_j^+]) < n$ must hold in view of $\mathcal{Z} \subseteq \operatorname{span}(X)$ and $k < n$, while $0 < \operatorname{card}((n_{j-1}^+, n_j^+])$ is obvious). The nonnegative definiteness assumption on $\check{\Omega}$ in Part 3 of Corollary 5.17 in Preinerstorfer and Pötscher (2016) is clearly satisfied. Obviously $\check{\Omega}(z) = 0$ holds for every $z \in \mathcal{Z}$ since $\mathcal{Z} \subseteq \operatorname{span}(X)$. It remains to establish that $R\hat{\beta}(z) \neq 0$ holds $\lambda_{\mathcal{Z}}$-everywhere: Clearly, $R\hat{\beta}(z) = 0$ for $z \in \mathcal{Z}$ occurs precisely for $z \in \mathcal{Z} \cap \mathfrak{M}_0^{lin}$ since $\mathcal{Z} \subseteq \operatorname{span}(X)$. But $\mathcal{Z} \cap \mathfrak{M}_0^{lin}$ is a $\lambda_{\mathcal{Z}}$-null set in view of the fact that $\mathcal{Z}$ is not contained in $\mathfrak{M}_0^{lin}$ as noted before (and hence $\mathcal{Z} \cap \mathfrak{M}_0^{lin}$ is a proper linear subspace of $\mathcal{Z}$). Part 3 of Corollary 5.17 in Preinerstorfer and Pötscher (2016) then proves the claim for $C > 0$. A fortiori it then also holds for all real C.

(b) This follows in the same way as Part (a) by applying Part 3 of Corollary 5.17 in Preinerstorfer and Pötscher (2016) now to T_{Het} (with $\check{\beta} = \hat{\beta}$, $\check{\Omega} = \hat{\Omega}_{Het}$, $N = \emptyset$, and $N^* = \mathsf{B}$). ■

We note that Propositions 5.7 and A.2 could also be proved by making use of Theorem 3.1 in Pötscher and Preinerstorfer (2019).

Remark C.2. (i) Condition (8) ((10), respectively) in Theorem 5.1 can equivalently be written as $\operatorname{span}(\{\pi_{(\mathfrak{M}_0^{lin})^\perp,i}\}) \not\subseteq \operatorname{span}(X)$ ($\not\subseteq \mathsf{B}$, respectively) for every $i \in I_1(\mathfrak{M}_0^{lin})$ as discussed in the proof. Since the spaces $\operatorname{span}(\{\pi_{(\mathfrak{M}_0^{lin})^\perp,i}\})$ are one-dimensional for $i \in I_1(\mathfrak{M}_0^{lin})$ and since $1 < n - k + q = n - \dim(\mathfrak{M}_0^{lin})$, it follows that these spaces are necessarily elements of $\mathbb{J}(\mathfrak{M}_0^{lin}, \mathfrak{C}_{Het})$; in fact, they are precisely the minimal elements of $\mathbb{J}(\mathfrak{M}_0^{lin}, \mathfrak{C}_{Het})$ w.r.t. the order induced by inclusion.

(ii) Condition (A.2) ((A.4), respectively) in Theorem A.1 can equivalently be written as

$$\operatorname{span}(\{\Pi_{(\mathfrak{M}_0^{lin})^\perp} e_i(n) : i \in (n_{j-1}^+, n_j^+]\}) \not\subseteq \operatorname{span}(X) \ (\not\subseteq \mathsf{B}, \text{ respectively})$$

for every $j = 1, \ldots, m$ with $(n_{j-1}^+, n_j^+] \cap I_1(\mathfrak{M}_0^{lin}) \neq \emptyset$ as discussed in the proof. However, in this more general case, it can happen that such a space appearing on the l.h.s. of the non-inclusion relation has a dimension not smaller than $n - \dim(\mathfrak{M}_0^{lin})$, and hence is not a member of $\mathbb{J}(\mathfrak{M}_0^{lin}, \mathfrak{C}_{(n_1,\ldots,n_m)})$. In light of the general results in Pötscher and Preinerstorfer (2018) (e.g., Corollary 5.6) one may wonder if requiring the non-inclusion condition in (A.2) (A.4, respectively) for such spaces does not add an unnecessary restriction. However, this is not so as this non-inclusion is easily seen to be automatically satisfied for such spaces.[73] Furthermore, the collection of all spaces of the form $\operatorname{span}(\{\Pi_{(\mathfrak{M}_0^{lin})^\perp} e_i(n) : i \in (n_{j-1}^+, n_j^+]\})$ for $j = 1, \ldots, m$, such that $(n_{j-1}^+, n_j^+] \cap I_1(\mathfrak{M}_0^{lin}) \neq \emptyset$ and such that the dimension of these spaces is smaller than $n - \dim(\mathfrak{M}_0^{lin})$ is precisely the collection of minimal elements of $\mathbb{J}(\mathfrak{M}_0^{lin}, \mathfrak{C}_{(n_1,\ldots,n_m)})$ w.r.t. the order induced by inclusion. [Note that $\mathbb{J}(\mathfrak{M}_0^{lin}, \mathfrak{C}_{(n_1,\ldots,n_m)})$ may be empty.]

Proposition C.3. *Suppose we are in the setting of Example 5.5 with $n_j \geq 2$ for all j. Then T_{Het} is size controllable over $\mathfrak{C}_{Het}$, i.e., (9) holds for every $0 < \alpha < 1$.*

Proof: Note that B is a subset of

$$\mathsf{S} := \{y \in \mathbb{R}^n : \hat{u}_i(y) = 0 \text{ for some } i = 1, \ldots, n\},$$

and that S is a $\lambda_{\mathbb{R}^n}$-null set, as it is a finite union of $\lambda_{\mathbb{R}^n}$-null sets (since $e_i(n) \notin \operatorname{span}(X)$ in view of $n_j \geq 2$ for all j). Also note that $S_j > 0$ holds for $y \notin \mathsf{S}$. Now, for $y \notin \mathsf{S}$, by the Sherman-Morrison formula, the inverse of $S_1 \iota\iota' + \operatorname{diag}(S_2, \ldots, S_k)$ equals

$$\operatorname{diag}(S_2^{-1}, \ldots, S_k^{-1}) - \operatorname{diag}(S_2^{-1}, \ldots, S_k^{-1})\iota\iota' \operatorname{diag}(S_2^{-1}, \ldots, S_k^{-1}) / \sum_{j=1}^{k} 1/S_j.$$

We may thus write

$$T_{Het}(y) = \sum_{j=2}^{k} \frac{(\bar{y}_{(1)} - \bar{y}_{(j)})^2}{S_j} - \left[\sum_{j=2}^{k} \frac{\bar{y}_{(1)} - \bar{y}_{(j)}}{S_j}\right]^2 / \sum_{j=1}^{k} 1/S_j \quad \text{for every } y \notin \mathsf{S}. \tag{C.1}$$

As noted in Remark 3.3, for any invertible $q \times q$-dimensional matrix A, the test statistic T_{Het} based on R and the analogous test statistic, but computed with AR instead of R, coincide everywhere (note $r = 0$). We apply this observation in the following way: fix $l \in \{2, \ldots, k\}$, and choose A with l-th column $(-1, \ldots, -1)'$, l-th row$(0, \ldots 0, -1, 0, \ldots, 0)$, and such that after deleting the l-th column and the l-th row we obtain I_{q-1}. Then

$$AR = RP_l,$$

where P_l is the $k \times k$ permutation matrix that interchanges the first and l-th coordinate (and keeps all other coordinates fixed). By a similar computation as the one that led to the expression in (C.1), but now with RP_l in place of R, we can now conclude that for every $l \in \{1, \ldots, k\}$ we

[73] Note that any such space is necessarily equal to $(\mathfrak{M}_0^{lin})^\perp$. If now $(\mathfrak{M}_0^{lin})^\perp$ were contained in $\operatorname{span}(X)$ (B, respectively), then $\mathbb{R}^n$ would also have to be contained in $\operatorname{span}(X)$ (B, respectively), since $\mathbb{R}^n$ can be written as the direct sum of $(\mathfrak{M}_0^{lin})^\perp$ and $\mathfrak{M}_0^{lin}$ and since $\operatorname{span}(X)$ (B, respectively) are invariant under addition of elements of $\mathfrak{M}_0^{lin}$. However, $\operatorname{span}(X)$ is a proper subspace of $\mathbb{R}^n$ (since we always assume $k < n$) and B is a finite union of proper linear subspaces of $\mathbb{R}^n$ under Assumption 1. This gives a contradiction.

have

$$T_{Het}(y) = \sum_{j=1, j\neq l}^{k} \frac{(\bar{y}_{(l)} - \bar{y}_{(j)})^2}{S_j} - \left[\sum_{j=1, j\neq l}^{k} \frac{\bar{y}_{(l)} - \bar{y}_{(j)}}{S_j} \right]^2 / \sum_{j=1}^{k} 1/S_j \quad \text{for every } y \notin \mathsf{S}.$$

For $y \notin \mathsf{S}$ we may thus upper bound $T_{Het}(y)$ by $\sum_{j\neq l}^{k} (\bar{y}_l - \bar{y}_j)^2/S_j$, and we are free to choose l. Setting $l = l(y) \in \arg\min_{j=1,\ldots,k} S_j$, the upper bound for $T_{Het}(y)$ just derived, together with $S_j \geq (S_j + S_l)/2 > 0$, gives for $y \notin \mathsf{S}$

$$\begin{aligned} T_{Het}(y) &\leq \sum_{j=1, j\neq l}^{k} \frac{(\bar{y}_{(l)} - \bar{y}_{(j)})^2}{S_j} \leq 2 \sum_{j=1, j\neq l}^{k} \frac{(\bar{y}_{(l)} - \bar{y}_{(j)})^2}{S_j + S_l} \\ &\leq 2 \sum_{i,j=1, i\neq j}^{k} \frac{(\bar{y}_{(i)} - \bar{y}_{(j)})^2}{S_i + S_j} = 2 \sum_{i,j=1, i\neq j}^{k} T_{i,j}(y), \end{aligned}$$

where $T_{i,j}(y) = (\bar{y}_{(i)} - \bar{y}_{(j)})^2/(S_i + S_j)$. Note that the quantity to the far right does not depend on our particular choice of l. For $y \in \mathsf{S}$, define $T_{i,j}$ by the same formula as long as $S_i + S_j > 0$ and as $T_{i,j} = 0$ else. Since S is a $\lambda_{\mathbb{R}^n}$-null set, we have for any $C > 0$

$$\sup_{\mu_0 \in \mathfrak{M}_0} \sup_{0<\sigma^2<\infty} \sup_{\Sigma \in \mathfrak{C}_{Het}} P_{\mu_0, \sigma^2\Sigma}(T_{Het} \geq C) \leq \sum_{i,j=1, i\neq j}^{k} \sup_{\mu_0 \in \mathfrak{M}_0} \sup_{0<\sigma^2<\infty} \sup_{\Sigma \in \mathfrak{C}_{Het}} P_{\mu_0, \sigma^2\Sigma}(T_{i,j} \geq C/2(k^2 - k)).$$

Now observe that $T_{i,j}$ depends only on the coordinates of y corresponding to groups i and j and furthermore coincides with the test statistic of the form (4) for a two sample mean comparison as considered in Example 5.4 (with sample size being equal to $n_i + n_j$). A simple argument then shows that the terms in the sum on the r.h.s of the preceding display can be rewritten as the sizes of the test statistic (4) as considered in Example 5.4 with sample size now being given by $n_i + n_j$. Hence, all these terms can be made arbitrarily small by choosing C large enough by what has been established in Example 5.4. ■

We provide here a further example, where the sufficient condition of Part (b) of Theorem 5.1 fails, but size control is possible.

Example C.1. Suppose we are given $k \geq 2$ integers n_j describing group sizes satisfying $n_1 \geq 2$ and $n_j \geq 1$ for $j \geq 2$. Sample size is $n = \sum_{j=1}^{k} n_j$. Clearly $k < n$ is then satisfied. The regressors x_{ti} indicate group membership, i.e., they satisfy $x_{ti} = 1$ for $\sum_{j=1}^{i-1} n_j < t \leq \sum_{j=1}^{i} n_j$ and $x_{ti} = 0$ otherwise. The heteroskedasticity model is again given by $\mathfrak{C}_{Het}$. Let $R = (1, 0, \ldots, 0)$, i.e., the coefficient of the first regressor is subject to test. Then $I_0(\mathfrak{M}_0^{lin}) = \{\sum_{l=1}^{j} n_l : n_j = 1,\ j = 2, \ldots, k\}$. With regard to T_{uc} we immediately see that $e_i(n) \notin \operatorname{span}(X)$ for $i \in I_1(\mathfrak{M}_0^{lin})$ holds, and thus the sufficient condition (8) for size control of T_{uc} is satisfied. Turning to T_{Het}, observe that Assumption 1 is satisfied as is easily seen. Furthermore, it is not difficult to see that $\mathsf{B} = \{y \in \mathbb{R}^n : y_1 = \ldots = y_{n_1}\}$. Note that $\operatorname{span}(X) \subseteq \mathsf{B}$, but $\mathsf{B} \neq \operatorname{span}(X)$, except if $n_j = 1$ for all $j \geq 2$ holds. In the latter case it is then easy to see that $e_i(n) \notin \operatorname{span}(X) = \mathsf{B}$ for every $i \in I_1(\mathfrak{M}_0^{lin})$ holds, and thus the sufficient condition (10) for size control of T_{Het} is satisfied. But if $n_j > 1$ for some $j \geq 2$ holds, then for any index i satisfying $\sum_{l=1}^{j-1} n_l < i \leq \sum_{l=1}^{j} n_l$ we have $i \in I_1(\mathfrak{M}_0^{lin})$ as well as $e_i(n) \in \mathsf{B}$. Consequently, the sufficient condition (10) for size control of T_{Het} is *not* satisfied and hence Theorem 5.1 does not inform us about size controllability of T_{Het} in this case. However, the following argument shows that size control for T_{Het} is possible also in this case: The test statistic T_{Het} for the given problem coincides with a corresponding test statistic (again of the form (6) for an appropriate choice of d_i's) in the "reduced" problem that one obtains by throwing away all data points for $t > n_1$ and by also deleting all regressors from the regression model but the first one. This leads one to the heteroskedastic location model discussed in Example 5.3 albeit with sample size reduced to n_1. It is now not difficult to see that the size of T_{Het} in the original formulation of the problem coincides with the size of the corresponding test statistic in the "reduced" problem, which – in light of the discussion in Example 5.3 – shows that size control for T_{Het} in the original problem is possible also in the case where $n_j > 1$ for some $j \geq 2$ holds. [If $n_1 = 1$ and if $n_j \geq 2$ for some j, condition (8) in Theorem 5.1(a) is violated, implying – in view of Proposition 5.7 – that the size of the rejection region $\{T_{uc} > C\}$ is 1 for every choice of C; and that the test statistic T_{Het} is identically zero (since Assumption 1 is violated and, in fact, $\hat{\Omega}_{Het}$ is identically zero). The case where all n_j are equal to 1 even falls outside of our framework since we always require $n > k$.]

