A Details on Computation of Lower Bounds on Size

Without loss of generality, let the order of the clusters be such that the 1-1 element of Γ_j increases weakly in j. In the generic regression design, we computed the null rejection probability for $\{\Psi_j\}_{j=1}^q$ of the form $\Psi_j = \mathbf{1}[j \leq l]vv'$, $l = 1, \ldots, q$ or $\Psi_j = \mathbf{1}[j \geq l]vv'$ for $l = 2, \ldots, q$ and $v = 1$ for $k = 1$, $v = (1, s_1)'$ for $k = 2$ and $v = (1, s_1, s_2)'$ for $k = 3$, where $s_i \in \{-1, 0, 1\}$, $i = 1, 2$. Thus, we computed the null rejection probability over $(2q-1)3^{k-1}$ values of $\{\Psi_j\}_{j=1}^q$. In the two-sample design, $\psi_j = \mathbf{1}[j \leq l]$ with $l = 1, \ldots, q$ or $\psi_j = \mathbf{1}[j \geq l]$ for $l = 2, \ldots, q$, so that we considered $2q - 1$ different values of $\{\Psi_j\}_{j=1}^q$.

To avoid numerical inaccuracies stemming from bootstrap randomization, we simply computed the exact CGM bootstrap distribution generated by 2^q equally likely draws of $\{U_j\}_{j=1}^q$.

To speed up the computations, for each realization of $\{\Gamma_j\}_{j=1}^q$, the null rejection probabilities were initially computed for 2,500 draws of $\{Z_j\}_{j=1}^q$ for the values of $\{\Psi_j\}_{j=1}^q$ specified above. We then generated 10,000 new Monte Carlo draws of $\{Z_j\}_{j=1}^q$ under the maximizing value of $\{\Psi_j\}_{j=1}^q$, and re-estimated the null rejection probability. The resulting 100 values (one for each realization of $\{\Gamma_j\}_{j=1}^q$) are summarized in Tables 1 and 2.

B Proof of Theorem 1

The following Lemma is used in the proof of Theorem 1.

Lemma 1 *Let $Z_0, \ldots, Z_k \sim iid\mathcal{N}(0,1)$, and l_1,l_2 non-negative integers such that $k = l_1 + l_2 + 1$, and $w_0, w_1, w_2 > 0$. Define the $[0,1] \mapsto \mathbb{R}$ functions $g(s) = s^{(l_1+l_2+3)/2}(s+w_0)^{-1/2}(s+w_1)^{-l_1/2}(s+w_2)^{-l_2/2}$, $F_1(s) = -2\sqrt{1-s}/\sqrt{s}$ and $F_2(s) = 2\arcsin(\sqrt{s})$. Then, for any positive integer N,*

$$P\left(\frac{Z_0^2}{w_0 Z_1^2 + w_1 \sum_{i=2}^{l_1+1} Z_i^2 + w_2 \sum_{i=l_1+2}^{k} Z_i^2} > 1\right)$$
$$\leq \frac{1}{\pi}\sum_{i=2}^{N}\left[ig\left(\frac{i-1}{N}\right) - (i-1)g\left(\frac{i}{N}\right)\right]\left[F_1\left(\frac{i}{N}\right) - F_1\left(\frac{i-1}{N}\right)\right]$$
$$+ \frac{N}{\pi}\sum_{i=1}^{N}\left[g\left(\frac{i}{N}\right) - g\left(\frac{i-1}{N}\right)\right]\left[F_2\left(\frac{i}{N}\right) - F_2\left(\frac{i-1}{N}\right)\right].$$