Remark C.4. Alternatively to the argument given in Example C.1 for the case where $n_j > 1$ for some $j \geq 2$ holds, size controllability of T_{Het} can also be established by the following reasoning: Keep the sample of size n, but replace the regressors $x_{\cdot i}$ for $2 \leq i \leq k$ by new regressors given by the standard basis vectors $e_j(n)$ for $j > n_1$ (the number of regressors now being $k^* = n - n_1 + 1 < n$ and $R = (1, 0, \ldots, 0)$ now being $1 \times k^*$). Then one observes that (i) this does not affect the test statistic, (ii) makes the set $\mathfrak{M}_0$ at most larger, and (ii) in the new model the sufficient condition (10) is now satisfied (as in the new model $n_j = 1$ holds for $j > n_1$). Hence, size control (even over the larger $\mathfrak{M}_0$) follows. A third possibility to establish the size-controllability result is to observe that the test statistic T_{Het} as well as the set B in the original model are – additional to being $G(\mathfrak{M}_0)$-invariant – also invariant w.r.t. addition of the elements $e_i(n)$ for $i > n_1$ and then to appeal to a generalization of Theorem 5.1 that exploits this additional invariance and provides sufficient conditions for size control that can be seen to be satisfied in the model considered in this example. Such a generalization of Theorem 5.1, which we refrain from stating, can be obtained from the general size control results presented in Pötscher and Preinerstorfer (2018).

Remark C.5. Example C.1 is an instance of the following observation: Suppose X is block-diagonal of rank k with blocks X_1 and X_2 where X_i is $n_i \times k_i$ with $n_1 + n_2 = n$ and $k_1 + k_2 = k$. Assume $k_1 < n_1$ (which entails $k < n$). Assume that the $q \times k$ restriction matrix R is of rank q and has the form $R = (R_1 : 0)$ with R_1 of dimension $q \times k_1$. The heteroskedasticity model is given by $\mathfrak{C}_{Het}$. Then, using the same reasoning as in Example C.1, we see that the question of size control of T_{Het} is equivalent to the question of size control of the corresponding test statistic in the "reduced" problem where one considers the regression model with regressor matrix equal to X_1 using only observations with $t \leq n_1$ (and as heteroskedasticity model the analogue of $\mathfrak{C}_{Het}$ for sample size n_1). As Example C.1 has shown, it is possible that the sufficient conditions for size control of T_{Het} are violated in the "original" problem, while at the same time the sufficient conditions may be satisfied in the "reduced" problem. Alternatively, one can argue similarly as in Remark C.4.

D Appendix: Proofs for Section 6 and Appendix A.2

Lemma D.1. *(a) Let $\mathcal{S}$ be a linear subspace of $\mathbb{R}^n$ and μ an element of $\mathbb{R}^n$ such that $\tilde{T}_{uc}$ restricted to $\mu+\mathcal{S}$ is not equal to a constant $\lambda_{\mu+\mathcal{S}}$-almost everywhere. Then $\lambda_{\mu+\mathcal{S}}(\tilde{T}_{uc} = C) = 0$ holds for every $C \in \mathbb{R}$.*

(b) $\lambda_{\mathbb{R}^n}(\tilde{T}_{uc} = C) = 0$ holds for every $C \in \mathbb{R}$.

(c) Let $\mathcal{S}$ be a linear subspace of $\mathbb{R}^n$ and μ an element of $\mathbb{R}^n$ such that $\tilde{T}_{Het}$ restricted to $\mu + \mathcal{S}$ is not equal to a constant $\lambda_{\mu+\mathcal{S}}$-almost everywhere. Then $\lambda_{\mu+\mathcal{S}}(\tilde{T}_{Het} = C) = 0$ holds for every $C \in \mathbb{R}$.

(d) Suppose Assumption 2 holds and $\tilde{T}_{Het}$ is not constant on $\mathbb{R}^n \backslash \tilde{\mathsf{B}}$. Then $\lambda_{\mathbb{R}^n}(\tilde{T}_{Het} = C) = 0$ holds for every $C \in \mathbb{R}$.

Proof: (a) Since $\tilde{T}_{uc}$ is constant on $\mathfrak{M}_0$ by definition, it follows that $\mu + \mathcal{S} \not\subseteq \mathfrak{M}_0$ must hold, and hence $\mathfrak{M}_0$ is a $\lambda_{\mu+\mathcal{S}}$-null set (cf. the argument in Remark 5.9(i) in Pötscher and Preinerstorfer (2018)). Consequently, $\tilde{T}_{uc}$ restricted to $(\mu + \mathcal{S})\backslash\mathfrak{M}_0$ is not constant. Suppose now there exists a $C \in \mathbb{R}$ so that $\lambda_{\mu+\mathcal{S}}(\{y \in \mathbb{R}^n : \tilde{T}_{uc}(y) = C\}) > 0$. Then, since $\mathfrak{M}_0$ is a $\lambda_{\mu+\mathcal{S}}$-null set as just shown, it follows that even $\lambda_{\mu+\mathcal{S}}(\{y \in \mathbb{R}^n\backslash\mathfrak{M}_0 : \tilde{T}_{uc}(y) = C\}) > 0$ must hold, which can be written as $\lambda_{\mu+\mathcal{S}}(\{y \in \mathbb{R}^n\backslash\mathfrak{M}_0 : p(y) = 0\}) > 0$, with the multivariate polynomial p given by $p(y) = (R\hat{\beta}(y) - r)' \left(R(X'X)^{-1}R'\right)^{-1} (R\hat{\beta}(y) - r) - C\tilde{\sigma}^2(y)$. This implies that p restricted to $\mu + \mathcal{S}$ vanishes on a set of positive $\lambda_{\mu+\mathcal{S}}$-measure. Since p restricted to $\mu + \mathcal{S}$ can clearly be expressed as a polynomial in coordinates parameterizing the affine space $\mu + \mathcal{S}$, it follows that p vanishes identically on $\mu + \mathcal{S}$. But this implies that $\tilde{T}_{uc}$ restricted to $(\mu + \mathcal{S})\backslash\mathfrak{M}_0$ is constant equal to C, a contradiction (as $\mathfrak{M}_0$ is a $\lambda_{\mu+\mathcal{S}}$-null set).

(b) Follows from Part (a) upon choosing $\mathcal{S} = \mathbb{R}^n$, if we can show that $\tilde{T}_{uc}$ is not $\lambda_{\mathbb{R}^n}$-almost everywhere constant. Given that $\tilde{T}_{uc}$ is continuous on $\mathbb{R}^n\backslash\mathfrak{M}_0$ (the complement of a proper affine subspace), it suffices to show that $\tilde{T}_{uc}$ is not constant on $\mathbb{R}^n\backslash\mathfrak{M}_0$. To this end consider first $y = X\beta$ with $R\beta - r \neq 0$ (such a β obviously exists). Observe that $\tilde{\sigma}^2(y) \neq 0$ as $y \notin \mathfrak{M}_0$ and that $R\hat{\beta}(y) - r = R\beta - r \neq 0$. Hence, $\tilde{T}_{uc}(y) \neq 0$ for this choice of y. Next, choose $y = X\beta + w$, where $R\beta - r = 0$ (such a β obviously exists) and where $w \neq 0$ is orthogonal to $\operatorname{span}(X)$ (which is possible since $k < n$ is always maintained). Then $\hat{\beta}(y) = \beta = \tilde{\beta}(y)$, implying $R\hat{\beta}(y) - r = R\beta - r = 0$ and $\tilde{\sigma}^2(y) = w'w/(n - (k - q)) \neq 0$. Note that $y \notin \mathfrak{M}_0$. It follows that $\tilde{T}_{uc}(y) = 0$ holds for this choice of y. This establishes non-constancy of $\tilde{T}_{uc}$ on $\mathbb{R}^n\backslash\mathfrak{M}_0$.

(c) Completely analogous to the proof of Part (a) except that $\tilde{T}_{uc}$ and $\mathfrak{M}_0$ are replaced by $\tilde{T}_{Het}$ and $\tilde{\mathsf{B}}$, respectively, and that p now takes the form $p(y) = (R\hat{\beta}(y) - r)' \operatorname{adj}(\tilde{\Omega}_{Het}(y))(R\hat{\beta}(y) - r) - C \det(\tilde{\Omega}_{Het}(y))$, where $\operatorname{adj}(\cdot)$ denotes the adjoint of the square matrix indicated, with the convention that the adjoint of a 1×1 dimensional matrix equals one. [We note that under the assumptions for Part (c) the set $\tilde{\mathsf{B}}$ cannot coincide with $\mathbb{R}^n$ (since otherwise $\tilde{T}_{Het}$ would be constant equal to zero), and thus Assumption 2 must hold.]

(d) Follows from Part (c) upon choosing $\mathcal{S} = \mathbb{R}^n$, if we can show that $\tilde{T}_{Het}$ is not $\lambda_{\mathbb{R}^n}$-almost everywhere constant. Given that $\tilde{T}_{Het}$ is continuous on $\mathbb{R}^n \backslash \tilde{\mathsf{B}}$ (the complement of a finite union of proper affine subspaces by Lemma 6.1), this follows from the assumed non-constancy on $\mathbb{R}^n \backslash \tilde{\mathsf{B}}$. ■

Remark D.2. The additional assumption that $\tilde{T}_{Het}$ is not constant on $\mathbb{R}^n \backslash \tilde{\mathsf{B}}$ in Part (d) of the preceding lemma can not be dropped as can be seen from the following example: Consider the case where $k = q = 1$, $R = 1$, $r = 0$, the regressor is given by $e_1(n)$, and the constants $\tilde{d}_i$ satisfy $\tilde{d}_i = 1$ for all i. Then $\mathfrak{M}_0 = \mathfrak{M}_0^{lin} = \{0\}$, Assumption 2 is satisfied, and $\tilde{\mathsf{B}} = \operatorname{span}(e_1(n))^{\perp}$. Furthermore, $\tilde{T}_{Het}(y) = 1$ for every $y \in \mathbb{R}^n \backslash \tilde{\mathsf{B}}$. As a point of interest we note that $\tilde{T}_{Het}$ is trivially size controllable for every $0 < \alpha < 1$, but that the condition (17) for size controllability is violated since $e_j(n) \in \tilde{\mathsf{B}}$ for $j > 1$. [Of course, neither a smallest size-controlling critical value exists (when considering rejection regions of the form $\{\tilde{T}_{Het} \geq C\}$) nor can exact size controllability be achieved for $0 < \alpha < 1$.] An extension of this example to the case $q = k > 1$ is discussed in the proof of Remark 6.10 given further below.

Lemma D.3. *Let C be a given critical value. Then the rejection probabilities $P_{\mu_0,\sigma^2\Sigma}(\tilde{T}_{uc} \geq C)$ as well as $P_{\mu_0,\sigma^2\Sigma}(\tilde{T}_{Het} \geq C)$ for $\mu_0 \in \mathfrak{M}_0$, $\sigma^2 \in (0,\infty)$, $\Sigma \in \mathfrak{C}_{Het}$, do not depend on r. [It is understood here that the constants $\tilde{d}_i$ appearing in the definition of $\tilde{T}_{Het}$ have been chosen independently of the value of r.]*

Proof: Fix $\mu_0 \in \mathfrak{M}_0$. Observe that $\tilde{T}_{Het}(y) = \tilde{T}^0_{Het}(y - \mu_0)$, where

$$\tilde{T}^0_{Het}(z) = \begin{cases} (R\hat{\beta}(z))' \left(\tilde{\Omega}^0_{Het}(z)\right)^{-1} (R\hat{\beta}(z)) & \text{if } \operatorname{rank} \tilde{B}^0(z) = q, \\ 0 & \text{if } \operatorname{rank} \tilde{B}^0(z) < q, \end{cases}$$

where

$$\tilde{\Omega}^0_{Het}(z) = R(X'X)^{-1}X' \operatorname{diag}\left(\tilde{d}_1(\tilde{u}^0_1(z))^2, \ldots, \tilde{d}_n(\tilde{u}^0_n(z))^2\right) X(X'X)^{-1}R',$$

where $\tilde{u}^0(z) = \Pi_{(\mathfrak{M}_0^{lin})^\perp} z$, and where $\tilde{B}^0(z) = R(X'X)^{-1}X' \operatorname{diag}(e_1'(n)\Pi_{(\mathfrak{M}_0^{lin})^\perp}(z), \ldots, e_n'(n)\Pi_{(\mathfrak{M}_0^{lin})^\perp}(z))$. Here we have made use of (13) and the fact that $\tilde{u}(y) = \tilde{u}^0(y - \mu_0)$. Now

$$P_{\mu_0,\sigma^2\Sigma}(\tilde{T}_{Het}(y) \geq C) = P_{\mu_0,\sigma^2\Sigma}(\tilde{T}^0_{Het}(y - \mu_0) \geq C) = P_{0,\sigma^2\Sigma}(\tilde{T}^0_{Het}(z) \geq C)$$

and the far right-hand side does not depend on r as $\tilde{T}^0_{Het}$ does not depend on r. The proof for

$\tilde{T}_{uc}$ is completely analogous, noting that $\tilde{T}_{uc}(y) = \tilde{T}^{0}_{uc}(y - \mu_0)$, where

$$\tilde{T}^{0}_{uc}(z) = \begin{cases} (R\hat{\beta}(z))' \left((\tilde{\sigma}^{0}(z))^2 R(X'X)^{-1}R'\right)^{-1} (R\hat{\beta}(z)) & \text{if } z \notin \mathfrak{M}_0^{lin}, \\ 0 & \text{if } z \in \mathfrak{M}_0^{lin}, \end{cases}$$

and where $(\tilde{\sigma}^{0}(z))^2 = (\tilde{u}^{0}(z))'\tilde{u}^{0}(z)/(n-(k-q))$. ■

Proof of Theorem 6.4: We first prove Part (b). We apply Part (b) of Theorem A.5 with $n_j = 1$ for $j = 1, \ldots, n = m$ observing that then $\mathfrak{C}_{(n_1,\ldots,n_m)} = \mathfrak{C}_{Het}$ and that condition (A.7) reduces to (17) (exploiting that $\tilde{\mathsf{B}} - \mu_0$ is a finite union of proper linear subspaces as discussed in Lemma 6.1). This establishes (16). The final claim in Part (b) of the theorem follows from Part (b) of Theorem A.5, if we can show that C^* given there can be written as claimed in Theorem 6.4: To this end we proceed as follows:[74] Choose an element μ_0 of $\mathfrak{M}_0$. Observe that $I_1(\mathfrak{M}_0^{lin}) \neq \emptyset$ (since $\dim(\mathfrak{M}_0^{lin}) = k - q < n$), and that for every $i \in I_1(\mathfrak{M}_0^{lin})$ the linear space $\mathcal{S}_i = \operatorname{span}(\Pi_{(\mathfrak{M}_0^{lin})^{\perp}} e_i(n))$ is 1-dimensional (since $\mathcal{S}_i = \{0\}$ is impossible in view of $i \in I_1(\mathfrak{M}_0^{lin})$), and belongs to $\mathbb{J}(\mathfrak{M}_0^{lin}, \mathfrak{C}_{Het})$ (since $n - k + q > 1 = \dim(\mathcal{S}_i)$ holds) in view of Proposition B.1 in Section B. Since $\tilde{T}_{Het}$ is $G(\mathfrak{M}_0)$-invariant (Remark 6.2), it follows that $\tilde{T}_{Het}$ is constant on $(\mu_0 + \mathcal{S}_i)\backslash\{\mu_0\}$, cf. the beginning of the proof of Lemma 5.11 in Pötscher and Preinerstorfer (2018). Hence, $\mathcal{S}_i$ belongs to $\mathbb{H}$ (defined in Lemma 5.11 in Pötscher and Preinerstorfer (2018)) and consequently for C^* as defined in that lemma

$$C^* \geq \max\left\{\tilde{T}_{Het}(\mu_0 + \Pi_{(\mathfrak{M}_0^{lin})^{\perp}} e_i(n)) : i \in I_1(\mathfrak{M}_0^{lin})\right\}$$

must hold. To prove the opposite inequality, let $\mathcal{S}$ be an arbitrary element of $\mathbb{H}$, i.e., $\mathcal{S} \in \mathbb{J}(\mathfrak{M}_0^{lin}, \mathfrak{C}_{Het})$ and $\tilde{T}_{Het}$ is $\lambda_{\mu_0+\mathcal{S}}$-almost everywhere equal to a constant $C(\mathcal{S})$, say. Then Proposition B.1 in Section B shows that $\mathcal{S}_i \subseteq \mathcal{S}$ holds for some $i \in I_1(\mathfrak{M}_0^{lin})$. Because of Condition (17) we have $\mu_0 + \mathcal{S}_i \nsubseteq \tilde{\mathsf{B}}$ since $\Pi_{(\mathfrak{M}_0^{lin})^{\perp}} e_i(n)$ and $e_i(n)$ differ only by an element of $\mathfrak{M}_0^{lin}$ and since $\tilde{\mathsf{B}} + \mathfrak{M}_0^{lin} = \tilde{\mathsf{B}}$. We thus can find $s \in \mathcal{S}_i$ such that $\mu_0 + s \notin \tilde{\mathsf{B}}$. Note that $s \neq 0$ must hold, since $\mu_0 \in \mathfrak{M}_0 \subseteq \tilde{\mathsf{B}}$ (see Lemma 6.1). In particular, $\tilde{T}_{Het}$ is continuous at $\mu_0 + s$, since $\mu_0 + s \notin \tilde{\mathsf{B}}$. Now, for every open ball A_ε in $\mathbb{R}^n$ with center s and radius $\varepsilon > 0$ we can find an element $a_\varepsilon \in A_\varepsilon \cap \mathcal{S}$ such that $\tilde{T}_{Het}(\mu_0 + a_\varepsilon) = C(\mathcal{S})$. Since $a_\varepsilon \to s$ for $\varepsilon \to 0$, it follows that $C(\mathcal{S}) = \tilde{T}_{Het}(\mu_0 + s)$. Since $s \neq 0$ and since $\tilde{T}_{Het}$ is constant on $(\mu_0 + \mathcal{S}_i)\backslash\{\mu_0\}$ as shown before, we can conclude that $C(\mathcal{S}) = \tilde{T}_{Het}(\mu_0 + s) = \tilde{T}_{Het}(\mu_0 + \Pi_{(\mathfrak{M}_0^{lin})^{\perp}} e_i(n))$, where we recall that $\Pi_{(\mathfrak{M}_0^{lin})^{\perp}} e_i(n) \neq 0$. But this now implies

$$C^* = \max\left\{\tilde{T}_{Het}(\mu_0 + \Pi_{(\mathfrak{M}_0^{lin})^{\perp}} e_i(n)) : i \in I_1(\mathfrak{M}_0^{lin})\right\}.$$

[74] Alternatively, one could base a proof on Lemma C.1 in Pötscher and Preinerstorfer (2019).

Using $G(\mathfrak{M}_0)$-invariance of $\tilde{T}_{Het}$ we conclude that

$$C^* = \max\left\{\tilde{T}_{Het}(\mu_0 + e_i(n)) : i \in I_1(\mathfrak{M}_0^{lin})\right\}.$$

We next prove Part (a): Apply Part (a) of Theorem A.5 with $n_j = 1$ for $j = 1, \ldots, n = m$, observing that then $\mathfrak{C}_{(n_1,\ldots,n_m)} = \mathfrak{C}_{Het}$. This establishes (15).[75] The final claim in Part (a) of the theorem follows similarly as the corresponding claim of Part (b) upon replacing the set $\tilde{\mathsf{B}}$ by $\mathfrak{M}_0$ in the argument, by noting that $\tilde{T}_{uc}$ is $G(\mathfrak{M}_0)$-invariant, and that $\mu_0 + \mathcal{S}_i \nsubseteq \mathfrak{M}_0$ holds because of $i \in I_1(\mathfrak{M}_0^{lin})$.

Part (c) follows from Part (c) of Theorem A.5 upon setting $n_j = 1$ for $j = 1, \ldots, n = m$ (and upon noting that then the conditions in Theorem A.5 reduce to the conditions of the present theorem). ■

Proof of Proposition 6.7: Follows from Part A.1 of Proposition 5.12 of Pötscher and Preinerstorfer (2018) and the sentence following this proposition. Note that the assumptions of this proposition have been verified in the proof of Theorem 6.4 (see also the proof of Theorem A.5, on which the proof of Theorem 6.4 is based), where it is also shown that the quantity C^* used in Proposition 5.12 of Pötscher and Preinerstorfer (2018) coincides with C^* defined in Theorem 6.4. ■

Proof of Remark 6.10: (i) From the definition of $\tilde{B}(y)$ and since here $\tilde{u}(y) = y - \mu_0$ we obtain for $i = 1, \ldots, n$

$$\tilde{B}(\mu_0 + e_i(n)) = R(X'X)^{-1}(0, \ldots, 0, x_{i\cdot}', 0, \ldots, 0)$$

where $x_{i\cdot}'$ appears in the i-th position (recall that $x_{i\cdot}'$ is the i-th column of X'). But then $\operatorname{rank}(\tilde{B}(\mu_0 + e_i(n))) \leq 1 < q$, implying that $\mu_0 + e_i(n) \in \tilde{\mathsf{B}}$. [In case $q = k = 1$, $\operatorname{rank}(\tilde{B}(\mu_0 + e_i(n))) = 1 = q$ for every $i = 1, \ldots, n$ whenever the matrix X has no zero entry. This then implies that (17) is satisfied. However, if X contains a zero at the m-th position, say, then $\tilde{B}(\mu_0 + e_m(n)) = 0 < 1 = q$, implying that $\mu_0 + e_m(n) \in \tilde{\mathsf{B}}$, thus leading to violation of (17) as $I_1(\mathfrak{M}_0^{lin}) = \{1, \ldots, n\}$.]

(ii) Define $\beta_0 = (X'X)^{-1}X'\mu_0$ and note that $R\beta_0 = r$ holds. Observing that R is nonsingular, that $\tilde{d}_i > 0$ for $1, \ldots, n$, and that $\tilde{u}(y) = y - \mu_0$, we obtain for $y \notin \tilde{\mathsf{B}}$

$$\begin{aligned}
\tilde{T}_{Het}(y) &= (y-\mu_0)' X \left[X' \operatorname{diag}\left(\tilde{d}_1\tilde{u}_1^2(y), \ldots, \tilde{d}_n\tilde{u}_n^2(y)\right) X\right]^{-1} X'(y-\mu_0) \\
&= \tilde{u}(y)'X\left[X' \operatorname{diag}\left(\tilde{d}_1\tilde{u}_1^2(y), \ldots, \tilde{d}_n\tilde{u}_n^2(y)\right) X\right]^{-1} X'\tilde{u}(y) \\
&\leq \left(\min_{1\leq i\leq n} \tilde{d}_i\right)^{-1} e'A(y)\left[A'(y)A(y)\right]^{-1} A'(y)e \leq n\left(\min_{1\leq i\leq n} \tilde{d}_i\right)^{-1}
\end{aligned}$$

where $e = (1, \ldots, 1)'$ and $A(y) = \operatorname{diag}(\tilde{u}_1(y), \ldots, \tilde{u}_n(y))X$. Note that $A'(y)A(y)$ is nonsingular

[75] This argument is actually superfluous since $\tilde{T}_{uc}$ is bounded as noted in Section 6.2.1.

for $y \notin \tilde{\mathsf{B}}$ and that the matrix in the quadratic form is a projection matrix. For $y \in \tilde{\mathsf{B}}$ we have $\tilde{T}_{Het}(y) = 0$. Hence, $\tilde{T}_{Het}(y)$ is bounded from above, and is trivially bounded from below as $\tilde{T}_{Het}(y) \geq 0$ for every $y \in \mathbb{R}^n$.

(iii) In the following examples we always set $\mu_0 = 0$ (i.e., $r = 0$) for the sake of simplicity. Remark D.2 provides an example where $\tilde{T}_{Het}$ is constant on $\mathbb{R}^n \backslash \tilde{\mathsf{B}}$. This example has $q = k = 1$. It can be easily extended to the case $q = k \geq 2$ by considering a design matrix X, the columns of which are given by the first k standard basis vectors, by setting $R = I_q$, and $\tilde{d}_i = 1$ for every $i = 1, \ldots, n$. Then $\tilde{T}_{Het}(y) = k$ for every $y \in \mathbb{R}^n \backslash \tilde{\mathsf{B}} = \{y \in \mathbb{R}^n : y_1 \neq 0, \ldots, y_k \neq 0\}$. An example where $\tilde{T}_{Het}$ is not constant on $\mathbb{R}^n \backslash \tilde{\mathsf{B}}$ is in case $q = k = 1$ given by the location model: Here $\tilde{T}_{Het}(y) = (\sum_{t=1}^n y_t)^2 / \sum_{t=1}^n y_t^2$ for every $y \in \mathbb{R}^n \backslash \tilde{\mathsf{B}} = \{y \in \mathbb{R}^n : y \neq 0\}$, which obviously is not constant (as $n > k = 1$).[76] This example can again be extended to the case $q = k > 1$ as follows: Let $R = I_q$ and let X be the design matrix where each of the columns correspond to a dummy variable describing membership in one of k disjoint groups G_j, each group of the same cardinality n_1 with $n_1 > 1$. Consequently, $n = kn_1$. W.l.o.g., we may assume that the elements G_1 have the lowest indices, followed by the elements of G_2, and so on. It is then easy to see that

$$\tilde{T}_{Het}(y) = \sum_{j=1}^{k} \left[(\sum_{t \in G_j} y_t)^2 / \sum_{t \in G_j} y_t^2 \right] \tag{D.1}$$

for $y \in \mathbb{R}^n \backslash \tilde{\mathsf{B}} = \bigcap_{j=1}^k \{y \in \mathbb{R}^n : y_t \neq 0 \text{ for at least one } t \in G_j\}$. Obviously, the expression in (D.1) is not constant: Choosing $y = e$ gives the value $kn_1 = n$, whereas choosing y such that $y_1 = y_{n_1+1} = y_{2n_1+1} = \ldots = y_{(k-1)n_1+1} = 1$ with all the other coordinates being zero gives a value of $k < n = kn_1$ since $n_1 > 1$.

Proof of Theorem 6.12: From the definition of C^* we see that C^* is nonnegative and finite. Let C be arbitrary but satisfying $C^* < C < \sup_{y \in \mathbb{R}^n} \tilde{T}_{Het}(y)$. We can then choose $y_0 \in \mathbb{R}^n$ with $\tilde{T}_{Het}(y_0) > C > 0$. In view of the definition of $\tilde{T}_{Het}$ it follows that $y_0 \notin \tilde{\mathsf{B}}$, and hence $\tilde{T}_{Het}$ is continuous at y_0. We can thus find an open neighborhood $U(y_0)$ of y_0 in $\mathbb{R}^n$ such that $\tilde{T}_{Het}$ is larger than C on $U(y_0)$. In particular, $P_{\mu_0,\Sigma}(\tilde{T}_{Het} \geq C) \geq P_{\mu_0,\Sigma}(U(y_0)) > 0$ for every $\mu_0 \in \mathfrak{M}_0$ and every $\Sigma \in \mathfrak{C}_{Het}$. This establishes $\alpha^* > 0$. Choose $\delta > 0$ such that $\delta \leq \alpha$ and $\delta < \alpha^*$. Then the size of the rejection region $\{\tilde{T}_{Het} \geq C_\Diamond(\delta)\}$ is exactly equal to δ by Parts (b) and (c) of Theorem 6.4. Consequently, $\{\tilde{T}_{Het} \geq C_\Diamond(\delta)\}$ is not a $\lambda_{\mathbb{R}^n}$-null set. By construction, $C_\Diamond(\alpha) \leq C_\Diamond(\delta)$ holds, and hence $\{\tilde{T}_{Het} \geq C_\Diamond(\alpha)\}$ contains $\{\tilde{T}_{Het} \geq C_\Diamond(\delta)\}$, which completes the proof. ■

Proof of Theorem A.5: We first prove Part (b). We wish to apply Part A of Proposition 5.12 of Pötscher and Preinerstorfer (2018) with $\mathfrak{C} = \mathfrak{C}_{(n_1,\ldots,n_m)}$, $T = \tilde{T}_{Het}$, $\mathcal{L} = \mathfrak{M}_0^{lin}$, and $\mathcal{V} = \{0\}$. First, note that $\dim(\mathfrak{M}_0^{lin}) = k - q < n$. Second, under Assumption 2, $\tilde{T}_{Het}$ is clearly Borel-measurable and is continuous on the complement of $\tilde{\mathsf{B}}$, where $\tilde{\mathsf{B}}$ is a closed $\lambda_{\mathbb{R}^n}$-null set

[76] In this example condition (17) is satisfied as $e_i(n) \notin \tilde{\mathsf{B}}$ for every $i = 1, \ldots, n$. To arrive at an example where again $\tilde{T}_{Het}$ is not constant on $\mathbb{R}^n \backslash \tilde{\mathsf{B}}$ but where condition (17) is not satisfied, consider the case where $X = (1, \ldots, 1, 0)'$ with $n \geq 2$. Observe that then $I_1(\mathfrak{M}_0^{lin}) = \{1, \ldots, n\}$.

(see Lemma 6.1 and the paragraph following this lemma). Because of Remark 6.2, we hence see that the general assumptions on $T = \tilde{T}_{Het}$, on $N^{\dagger} = \tilde{\mathsf{B}}$, on $\mathcal{L} = \mathfrak{M}_0^{lin}$, as well as on $\mathcal{V} = \{0\}$ in Proposition 5.12 of Pötscher and Preinerstorfer (2018) are satisfied. Next, observe that the validity of condition (A.7) clearly does not depend on the choice of $\mu_0 \in \mathfrak{M}_0$ since $\tilde{\mathsf{B}} + \mathfrak{M}_0^{lin} = \tilde{\mathsf{B}}$ as shown in Lemma 6.1. For the same reason condition (A.7) can equivalently be written as

$$\mu_0 + \operatorname{span}\left(\left\{\Pi_{\left(\mathfrak{M}_0^{lin}\right)^{\perp}} e_i(n) : i \in (n_{j-1}^{+}, n_j^{+}]\right\}\right) \nsubseteq \tilde{\mathsf{B}}$$

for every $j = 1, \ldots, m$, such that $(n_{j-1}^{+}, n_j^{+}] \cap I_1(\mathfrak{M}_0^{lin}) \neq \emptyset$, since $\Pi_{\left(\mathfrak{M}_0^{lin}\right)^{\perp}} e_i(n)$ and $e_i(n)$ differ only by an element of $\mathfrak{M}_0^{lin}$. In view of Proposition B.2 in Appendix B, this implies that $\mu_0 + \mathcal{S}$ for any $\mathcal{S} \in \mathbb{J}(\mathfrak{M}_0^{lin}, \mathfrak{C}_{(n_1,\ldots,n_m)})$ is not contained in $\tilde{\mathsf{B}}$, and thus not in $N^{\dagger}$. Since $\mu_0 + \mathcal{S}$ is an affine space and $N^{\dagger} = \tilde{\mathsf{B}}$ is a finite union of proper affine spaces under Assumption 2 as discussed in Lemma 6.1, we may conclude (cf. Corollary 5.6 in Pötscher and Preinerstorfer (2018) and its proof) that $\lambda_{\mu_0+\mathcal{S}}(N^{\dagger}) = 0$ for every $\mathcal{S} \in \mathbb{J}(\mathfrak{M}_0^{lin}, \mathfrak{C}_{(n_1,\ldots,n_m)})$ and every $\mu_0 \in \mathfrak{M}_0$. This completes the verification of the assumptions of Proposition 5.12 in Pötscher and Preinerstorfer (2018) that are not specific to Part A (or Part B) of this proposition. We next verify the assumptions specific to Part A of this proposition: Assumption (a) is satisfied (even for every $C \in \mathbb{R}$) as a consequence of Part (d) of Lemma D.1 (note that we have assumed that $\tilde{T}_{Het}$ is not constant on $\mathbb{R}^n \backslash \tilde{\mathsf{B}}$). And Assumption (b) in Part A follows from Part (c) of Lemma D.1. Part A of Proposition 5.12 of Pötscher and Preinerstorfer (2018) now immediately delivers claim (A.6), since $C^* < \infty$ as noted in that proposition. That C^* and α^* do not depend on the choice of $\mu_0 \in \mathfrak{M}_0$ is an immediate consequence of $G(\mathfrak{M}_0)$-invariance of $\tilde{T}_{Het}$. Also note that α^* as defined in the theorem coincides with α^* as defined in Proposition 5.12 of Pötscher and Preinerstorfer (2018) in view of $G(\mathfrak{M}_0)$-invariance of $\tilde{T}_{Het}$. In case $\alpha < \alpha^*$, the remaining claim in Part (b) of the theorem, namely that equality can be achieved in (A.3), follows from the definition of C^* in Lemma 5.11 of Pötscher and Preinerstorfer (2018) and from Part A.2 of Proposition 5.12 of Pötscher and Preinerstorfer (2018) (and the observation immediately following that proposition allowing one to drop the suprema w.r.t. μ_0 and σ^2, and to set $\sigma^2 = 1$); in case $\alpha = \alpha^* < 1$, it follows from Remarks 5.13(i),(ii) in Pötscher and Preinerstorfer (2018) using Part (d) of Lemma D.1. [In case $\alpha^* = 0$, there is nothing to prove.]

The proof of Part (a) proceeds similarly, but with some differences: Noting that $\tilde{T}_{uc}$ is clearly Borel-measurable and is continuous on the complement of $\mathfrak{M}_0$, where $\mathfrak{M}_0$ is a closed $\lambda_{\mathbb{R}^n}$-null set, and using Remark 6.2, we now see that the general assumptions on $T = \tilde{T}_{uc}$, on $N^{\dagger} = \mathfrak{M}_0$, on $\mathcal{L} = \mathfrak{M}_0^{lin}$, as well as on $\mathcal{V} = \{0\}$ in Proposition 5.12 of Pötscher and Preinerstorfer (2018) are satisfied (again with $\mathfrak{C} = \mathfrak{C}_{(n_1,\ldots,n_m)}$). Let now $\mathcal{S} \in \mathbb{J}(\mathfrak{M}_0^{lin}, \mathfrak{C}_{(n_1,\ldots,n_m)})$. In view of Proposition B.2 in Appendix B, $\mathcal{S}$ must then contain an element of the form $\Pi_{\left(\mathfrak{M}_0^{lin}\right)^{\perp}} e_i(n)$ for some $i \in I_1(\mathfrak{M}_0^{lin})$. Observe that $\Pi_{\left(\mathfrak{M}_0^{lin}\right)^{\perp}} e_i(n) \notin \mathfrak{M}_0^{lin}$ must hold, since otherwise we would have $e_i(n) \in \mathfrak{M}_0^{lin}$, contradicting $i \in I_1(\mathfrak{M}_0^{lin})$. It follows that $\mathcal{S} \nsubseteq \mathfrak{M}_0^{lin}$, and thus $\mu_0 + \mathcal{S} \nsubseteq \mathfrak{M}_0$ for every $\mu_0 \in \mathfrak{M}_0$. Since $\mu_0 + \mathcal{S}$ is an affine space and $N^{\dagger} = \mathfrak{M}_0$ is a proper affine space we may conclude (cf. Corollary 5.6 in Pötscher and Preinerstorfer (2018) and its proof) that $\lambda_{\mu_0+\mathcal{S}}(N^{\dagger}) = 0$ for every $\mathcal{S} \in \mathbb{J}(\mathfrak{M}_0^{lin}, \mathfrak{C}_{(n_1,\ldots,n_m)})$ and every $\mu_0 \in \mathfrak{M}_0$. We have thus now completed the verification of the assumptions of Proposition 5.12 of Pötscher and Preinerstorfer (2018) that are not specific to Part A (or Part B) of this proposition. We next verify the assumptions specific to Part A of this proposition: Verification of Assumptions (a) and (b) in Part A of Proposition 5.12 of Pötscher and Preinerstorfer (2018) proceeds similar as before except for now using Parts (b) and (a) of Lemma D.1. Part A of Proposition 5.12 of Pötscher and Preinerstorfer (2018) now immediately delivers claim (A.5), again since $C^* < \infty$ as noted in that proposition.[77] Again, $G(\mathfrak{M}_0)$-invariance of $\tilde{T}_{uc}$ implies that C^* and α^* do not depend on the choice of $\mu_0 \in \mathfrak{M}_0$, and that α^* as defined in the theorem coincides with α^* as defined in Proposition 5.12 of Pötscher and Preinerstorfer (2018). The remaining claim in Part (a) is proved completely analogous as the corresponding claim in Part (b) except for now using Part (b) of Lemma D.1.

We finally prove Part (c): The claims follow from Remark 5.10 in Pötscher and Preinerstorfer (2018) and Lemma D.1; cf. also Appendix A.3. ■

E Appendix: Algorithms

In this appendix, we discuss in more detail algorithms for determining (i) rejection probabilities, (ii) the size of a test based on one of the test statistics T_{Het}, T_{uc}, $\tilde{T}_{Het}$, or $\tilde{T}_{uc}$ together with a given candidate critical value, and (iii) size-controlling critical values. We discuss these algorithms under the Gaussianity assumption made in Section 2, but recall from Section 7.1 that the algorithms as given here can also be used to calculate *null* rejection probabilities, size, and size-controlling critical values in the elliptically symmetric case *without any changes*; similarly, the algorithms given here can also be used to calculate the size and size-controlling critical values in the semiparametric model discussed in (iv) of Section 7.1 as they stand. Furthermore, we restrict ourselves to the heteroskedasticity model $\mathfrak{C}_{Het}$; adapting the algorithms to subsets $\mathfrak{C}$ of $\mathfrak{C}_{Het}$ is rather straightforward (basically one has to appropriately constrain the optimization routines involved, appropriately redefine some of the quantities like C_{low}, and refer to the size-control conditions pertinent to the given heteroskedasticity model $\mathfrak{C}$).

E.1 Computing rejection probabilities

Suppose that a $G(\mathfrak{M}_0)$-invariant test statistic $T : \mathbb{R}^n \to \mathbb{R}$ has the following property: for some (and hence any) $\mu_0 \in \mathfrak{M}_0$ and a critical value $C \in \mathbb{R}$, there exists a symmetric $n \times n$ matrix A_C, such that

$$T(\mu_0 + z) \geq C \Leftrightarrow z' A_C z \geq 0 \text{ holds for } \lambda_{\mathbb{R}^n}\text{-almost every } z \in \mathbb{R}^n. \tag{E.1}$$

[77] This argument is actually superfluous since $\tilde{T}_{uc}$ is bounded as noted in Section 6.2.1. However, verification of the assumptions of Proposition 5.12 in Pötscher and Preinerstorfer (2018) is essential for the proof of the other claims in Part(a) of Theorem A.5.

If this property is satisfied, then for all choices of $\Sigma \in \mathfrak{C}_{Het}$, $\mu_0 \in \mathfrak{M}_0$, $\mu \in \mathfrak{M}$, and $\sigma^2 \in (0, \infty)$, setting $\nu := \sigma^{-1}\Sigma^{-1/2}(\mu - \mu_0)$, we may write

$$P_{\mu,\sigma^2\Sigma}(\{z \in \mathbb{R}^n : T(z) \geq C\}) = P_{\nu,I_n}(\{\zeta \in \mathbb{R}^n : \zeta'\Sigma^{1/2}A_C\Sigma^{1/2}\zeta \geq 0\}); \tag{E.2}$$

in case $\mu \in \mathfrak{M}_0$, we may set $\mu_0 = \mu$ to further simplify the right-hand-side in (E.2) to

$$P_{0,I_n}(\{\zeta \in \mathbb{R}^n : \zeta'\Sigma^{1/2}A_C\Sigma^{1/2}\zeta \geq 0\}). \tag{E.3}$$

The probability that a Gaussian quadratic form is not less than 0 (such as (E.2) or (E.3)) can numerically be determined by standard algorithms such as Davies (1980). Relation (E.1) can thus be exploited for efficiently computing rejection probabilities (for a given critical value), and thus plays an instrumental rôle in numerically determining the size of a test, size-controlling critical values, or the power function of a test.

For the important case $q = 1$ we now show that the above approach can indeed be used. It follows from the subsequent lemma that for any critical value C the property in (E.1) holds for the following test statistics: (i) T_{Het} provided Assumption 1 holds; (ii) T_{uc}; (iii) $\tilde{T}_{Het}$ provided Assumption 2 holds; (iv) $\tilde{T}_{uc}$. Recall from Lemmata 3.1 and 6.1 that under Assumption 1 (Assumption 2, respectively), the set B ($\tilde{\mathsf{B}}$, respectively) is a $\lambda_{\mathbb{R}^n}$-null set. Note that v defined in the lemma satisfies $v \neq 0$.

Lemma E.1. *Suppose $q = 1$. Let $v = v_{R,X} := X(X'X)^{-1}R'$. Then, for every $C \in \mathbb{R}$ and every $\mu_0 \in \mathfrak{M}_0$, we have:*

(a) If $\mu_0 + z \notin \mathsf{B}$, then $T_{Het}(\mu_0 + z) \geq C$ ($\leq C$) is equivalent to $z'A_{Het,C}z \geq 0$ (≤ 0), where

$$A_{Het,C} := vv' - C\Pi_{\mathrm{span}(X)^\perp} \operatorname{diag}\left(v_1^2 d_1, \ldots, v_n^2 d_n\right) \Pi_{\mathrm{span}(X)^\perp}. \tag{E.4}$$

(b) If $\mu_0 + z \notin \mathrm{span}(X)$, then $T_{uc}(\mu_0 + z) \geq C$ ($\leq C$) is equivalent to $z'A_{uc,C}z \geq 0$ (≤ 0), where

$$A_{uc,C} := vv' - C\frac{v'v}{n-k}\Pi_{\mathrm{span}(X)^\perp}. \tag{E.5}$$

(c) If $\mu_0 + z \notin \tilde{\mathsf{B}}$, then $\tilde{T}_{Het}(\mu_0 + z) \geq C$ ($\leq C$) is equivalent to $z'\tilde{A}_{Het,C}z \geq 0$ (≤ 0), where

$$\tilde{A}_{Het,C} := vv' - C\Pi_{(\mathfrak{M}_0^{lin})^\perp} \operatorname{diag}\left(v_1^2 \tilde{d}_1, \ldots, v_n^2 \tilde{d}_n\right) \Pi_{(\mathfrak{M}_0^{lin})^\perp}. \tag{E.6}$$

(d) If $\mu_0 + z \notin \mathfrak{M}_0$, then $\tilde{T}_{uc}(\mu_0 + z) \geq C$ ($\leq C$) is equivalent to $z'\tilde{A}_{uc,C}z \geq 0$ (≤ 0), where

$$\tilde{A}_{uc,C} := vv' - C\frac{v'v}{n-(k-1)}\Pi_{(\mathfrak{M}_0^{lin})^\perp}. \tag{E.7}$$

Proof: We first observe that there is nothing to prove in Part (a) (Part (c), respectively) if Assumption 1 (Assumption 2, respectively) is violated, since then $\mathsf{B} = \mathbb{R}^n$ ($\tilde{\mathsf{B}} = \mathbb{R}^n$, respectively) by Lemma 3.1 (Lemma 6.1, respectively). In the following we hence may assume for Part (a) (Part (c), respectively) that Assumption 1 (Assumption 2, respectively) hold, in which case B ($\tilde{\mathsf{B}}$, respectively) is a $\lambda_{\mathbb{R}^n}$-null set. The expressions in (E.4), (E.5), (E.6), and (E.7) now follow directly from the definitions of the test statistics since $q = 1$, recalling in particular that $\hat{u}(\mu_0+z) = \Pi_{\mathrm{span}(X)^\perp}(\mu_0+z) = \Pi_{\mathrm{span}(X)^\perp} z$, and $\tilde{u}(\mu_0+z) = \Pi_{(\mathfrak{M}_0^{lin})^\perp}((\mu_0+z)-\mu_0) = \Pi_{(\mathfrak{M}_0^{lin})^\perp} z$, and noting that for $q = 1$

$$R\hat{\beta}(\mu_0 + z) = r + v'z,$$
$$\hat{\Omega}_{Het}(\mu_0 + z) = z'\Pi_{\mathrm{span}(X)^\perp} \operatorname{diag}(v_1^2 d_1, \ldots, d_n v_n^2)\Pi_{\mathrm{span}(X)^\perp} z,$$
$$\tilde{\Omega}_{Het}(\mu_0 + z) = z'\Pi_{(\mathfrak{M}_0^{lin})^\perp} \operatorname{diag}(v_1^2 \tilde{d}_1, \ldots, \tilde{d}_n v_n^2)\Pi_{(\mathfrak{M}_0^{lin})^\perp} z,$$
$$\hat{\sigma}^2(\mu_0 + z) = \frac{z'\Pi_{\mathrm{span}(X)^\perp} z}{n-k},\ \tilde{\sigma}^2(z) = \frac{z'\Pi_{(\mathfrak{M}_0^{lin})^\perp} z}{n-(k-1)}$$

hold. ■

Remark E.2. The algorithm in Davies (1980) applied to (E.2) requires that the matrix A_C is not the zero matrix. In (i)-(iii) below we always have $q = 1$.

(i) It is easy to see that $A_{Het,C}$, $A_{uc,C}$, and $\tilde{A}_{uc,C}$ are never equal to the zero matrix: Note that $v'A_{Het,C}v = (v'v)^2 > 0$, since $v \in \mathrm{span}(X)$ and $v \neq 0$. The same argument applies to $A_{uc,C}$. Furthermore, for $C = 0$ the matrix $\tilde{A}_{uc,C}$ is obviously not the zero matrix; for $C \neq 0$ let $w \in (\mathfrak{M}_0^{lin})^\perp$, $w \neq 0$, w orthogonal to v, then $w'\tilde{A}_{uc,C}w = -w'wCv'v/(n-(k-1)) \neq 0$ (note that such a w exists, since $v \in (\mathfrak{M}_0^{lin})^\perp$ and $\dim((\mathfrak{M}_0^{lin})^\perp) = n-(k-q) > n-k \geq 1$ hold).

(ii) For $\tilde{A}_{Het,C}$ we have the following: Since $v \in (\mathfrak{M}_0^{lin})^\perp$ holds, $v'\tilde{A}_{Het,C}v = (v'v)^2 - Cv' \operatorname{diag}(v_1^2\tilde{d}_1, \ldots, v_n^2\tilde{d}_n)v$, which is zero only for $C = C_0$ where $C_0 = \sum_{i=1}^n v_i^2 / \sum_{i=1}^n v_i^4 \tilde{d}_i$ (note that the ratio is well-defined since all the $\tilde{d}_i$ are positive and since $v \neq 0$). Hence, $\tilde{A}_{Het,C}$ is not the zero matrix, except possibly for $C = C_0$. We now show that – in case Assumption 2 is satisfied – $\tilde{A}_{Het,C_0} = 0$ is equivalent to $\tilde{T}_{Het}(y)$ being constant for $y \in \mathbb{R}^n \backslash \tilde{\mathsf{B}}$: Suppose $\tilde{A}_{Het,C_0} = 0$. Since $C_0 > 0$, we obtain $\Pi_{(\mathfrak{M}_0^{lin})^\perp} \operatorname{diag}(v_1^2\tilde{d}_1, \ldots, v_n^2\tilde{d}_n)\Pi_{(\mathfrak{M}_0^{lin})^\perp} = vv'/C_0$ and thus $\tilde{A}_{Het,C} = vv'(1 - C/C_0)$. Fix $\mu_0 \in \mathfrak{M}_0$ arbitrary. For every $C > C_0$ we have $z'\tilde{A}_{Het,C}z \leq 0$ for every z, and hence for every z with $\mu_0 + z \notin \tilde{\mathsf{B}}$ (note that $\mathbb{R}^n \backslash \tilde{\mathsf{B}}$ is nonempty under Assumption 2). By Lemma E.1 we can conclude that $\tilde{T}_{Het}(\mu_0 + z) \leq C$ for every $\mu_0 + z \notin \tilde{\mathsf{B}}$. By the same token, we obtain that $\tilde{T}_{Het}(\mu_0 + z) \geq C$ for every $\mu_0 + z \notin \tilde{\mathsf{B}}$ when $C < C_0$ holds. We conclude that $\tilde{T}_{Het}(\mu_0 + z) = C_0$ for every $\mu_0 + z \notin \tilde{\mathsf{B}}$, i.e., $\tilde{T}_{Het}(y) = C_0$ for every $y \in \mathbb{R}^n \backslash \tilde{\mathsf{B}}$. To prove the converse, assume $\tilde{T}_{Het}(y) = C_1$ for every $y \notin \tilde{\mathsf{B}}$. Fix $\mu_0 \in \mathfrak{M}_0$ arbitrary. Then $\tilde{T}_{Het}(\mu_0 + z) = C_1$ for every z with $\mu_0 + z \notin \tilde{\mathsf{B}}$. By Lemma E.1 we get $z'\tilde{A}_{Het,C}z \geq 0$ (≤ 0, respectively) for $C \leq C_1$ ($C \geq C_1$, respectively) for every $z \notin \tilde{\mathsf{B}} - \mu_0$. Under Assumption 2 the set $\tilde{\mathsf{B}} - \mu_0$ is a $\lambda_{\mathbb{R}^n}$-null set, hence its complement is dense in $\mathbb{R}^n$. By continuity of the quadratic forms, we get $z'\tilde{A}_{Het,C}z \geq 0$ (≤ 0, respectively) for $C \leq C_1$ ($C \geq C_1$, respectively) for all $z \in \mathbb{R}^n$. We thus obtain $z'\tilde{A}_{Het,C_1}z = 0$ for every $z \in \mathbb{R}^n$. Since $\tilde{A}_{Het,C_1}$ is symmetric, $\tilde{A}_{Het,C_1} = 0$ follows and $C_1 = C_0$ must hold.

(iii) Before applying the algorithm in Davies (1980) to (E.2) with $T = T_{Het}$ and $A_C = A_{Het,C}$ we first check that Assumption 1 holds since otherwise Part (a) of the preceding lemma does not apply. In case of $T = \tilde{T}_{Het}$ and $A_C = \tilde{A}_{Het,C}$ we check that Assumption 2 holds for similar reasons; and, in case this assumption is satisfied, we then always also compute C_0 and check numerically that $\tilde{A}_{Het,C_0}$ (and hence any $\tilde{A}_{Het,C}$) is not the zero matrix.

In case $q > 1$, the algorithm in Davies (1980) could also be used to compute rejection probabilities for the tests based on T_{uc} and $\tilde{T}_{uc}$ as is easy to see. Since this is not so for T_{Het} and $\tilde{T}_{Het}$, we do not proceed in this way for reasons of comparability. In case $q > 1$ we thus compute the required rejection probabilities by Monte Carlo.

E.2 Determining the size of a test

For simplicity throughout this subsection T denotes any one of the test statistics UC, HC0-HC4, UCR, HC0R-HC4R. In case of HC0-HC4 we assume in our discussion that the design matrix X and R are such that Assumption 1 is satisfied, and in case of HC0R-HC4R we assume that Assumption 2 holds and that the test statistic is not constant on $\mathbb{R}^n \backslash \tilde{\mathsf{B}}$.[78] These conditions should be checked either theoretically or numerically before using the algorithms described below. Such numerical checks are implemented in the R-package **hrt** (Preinerstorfer (2021)) realizing these algorithms.

We now discuss algorithms for determining the size (over $\mathfrak{C}_{Het}$) of the test that rejects if $T \geq C$ for a given critical value $C > 0$ (note that any $C \leq 0$ leads to a trivial test that always rejects). By $G(\mathfrak{M}_0)$-invariance of T, for any given $\mu_0 \in \mathfrak{M}_0$, the size of this test simplifies to

$$\sup_{\Sigma \in \mathfrak{C}_{Het}} P_{\mu_0,\Sigma}(T \geq C), \tag{E.8}$$

which is what the algorithms described below compute numerically.

Before trying to determine the size numerically, it is advisable to check whether C is not less than the pertinent lower bound C^* for size-controlling critical values obtained in our theoretical results in Propositions 5.5 and 6.7 (and the attending footnotes), since otherwise one already knows that the size of the test is equal to 1, and hence there is no need to run the algorithm. The implementations of the algorithms in the R-package **hrt** (Preinerstorfer (2021)) have an option that provides such a check and outputs 1 if the check fails without running the algorithm.

Of course, the design matrix X, the restriction (R, r), and the particular choice of test statistic from the above list, are inputs to all the algorithms that are discussed in this and the subsequent section E.3, but we do not show these inputs explicitly in the descriptions of the algorithms given further down.

[78] This rules out trivial cases only.

E.2.1 Case $q = 1$

In the important special case $q = 1$ we can use (E.3) and Lemma E.1 to compute the rejection probabilities $P_{\mu_0,\Sigma}(T \geq C)$ appearing in (E.8) efficiently via, e.g., Davies (1980) (referred to as DA in what follows). A generic algorithm based on this observation is summarized in Algorithm 1.

Algorithm 1 Computing the size in case $q = 1$.

1: **Input** A real number $C > 0$ and positive integers $M_0 \geq M_1 \geq M_2$.
2: **Stage 0: Initial value search**
3: **for** $j = 1$ `to` $j = M_0$ **do**
4: Choose a candidate $\Sigma_j \in \mathfrak{C}_{Het}$.
5: Obtain $\tilde{p}_j := P_{\mu_0,\Sigma_j}(T \geq C)$ using DA.
6: **end for**
7: Rank the candidates Σ_j according to the value (from largest to smallest) of the corresponding quantities $\tilde{p}_j$ to obtain $\Sigma_{1:M_0}, \ldots, \Sigma_{M_1:M_0}$, the initial values for the next stage.
8: **Stage 1: Coarse localized optimizations**
9: **for** $j = 1$ `to` $j = M_1$ **do**
10: Obtain Σ_j^* by running a numerical algorithm for the optimization problem (E.8) initialized at $\Sigma_{j:M_0}$ and obtain $\bar{p}_{j,\Sigma_j^*} := P_{\mu_0,\Sigma_j^*}(T \geq C)$ (using DA to evaluate probabilities).
11: **end for**
12: Rank the obtained matrices Σ_j^* according to the value (from largest to smallest) of the corresponding $\bar{p}_{j,\Sigma_j^*}$ to obtain $\Sigma_{1:M_1}^*, \ldots, \Sigma_{M_2:M_1}^*$, the initial values for the next stage.
13: **Stage 2: Refined localized optimization**
14: **for** $j = 1$ `to` $j = M_2$ **do**
15: Obtain Σ_j^{**} by running a (refined) numerical algorithm for the optimization problem (E.8) initialized at $\Sigma_{j:M_1}^*$ and obtain $\bar{\bar{p}}_{j,\Sigma_j^{**}} := P_{\mu_0,\Sigma_j^{**}}(T \geq C)$ (using DA to evaluate probabilities).
16: **end for**
17: **Return** $\max_{j=1,\ldots,M_2} \bar{\bar{p}}_{j,\Sigma_j^{**}}$

Remark E.3. The initial values Σ_j in Stage 0 of Algorithm 1 can, for example, be obtained randomly (e.g., by sampling the diagonal elements of Σ_j from a uniform distribution on the unit simplex in $\mathbb{R}^n$). Such random choices may then be supplemented by "special" elements of $\mathfrak{C}_{Het}$, e.g., matrices that are close to $e_i(n)e_i(n)'$, $i = 1, \ldots, n$, or the matrix $n^{-1}I_n$, or a matrix Σ that maximizes the expectation of the quadratic form $y \mapsto y'\Sigma^{1/2}A_C\Sigma^{1/2}y$ under P_{0,I_n} (where A_C is obtained via Lemma E.1, cf. also the discussion preceding that lemma), the latter choice being motivated by (E.3). For the particular choice of initial values used in the R-package **hrt** and in our numerical calculations see Preinerstorfer (2021) and Appendix F.

Remark E.4. If Algorithm 1 is to be applied to a relatively large critical value C (say C larger than 5 times the $(1-\alpha)$-quantile of the cdf of $P_{0,I_N} \circ T$), then one may run Algorithm 1 on a smaller critical value first (e.g., the just mentioned quantile), and use the covariance matrix realizing the maximal rejection probability for this smaller critical value (in line 17 of Algorithm 1) as an additional initial value when running Algorithm 1 for determining the size corresponding to the originally given C. This can help to ameliorate numerical difficulties due to the rejection probabilities being close to zero over large portions of $\mathfrak{C}_{Het}$. The just described procedure is available as an option in the R-package **hrt**.

Remark E.5. The concrete choice of the numerical optimization algorithm used in Stages 1 and 2 of Algorithm 1 is left unspecified here, but may, for example, be a constrained Nelder and Mead (1965) algorithm (as provided in R's "constrOptim" function), where in Stage 2 the parameters in this algorithm (and in principle also in DA) should be chosen to guarantee a higher accuracy. For the particular choice of optimization routines used in the R-package **hrt** and in our numerical calculations see Preinerstorfer (2021) and Appendix F.

Remarks E.3 and E.5 also apply to other algorithms introduced further down, and will not be repeated.

E.2.2 General case

An algorithm that is similar to Algorithm 1, but uses Monte-Carlo simulation instead of DA to compute the rejection probabilities $P_{\mu_0,\Sigma}(T \geq C)$ is discussed in Algorithm 2; this algorithm is a modification of Algorithm 2 in Pötscher and Preinerstorfer (2018).[79] In Algorithm 2 the number of replications used in the Monte-Carlo simulations (and thus their accuracy but also their runtime) is increased in each stage, leading to an improved accuracy in the rejection probabilities computed. While this algorithm is also applicable in case $q = 1$, Algorithm 1 is to be preferred (and is automatically applied by the R-package **hrt** in this case), as it is based on a preferable way of computing the rejection probabilities.

E.3 Determining smallest size-controlling critical values

Again, in this subsection T denotes any one of the test statistics UC, HC0-HC4, UCR, HC0R-HC4R. In case of HC0-HC4 we assume in our discussion that the design matrix X and R are such that Assumption 1 is satisfied, and in case of HC0R-HC4R we assume that Assumption 2 holds and that the test statistic is not constant on $\mathbb{R}^n \backslash \tilde{\mathsf{B}}$.[80] *Furthermore, we assume that size-controlling critical values exist.* These conditions should be checked either theoretically or numerically before using the algorithms described below. The last mentioned existence can be guaranteed by checking (theoretically or numerically) the respective sufficient conditions for size control in Theorems 5.1 and 6.4.[81] We note that the implementations of the algorithms presented below in the R-package **hrt** (Preinerstorfer (2021)) include such numerical checks.

[79] This algorithm involves evaluating the test statistic T. Since the definition of T depends on invertibility of a covariance matrix estimator, an invertibility check is required. We use the same invertibility check as discussed in the second paragraph in Appendix E.3 of Pötscher and Preinerstorfer (2022), with a tolerance parameter that can be specified by the user.

[80] This rules out trivial cases only.

[81] In case the respective sufficient conditions are violated, but size-controlling critical values nevertheless exist (as, e.g., in Example 5.5 or in Remark 6.10), the algorithm still works.

Algorithm 2 Computing the size for general q.

Input A real number $C > 0$ and positive integers $M_0 \geq M_1 \geq M_2$, $N_0 \leq N_1 \leq N_2$.
Stage 0: Initial value search
for $j = 1$ `to` $j = M_0$ **do**
    Generate a pseudorandom sample $Z_1, \ldots, Z_{N_0}$ from P_{0,I_n}.
    Obtain a candidate $\Sigma_j \in \mathfrak{C}_{Het}$.
    Compute $\tilde{p}_j = N_0^{-1} \sum_{i=1}^{N_0} \mathbf{1}_{[C,\infty)}(T(\mu_0 + \Sigma_j^{1/2} Z_i))$.
end for
Rank the candidates Σ_j according to the value (from largest to smallest) of the corresponding quantities $\tilde{p}_j$ to obtain $\Sigma_{1:M_0}, \ldots, \Sigma_{M_1:M_0}$, the initial values for the next stage.
Stage 1: Coarse localized optimizations
for $j = 1$ `to` $j = M_1$ **do**
    Generate a pseudorandom sample $Z_1, \ldots, Z_{N_1}$ from P_{0,I_n}.
    Define $\bar{p}_{j,\Sigma} = N_1^{-1} \sum_{i=1}^{N_1} \mathbf{1}_{[C,\infty)}(T(\mu_0 + \Sigma^{1/2} Z_i))$ for $\Sigma \in \mathfrak{C}_{Het}$.
    Obtain Σ_j^* by running a numerical optimization algorithm for the problem $\sup_{\Sigma \in \mathfrak{C}_{Het}} \bar{p}_{j,\Sigma}$ initialized at $\Sigma_{j:M_0}$.
end for
Rank the obtained numbers Σ_j^* according to the value (from largest to smallest) of the corresponding $\bar{p}_{j,\Sigma_j^*}$ to obtain $\Sigma_{1:M_1}^*, \ldots, \Sigma_{M_2:M_1}^*$, the initial values for the next stage.
Stage 2: Refined localized optimization
for $j = 1$ `to` $j = M_2$ **do**
    Generate a pseudorandom sample $Z_1, \ldots, Z_{N_2}$ from P_{0,I_n}.
    Define $\bar{\bar{p}}_{j,\Sigma} = N_2^{-1} \sum_{i=1}^{N_2} \mathbf{1}_{[C,\infty)}(T(\mu_0 + \Sigma^{1/2} Z_i))$ for $\Sigma \in \mathfrak{C}_{Het}$.
    Obtain Σ_j^{**} by running a numerical optimization algorithm for the problem $\sup_{\Sigma \in \mathfrak{C}_{Het}} \bar{\bar{p}}_{j,\Sigma}$ initialized at $\Sigma_{j:M_1}^*$.
end for
Return $\max_{j=1,\ldots,M_2} \bar{\bar{p}}_{j,\Sigma_j^{**}}$.

We now proceed to discussing several algorithms for determining the *smallest* critical value $C_\Diamond(\alpha)$ such that the size of the test, which rejects if $T \geq C_\Diamond(\alpha)$, does not exceed α $(0 < \alpha < 1)$.[82] [In fact, for $C_\Diamond(\alpha)$ the size then equals α *provided* a critical value that results in size *equal* to α actually exists.] Note that $C_\Diamond(\alpha) > 0$ must hold, in view of Remarks 5.4 and 6.6 since $\alpha < 1$. By $G(\mathfrak{M}_0)$-invariance, for some fixed $\mu_0 \in \mathfrak{M}_0$, the algorithms numerically compute the smallest critical value that satisfies

$$\sup_{\Sigma \in \mathfrak{C}_{Het}} P_{\mu_0,\Sigma}(T \geq C) \leq \alpha, \tag{E.9}$$

cf. the discussion surrounding (E.8). For later use we denote by F_Σ the cdf of $P_{\mu_0,\Sigma} \circ T$, which by $G(\mathfrak{M}_0)$-invariance does not depend on the particular choice for $\mu_0 \in \mathfrak{M}_0$.

E.3.1 Computing smallest size-controlling critical values via line search based on algorithms in Section E.2

Given an algorithm $\mathsf{A} : (0, \infty) \to [0, 1]$ that for $C > 0$ returns the size of the test that rejects if $T \geq C$, one can use a line-search algorithm to determine the smallest critical value $C = C_\Diamond(\alpha)$ satisfying $\mathsf{A}(C) \leq \alpha$. To this end, one starts at the lower bound $C_{low} = \max(C^*, C_{\text{hom}})$, where C^* is given in the pertinent parts of Theorems 5.1 and 6.4, respectively (cf. also Propositions 5.5 and 6.7, respectively, and the attending footnotes), and C_{hom} denotes the *smallest* $1 - \alpha$ quantile of F_{I_n}, i.e., of the cdf of the test statistic under homoskedasticity. Note that then $P_{\mu_0,I_n}(T \geq C_{\text{hom}}) = \alpha$ and that $P_{\mu_0,I_n}(T \geq C) > \alpha$ for $C < C_{\text{hom}}$ (to see this note that F_{I_n} is continuous as $\{T = C\}$ is a $\lambda_{\mathbb{R}^n}$-null set for all real C, cf. Lemma 5.16 in Pötscher and Preinerstorfer (2018) and Lemma D.1 in Appendix D). Furthermore, $C_{\text{hom}} > 0$ (since $T \geq 0$ and $\{T = 0\}$ is a $\lambda_{\mathbb{R}^n}$-null set), and consequently $C_{low} > 0$ holds. Starting from C_{low}, one then keeps increasing the critical value "in a reasonable way" until one obtains, for the first time, a C such that $\mathsf{A}(C) \leq \alpha$ holds. This procedure is summarized in Algorithm 3, in which the particular algorithm A used is an input to Algorithm 3. For A one may either use Algorithm 1 if $q = 1$, or Algorithm 2 for general q. Note that one may need to terminate the while-loop after a maximal number of iterations.

Remark E.6. (i) Note that a matrix Σ^{**} as required for the while-loop in Algorithm 3 can easily be obtained by implementing Algorithm 1 or 2 in such a fashion as to also return the covariance matrix for which the maximal rejection probability is attained in the respective Stage 2.

(ii) A smallest C_+ as required in line 5 of Algorithm 3 indeed exists since $\{T = C\}$ is a $\lambda_{\mathbb{R}^n}$-null set for all real C as noted before.

(iii) For details regarding the computation of C_{low} in the R-package **hrt** see Preinerstorfer (2021) and Appendix F.2.

[82]Such a *smallest* size-controlling critical value indeed exists under the assumptions of this subsection (which includes existence of a size-controlling critical value) in view of Appendix A.3. [Under the sufficient conditions for size control in the respective theorems, this can also be read off directly from these theorems.]

Algorithm 3 Numerical approximation of the smallest size-controlling critical value via a line search algorithm.

Input $\alpha \in (0,1)$, A, C_{low}, $\epsilon \in [0, 1-\alpha)$ (ϵ a small tolerance parameter).
$C \leftarrow C_{low}$
while $\mathsf{A}(C) > \alpha + \epsilon$ **do**
    Let Σ^{**} be such that $P_{\mu_0,\Sigma^{**}}(T \geq C) \approx \mathsf{A}(C)$.
    Determine, by an upward line search initialized at C, the smallest value C_+ such that $P_{\mu_0,\Sigma^{**}}(T \geq C_+) \leq \alpha$.
    $C \leftarrow C_+$.
end while
return C

E.3.2 Computing smallest size-controlling critical values via quantile maximization

For completeness and comparison with Pötscher and Preinerstorfer (2018), we briefly describe an algorithm that is a modification of Algorithm 1 in Pötscher and Preinerstorfer (2018). In contrast to the algorithm discussed in the previous section, it does not make use of size-computations, but determines the smallest size-controlling critical value as

$$\sup_{\Sigma \in \mathfrak{C}_{Het}} F_{\Sigma}^{-1}(1-\alpha) \tag{E.10}$$

where F_{Σ}^{-1} denotes the quantile function of the cdf F_{Σ}. That (E.10) indeed gives the smallest size-controlling critical value is not difficult to see keeping in mind that $P_{\mu_0,\Sigma}(T = C) = 0$ for every real C, every $\mu_0 \in \mathfrak{M}_0$, and every $\Sigma \in \mathfrak{C}_{Het}$ (in view of $\lambda_{\mathbb{R}^n}(\{T = C\}) = 0$ as noted before). The algorithm is summarized in Algorithm 4.

F Appendix: Details concerning numerical computations in Section 11

F.1 Details concerning Section 11.1

To obtain Tables 1 and 2, for each of the test statistics UC, HC0-HC4, UCR, HC0R-HC4R, we repeated the procedure summarized in Algorithm 5 below 15 times (recall that $n = 25$, $R = (0,1)$, and $r = 0$). Each time this algorithm returned a design matrix, the corresponding size of the rejection region $\{T \geq C_{\chi^2,0.05}\}$ was obtained for the specific test statistic used, as well as a corresponding lower bound for the smallest size-controlling critical value. Then, we computed the maximum out of the 15 lower bounds, which (for each test statistic) is reported in Table 1. We also computed the maximum out of the 15 sizes, which (for each test statistic) is reported in Table 2. We also did the same with the critical value $C_{\chi^2,0.05}$ replaced by the 95%-quantile of an $F_{1,n-k}$-distribution ($n - k = 23$), the corresponding results being reported in Table 3.

Algorithm 4 Numerical approximation of the smallest size-controlling critical value via quantiles.

Input Positive integers $M_0 \geq M_1 \geq M_2$, $N_0 \leq N_1 \leq N_2$.
Stage 0: Initial value search
for $j = 1$ `to` $j = M_0$ **do**
  Generate a pseudorandom sample $Z_1, \ldots, Z_{N_0}$ from P_{0,I_n}.
  Obtain a candidate $\Sigma_j \in \mathfrak{C}_{Het}$.
  Compute $\tilde{F}_j^{-1}(1-\alpha)$ where $\tilde{F}_j(x) = N_0^{-1} \sum_{i=1}^{N_0} \mathbf{1}_{(-\infty,x]}(T(\mu_0 + \Sigma_j^{1/2} Z_i))$ for $x \in \mathbb{R}$.
end for
Rank the candidates Σ_j according to the value (from largest to smallest) of the corresponding quantities $\tilde{F}_j^{-1}(1-\alpha)$ to obtain $\Sigma_{1:M_0}, \ldots, \Sigma_{M_1:M_0}$, the initial values for the next stage.
Stage 1: Coarse localized optimizations
for $j = 1$ `to` $j = M_1$ **do**
  Generate a pseudorandom sample $Z_1, \ldots, Z_{N_1}$ from P_{0,I_n}.
  Define $\bar{F}_{j,\Sigma}(x) = N_1^{-1} \sum_{i=1}^{N_1} \mathbf{1}_{(-\infty,x]}(T(\mu_0 + \Sigma^{1/2} Z_i))$ for $x \in \mathbb{R}$ and $\Sigma \in \mathfrak{C}_{Het}$.
  Obtain Σ_j^* by running a numerical optimization algorithm for the problem $\sup_{\Sigma \in \mathfrak{C}_{Het}} \bar{F}_{j,\Sigma}^{-1}(1-\alpha)$ initialized at $\Sigma_{j:M_0}$.
end for
Rank the obtained Σ_j^* according to the value (from largest to smallest) of the corresponding $\bar{F}_{j,\Sigma_j^*}^{-1}(1-\alpha)$ to obtain $\Sigma_{1:M_1}^*, \ldots, \Sigma_{M_2:M_1}^*$, the initial values for the next stage.
Stage 2: Refined localized optimization
for $j = 1$ `to` $j = M_2$ **do**
  Generate a pseudorandom sample $Z_1, \ldots, Z_{N_2}$ from P_{0,I_n}.
  Define $\bar{\bar{F}}_{j,\Sigma}(x) = N_2^{-1} \sum_{i=1}^{N_2} \mathbf{1}_{(-\infty,x]}(T(\mu_0 + \Sigma^{1/2} Z_i))$ for $x \in \mathbb{R}$ and $\Sigma \in \mathfrak{C}_{Het}$.
  Obtain Σ_j^{**} by running a numerical optimization algorithm for the problem $\sup_{\Sigma \in \mathfrak{C}_{Het}} \bar{\bar{F}}_{j,\Sigma}^{-1}(1-\alpha)$ initialized at $\Sigma_{j:M_1}^*$.
end for
Return $\max_{j=1,\ldots,M_2} \bar{\bar{F}}_{j,\Sigma_j^{**}}^{-1}(1-\alpha)$.

In the description of Algorithm 5, the function $f(x)$ is an abbreviation for $C^* = \max\{T(\mu_0 + e_i(n)) : i \in I_1(\mathfrak{M}_0^{lin})\}$, the lower bound for the size-controlling critical values (cf. Propositions 5.5, 6.7, and the attending footnotes), with the $n \times 2$ design matrix X given by an intercept e, say, as the first column and a regressor x as the second one. Note that computing C^* necessitates the evaluation of the test statistic on a finite set of elements of $\mathbb{R}^n$, and then determining the maximum among the values obtained.[83] Concerning the evaluation of test statistics, the definition of which depends on the invertibility of a covariance matrix estimator, we used the same invertibility check as discussed in the second paragraph in Appendix E.3 of Pötscher and Preinerstorfer (2022) with a tolerance parameter of 10^{-8}. For $R = (0, 1)$ and for each matrix X returned by Algorithm 5 all relevant assumptions (i.e., the assumptions in the pertinent parts of Theorems 5.1 and 6.4, respectively) have been checked numerically.

Algorithm 5 Search procedure used for generating Tables 1, 2, and 3.

1: Initialize $x \leftarrow 0 \in \mathbb{R}^n$.
2: **for** $i = 1$ `to` $i = 5$ **do**
3: Generate an n-dimensional pseudo-random vector z of independent coordinates each from a log-standard normal distribution.
4: Run a Nelder and Mead (1965) algorithm initialized at z to maximize f over $\mathbb{R}^n$ (with a maximal number of iterations of 50, and otherwise the default parameters in R's "optim" function) to obtain z^*, say.
5: **if** $i = 1$, or $i \geq 2$ and $f(z^*) > f(x)$ **then**
6: $x \leftarrow z^*$.
7: **end if**
8: **if** $f(x) > 4$ **then**
9: Go to line 12.
10: **end if**
11: **end for**
12: Use Algorithm 1 to determine the size of the test for the test statistic under consideration for the design matrix (e, x) and based on either of the following two critical values: (i) $C_{\chi^2,0.05}$ and (ii) the 95% quantile of an $F_{1,n-k}$ distribution.
13: **return** x, $f(x)$, and the two sizes determined in the previous step.

Algorithm 5 uses Algorithm 1 in determining the size of a given test. We made the following choices concerning the parameters required in Algorithm 1 (and used default settings if not mentioned otherwise):

1. The candidates in Stage 0 of Algorithm 1 were determined by combining the suggestions in Remarks E.3 and E.4. That is, denoting $M_p = 200\ 000$, we combined: (i) sampling $M_p/4 - 1$ points from the unit simplex in $\mathbb{R}^n$, each corresponding to the diagonal of a

[83] In the present context $\mathfrak{M}_0^{lin}$ is spanned by the intercept. Thus, $I_1(\mathfrak{M}_0^{lin}) = \{1, \ldots, n\}$ holds since $n \geq 2$. In general, to determine $I_1(\mathfrak{M}_0^{lin})$ numerically, the algorithm implemented in the R-package **hrt** (Preinerstorfer (2021)) first obtains a basis for $\mathfrak{M}_0^{lin}$, and then checks for every $i = 1, \ldots, n$ whether or not the rank of the matrix obtained by appending the basis with $e_i(n)$ increases. This is done by a rank computation analogous to the one described in the last-but-one paragraph of Appendix E.3 of Pötscher and Preinerstorfer (2022), using the same function "rank" referred to there, and with tolerance parameter 10^{-8}.

matrix in $\mathfrak{C}_{Het}$, and sampling $3M_p/4+1$ points $\xi = (\xi_1, \ldots, \xi_n)$, say, analogously, each point ξ giving rise to a diagonal of a matrix in $\mathfrak{C}_{Het}$ via $(\xi_1^2, \ldots, \xi_n^2)/\sum_{i=1}^n \xi_i^2$; (ii) trying all diagonal matrices with a single dominant coordinate 0.9999 and the other coordinates all equal to $0.0001/(n-1)$, so that the trace equals 1; (iii) $n^{-1}I_n$; (iv) using a maximizer of the quadratic form described in Remark E.3; and (v) using an additional initial value in case of a "large" critical value C as described in Remark E.4, making use of the conventions discussed in parentheses in that remark. This results in $M_0 = M_p + n + 2$ and possible one more (in case C is large) candidates for initial values.

2. M_1 was chosen as 500, the optimization algorithm run in Stage 1 was a constrained Nelder and Mead (1965) algorithm (the default in R's "constrOptim" function), which was run with a relative tolerance parameter of 10^{-2} and a maximal number of iterations of $20n$.

3. M_2 was chosen as 1, the optimization algorithm run in Stage 1 was a constrained Nelder and Mead (1965) algorithm (the default in R's "constrOptim" function), which was run with a relative tolerance parameter of 10^{-3} and a maximal number of iterations of $30n$.

4. DA (used by Algorithm 1) was run with the parameters "acc = 10^{-3}" and "lim = 30000" using the function "davies" of the package **CompQuadForm**.

F.2 Details concerning Section 11.2

The smallest size-controlling critical values reported in Tables 4 and 5 in Section 11.2 were obtained by running Algorithm 3 (with algorithm A given by Algorithm 1 and a maximal number of 25 iterations in the while loop) as implemented in the R-package **hrt** (Preinerstorfer (2021)) version 1.0.0. Concerning A, the same input parameters as described in the enumeration at the end of Appendix F.1 were used but with $M_p = 500\,000$ (and with $n = 30$). Concerning Algorithm 3 we made the following choices for the required inputs:

1. $C_{low} = \max(C^*, C_{\text{hom}})$ is determined as follows: C_{hom} is determined by a line-search algorithm (using R's uniroot function and monotonicity of the rejection probabilities in the critical value) with the rejection probabilities obtained from DA (in case $q = 1$) or via Monte Carlo, whereas C^* is determined as described in Appendix F.1. For more detail see Preinerstorfer (2021).

2. ϵ was set to 10^{-3}.

For computing the power functions in Section 11.2, we made use of (E.2) with the matrices A_C given in Lemma E.1 together with the implementation of the algorithm by Davies (1980) in the R-package **CompQuadForm** (Duchesne and de Micheaux (2010)) version 1.4.3 and with default parameters.

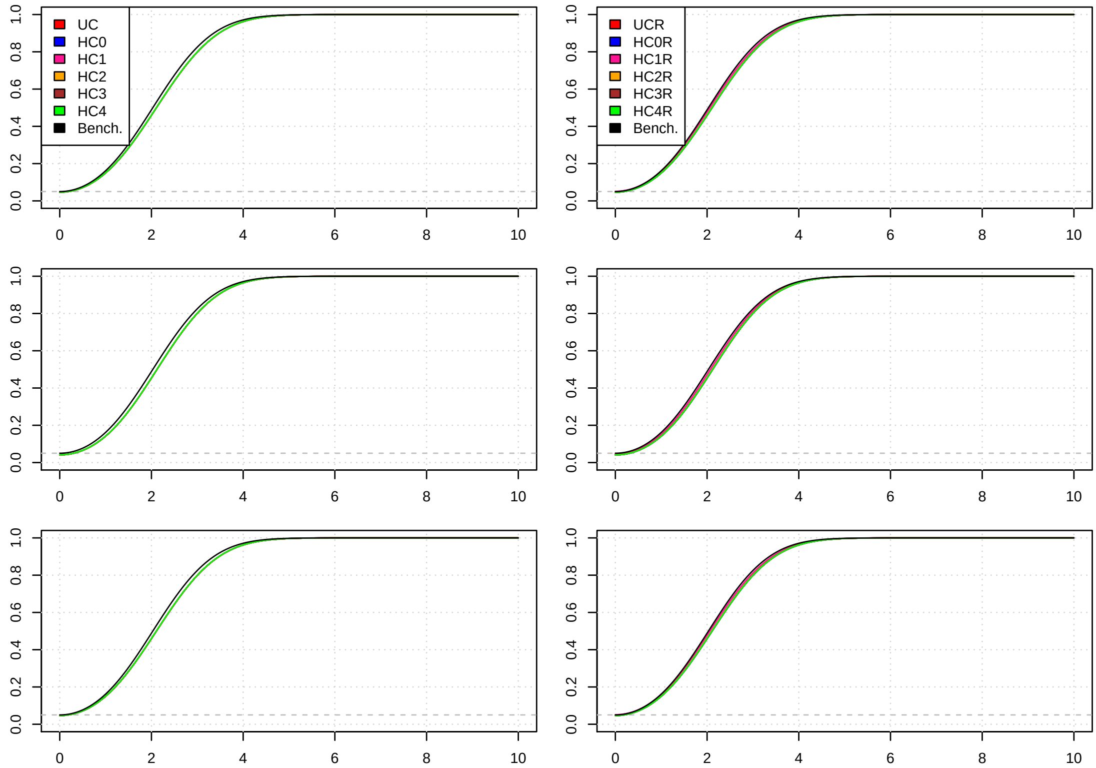

Figure F.1: Power functions for $n_1 = 15$. Left column: tests based on unrestricted residuals (cf. legend). Right column: tests based on restricted residuals (cf. legend). The rows corresponds to Σ_a for $a = 1, 5, 9$ from top to bottom. The abscissa shows δ. In the left panel the HC4-curve lies on top of the HC0–HC3-curves and the UC-curve. In the right panel the HC4R-curve lies on top of the HC0R-HC3R-curves and the UCR-curve. See the text for an explanation.

F.3 Additional figures for Section 11.2

The power functions for $n_1 = 15$ are given in Figure F.1.

G Appendix: Comments on Chu et al. (2021) and Hansen (2021)

In the special case of testing only one restriction (i.e., $q = 1$), Chu et al. (2021) and Hansen (2021) recently considered an interesting alternative approach to obtain tests based on the test statistics T_{Het} (for the commonly used choices of the weights d_i). Their suggestions are based on the observation (cf. also Section E.1 above) that, assuming Gaussianity of the errors, the null rejection probability of the test that rejects if T_{Het} exceeds a given critical value C can be rewritten as the probability that a quadratic form in Gaussian variables is nonnegative, which can efficiently be determined numerically for any given $\Sigma \in \mathfrak{C}_{Het}$ by a number of methods.[84] That is, if Σ were known, one could use this observation to numerically determine a critical value (an observation that is also exploited by our algorithms in case $q = 1$) or a p-value. Because Σ is, however, not known, this approach is infeasible. One solution, put forward in the present paper, is to work instead with a "worst-case" critical value, i.e., the smallest critical value that controls size (if such a critical value exists). In contrast, the idea in Chu et al. (2021) and Hansen (2021) to obtain a feasible test is a parametric bootstrap idea (cf. their papers for details):[85] (i) replace Σ by an estimate $\hat{\Sigma}$; e.g.,

$$\operatorname{diag}\left(d_1 \hat{u}_1^2\left(y\right), \ldots, d_n \hat{u}_n^2\left(y\right)\right) \tag{G.1}$$

based on typical choices of d_i; (ii) numerically determine a critical value (or p-value) from the cdf of the test statistic acting as if $\Sigma = \hat{\Sigma}$ (e.g., as outlined above); and (iii) reject the null hypothesis if the observed test statistic exceeds the so-computed critical value (or, equivalently, if the corresponding p-value obtained is less than the desired significance level). Note that the critical value in (ii) depends on the data Y through $\hat{\Sigma}$ (and is thus data-dependent in this sense).

No theoretical guarantees concerning the size of the tests proposed in Chu et al. (2021) and Hansen (2021) are given in these papers. Numerical results in both papers suggest that these parametric bootstrap tests can work well for certain design matrices X and hypotheses (R, r), but the authors also document some situations where the tests are considerably oversized. Hence, these tests are not valid, in general, which is in contrast to the procedure we suggest in the present paper. That a parametric bootstrap approach does not deliver size control is in line with results in Loh (1985) (see also Leeb and Pötscher (2017)) showing that under appropriate conditions parametric bootstrap procedures are oversized. It is also in line with a large body of literature on size distortions of (other) bootstrap-based tests for the testing problem under consideration, cf. Section 1 and Pötscher and Preinerstorfer (2022). As an aside we note that any *valid* data-dependent critical value, i.e., one that leads to a test with correct size (which is *not* the case for the proposals in Chu et al. (2021) and Hansen (2021)), must exceed the smallest size-controlling critical value with positive probability (or must be equal to the smallest size-controlling critical value with probability 1). Hence, a *valid* data-dependent critical value cannot always be smaller than the smallest size-controlling critical values, an observation that seems to have gone unnoticed in the discussion of the present article given in the introduction of Hansen (2021) (a discussion that also overlooks that one needs to take the square root of our critical values and lower bounds when discussing them in the context of the corresponding t-statistics).

To demonstrate further that the parametric bootstrap tests in Chu et al. (2021) and Hansen (2021) can be considerably oversized, we now report some numerical results for these tests. In particular, we report null rejection probabilities for a selection of points in the null hypothesis (i.e., for a selection of Σ's) and demonstrate that procedures suggested by Chu et al. (2021) and

[84] A reader has pointed out that the results in Phillips (1993) could also be developed into numerical approximations similar to the ones in Hansen (2021).

[85] A similar approach has already been put forward earlier by Welch (1938, 1951) and Satterthwaite (1946).

Hansen (2021) are not valid in the sense that these null rejection probabilities are considerably larger than the nominal significance level $\alpha = 0.05$ that is being used. Note that what we report are lower bounds for the size of the procedures investigated, which can even be larger, i.e., the overrejection problem can, in fact, be even more serious than what is seen in the tables below. Throughout, we study the following procedures

C: the procedure in Chu et al. (2021), when T_{Het} based on HC0-HC4 weights, respectively, is combined with the estimator $\hat{\Sigma}$ in (G.1) using the same weights as in the construction of the test statistic;

C3: the procedure as above, but where the estimator $\hat{\Sigma}$ in (G.1) always makes use of the HC3 weights;

H: the procedure in Hansen (2021) when T_{Het} is based on HC0-HC4 weights, respectively, and where for $\hat{\Sigma}$ the estimator suggested in Section 7 of Hansen (2021) is used.

We note here that procedure C3 is not considered in Chu et al. (2021); we include it, because $\hat{\Sigma}$ based on HC3 weights can be expected to perform better than if, e.g., HC0 weights are used. We also point out that Hansen (2021) only considers T_{Het} based on HC0-HC3 weights, but *not* on HC4; we also report rejection probabilities for the latter choice, because, in the examples we consider, it actually works better in terms of size than the choices considered in Hansen (2021).

Our implementations of the procedures in Chu et al. (2021) and Hansen (2021) rely on the algorithm in Davies (1980) (cf. Section E.1) to decide whether or not to reject (i.e., in Step (ii) of the description of that approach given further above in this section). To compute the rejection probabilities for the tests we used a Monte Carlo sample of size 100.000 for each of them. The nominal significance level used is $\alpha = 0.05$ throughout.

We consider three testing problems: The testing problems considered in Sections 11.2.1 and 11.2.2, as well as an additional one. Note that in all these examples the test statistics are size controllable and thus our test procedures based on smallest size-controlling critical values are applicable. [For Examples G.1 and G.2 this has already been discussed in Sections 11.2.1 and 11.2.2, respectively. For Example G.3 validity of Assumption 1 is obvious while condition (10) we have verified numerically.]

Example G.1. *(Comparing the means of two groups)* We here consider the same testing problem and setting (same n, n_1, n_2, α) as in Section 11.2.1. Table G.1 below shows the null rejection probabilities for the procedures C, C3, and H for the case $n_1 = 3$ and $a = 3$ (i.e., for $\Sigma = \Sigma_3$ defined in Section 11.2.1). We see from that table that the procedures suggested in Chu et al. (2021), i.e., procedures C, as well as the modification C3 are all considerably oversized (i.e., show rejection probabilities greater or equal to 2α). The methods using the idea in Hansen (2021) (including the case using HC4 weights not considered in Hansen (2021)) are slightly oversized in this example.

	HC0	HC1	HC2	HC3	HC4
C	0.14	0.14	0.12	0.11	0.10
C3	0.12	0.12	0.12	0.11	0.10
H	0.08	0.08	0.08	0.07	0.06

Table G.1: Null-rejection probabilities of the procedures C, C3, and H for comparing the means of two groups when $n_1 = 3$ and $a = 3$.

Example G.2. *(High-leverage design)* We here consider the same testing problem and setting (same n, α, X) as in Section 11.2.2. Table G.2 shows the null rejection probabilities for the procedures C, C3, and H for the case $a = 1$ (i.e., for $\Sigma = \Sigma_1^*$ defined in Section 11.2.2). The methods based on the approach in Chu et al. (2021) i.e., procedures C, as well as the modification C3 are all considerably oversized also in this example. The methods using the idea in Hansen (2021) are now also considerably oversized. The test using the HC4 estimator (which was not considered in Hansen (2021)) performs somewhat better and has a null rejection probability that exceeds the nominal significance level $\alpha = 0.05$ by a factor of 1.6.

	HC0	HC1	HC2	HC3	HC4
C	0.65	0.65	0.30	0.14	0.09
C3	0.18	0.18	0.17	0.14	0.10
H	0.16	0.16	0.15	0.12	0.08

Table G.2: Null-rejection probabilities of the procedures C, C3, and H for the high-leverage design matrix when $a = 1$.

The tables in the two preceding examples already show that the tests proposed by Chu et al. (2021) and Hansen (2021) can be considerably oversized. Note that the overrejection problem potentially is even more serious than what is seen from the tables as we have not searched over the space of Σ matrices, i.e., we have not reported size but only the null rejection probability at a particular value of Σ. Also, we have not made any attempt to search for design matrices X where overrejection is even more pronounced, but have only used design matrices from Sections 11.2.1 and 11.2.2.

We have seen in the preceding examples that pairing the method in Hansen (2021) with a HC4 based T_{Het} statistic performs more reasonably in these settings (it also is oversized, but less so). The question then arises whether there is some hope that this generalizes to other settings. The next example shows that this is unfortunately not the case.

Example G.3. We consider the same model and null hypothesis as in Section 11.2.2 except that the regressor x ($x \in \mathbb{R}^n$, $n = 30$) is different. Its entries x_i can be found plotted (against the index i) in Figure G.1. For this scenario one can prove (using similar arguments as in Pötscher and Preinerstorfer (2022)) that the size of the test obtained from pairing Hansen (2021)'s method with a HC4 based T_{Het} statistic actually equals 1. We do not give the details here but rather compute the null rejection probability of this test for Σ equal to the diagonal matrix with 0.999

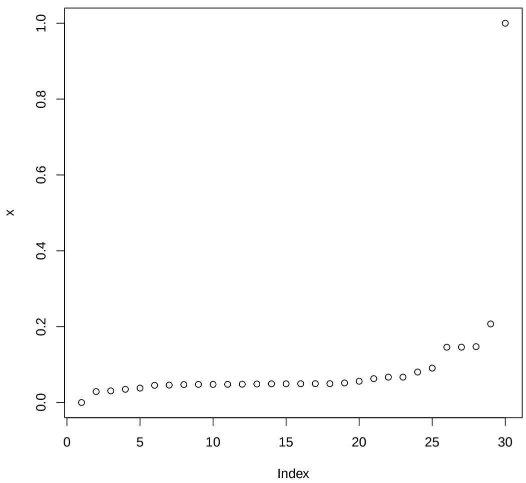

Figure G.1: Regressor used in Example G.3.

at the 21^{st} entry and the other diagonal entries constant so that the diagonal sums up to one. We used a Monte Carlo simulation (with 100.000 replications) and obtained a null rejection probability of 0.28, which is more than the five-fold nominal significance level.

References

Belloni, A. and Didier, G. (2008). On the Behrens-Fisher problem: a globally convergent algorithm and a finite-sample study of the Wald, LR and LM tests. *The Annals of Statistics*, **36** 2377–2408.

Chu, J., Lee, T.-H., Ullah, A. and Xu, H. (2021). Exact distribution of the F-statistic under heteroskedasticity of unknown form for improved inference. *Journal of Statistical Computation and Simulation*, **91** 1782–1801.

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

Duchesne, P. and de Micheaux, P. L. (2010). Computing the distribution of quadratic forms: Further comparisons between the Liu-Tang-Zhang approximation and exact methods. *Computational Statistics and Data Analysis*, **54** 858–862.

Hansen, B. (2021). The exact distribution of the White t-ratio. *Working paper, University of Wisconsin-Madison.*

Kim, S.-H. and Cohen, A. S. (1998). On the Behrens-Fisher problem: A review. *Journal of Educational and Behavioral Statistics*, **23** 356–377.

Leeb, H. and Pötscher, B. M. (2017). Testing in the presence of nuisance parameters: some comments on tests post-model-selection and random critical values. In *Big and complex data analysis*. Contrib. Stat., Springer, Cham, 69–82.

Lehmann, E. L. and Romano, J. P. (2005). *Testing Statistical Hypotheses.* 3rd ed. Springer Texts in Statistics, Springer, New York.

Loh, W.-Y. (1985). A new method for testing separate families of hypotheses. *Journal of the American Statistical Association*, **80** 362–368.

Mickey, M. R. and Brown, M. B. (1966). Bounds on the distribution functions of the Behrens-Fisher statistic. *Annals of Mathematical Statistics*, **37** 639–642.

Nelder, J. A. and Mead, R. (1965). A simplex method for function minimization. *The Computer Journal*, **7** 308–313.

Phillips, P. C. (1993). Operational algebra and regression t-tests. In *Models, Methods and Applications of Econometrics: Essays in Honor of A.R. Bergstrom* (P. C. Phillips, ed.). Oxford: Basil Blackwell, 140–152.

Pötscher, B. M. and Preinerstorfer, D. (2018). Controlling the size of autocorrelation robust tests. *Journal of Econometrics*, **207** 406–431.

Pötscher, B. M. and Preinerstorfer, D. (2019). Further results on size and power of heteroskedasticity and autocorrelation robust tests, with an application to trend testing. *Electronic Journal of Statistics*, **13** 3893–3942.

Pötscher, B. M. and Preinerstorfer, D. (2022). How reliable are bootstrap-based heteroskedasticity robust tests? *Econometric Theory,* first published online 22 April 2022. DOI.10.1017/S0266466622000184.

Preinerstorfer, D. (2021). *hrt: Heteroskedasticity Robust Testing.* R package version 1.0.0.

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

Ruben, H. (2002). A simple conservative and robust solution of the Behrens-Fisher problem. *Sankhyā Ser. A*, **64** 139–155.

Satterthwaite, F. E. (1946). An approximate distribution of estimates of variance components. *Biometrics Bulletin*, **2** 110–114.

Welch, B. L. (1938). The significance of the difference between two means when the population variances are unequal. *Biometrika*, **29** 350–362.

Welch, B. L. (1951). On the comparison of several mean values: an alternative approach. *Biometrika*, **38** 330–336.