Proof. By Lemma 2 and the development on page 10 of Bakirov and Székely (2005), the left hand side is equal to

$$\frac{1}{\pi}\int_0^1 s^{-3/2}(1-s)^{-1/2}g(s)ds.$$

The function $g : [0,1] \mapsto \mathbb{R}$ is convex, as can be checked by computing its second derivative. Thus, for any choice of $N \geq 1$, the piecewise linear function $\bar{g}_N$ that interpolates g at $s = 0/N, 1/N, \ldots, N/N$ satisfies $\bar{g}(s) \geq g(s)$ uniformly in $s \in [0,1]$. Therefore

$$\int_0^1 s^{-3/2}(1-s)^{-1/2}g(s)ds \leq \sum_{i=1}^{N}\int_{(i-1)/N}^{i/N} s^{-3/2}(1-s)^{-1/2}\bar{g}_n(s)ds.$$

The result follows, since the antiderivative of $s^{-3/2}(1-s)^{-1/2}$ is F_1, and the antiderivative of $s^{-1/2}(1-s)^{-1/2}$ is F_2. ■

Proof of Theorem 1:

The proof consists of 4 steps. First, we show that, by arguments similar to Bakirov (1998) and Bakirov and Székely (2005) ("B&S(05)" in the following), the supremum of the rejection probability is necessarily taken on for values of the variances that are either zero or take on some common value in each of the two groups. The only non-discrete element in the remaining parameter space is therefore the ratio of the non-zero variances of the two groups.

In the second step, we analyze the behavior of the t-statistic as a function of the number of non-zero variances in each group and the ratio of non-zero variances between the two groups.

In a third step, we construct an upper bound for the rejection probability for cases where the inequality $P(|t| > \text{cv}) \leq \alpha$ has some slack (equality is taken on when the smaller group has all equal and positive variances, and the larger group has all variances equal to zero). Given the slack, this can be done relatively easily by deriving upper bounds of the rejection probability on small ranges for the value of the ratio of the variances. These upper bounds are then shown to be below α via Lemma 1 and a direct computation.

The fourth step deals with the remaining "hard" part of the parameter space, where $P(|t| > \text{cv})$ is tangent to α. Here, an analytical inequality is derived (with no slack at the point $P(|t| > \text{cv}) \leq \alpha$). This inequality is then shown to be convex in the ratio of the variances, reducing the problem to checking the inequality for the extreme values. The result then follows from what is obtained in B&S(05) (since the two sample t-statistic reduces to the one sample t-statistic if the ratio of the variances is zero), and another direct computation via Lemma 1.

Step 1: Without loss of generality, set $\mu_1 = \mu_2 = 0$. Let $n = q_1$ and $m = q_2$, $D = \operatorname{diag}(\sigma_{1,1}, \cdots, \sigma_{1,n}, \sigma_{2,1}, \cdots, \sigma_{2,m})$, and $Z \sim \mathcal{N}(0, I_{n+m})$, in notation similar to B(98). Let e_l be a $l \times 1$ vector of ones, and $M_l = I_l - l^{-1} e_l e_l'$. For any critical value cv

$$\begin{aligned}
P(|t| > \mathrm{cv}) &= \Pr\left((\bar{X}_1 - \bar{X}_2)^2 > \mathrm{cv}\left(\frac{s_1^2}{q_1} + \frac{s_2^2}{q_2}\right)\right) \\
&= P\left(Z'D\left(\begin{array}{c} e_n/n \\ -e_m/m \end{array}\right)\left(\begin{array}{c} e_n/n \\ -e_m/m \end{array}\right)'\right. \\
&\qquad\qquad \left. - \mathrm{cv}\left(\begin{array}{cc} M_n/(n(n-1)) & 0 \\ 0 & M_m/(m(m-1)) \end{array}\right)\Bigg] DZ > 0\right) \\
&= (PZ'DADZ \leq 0)
\end{aligned}$$

with

$$A = \left(\begin{array}{cc} \mathrm{cv}\, M_n m/(n-1) - e_n e_n' m/n & e_n e_m' \\ e_m e_n' & \mathrm{cv}\, M_m n/(m-1) - e_m e_m' n/m \end{array}\right).$$

Further, the characteristic polynomial of DAD is

$$\begin{aligned}
f(\lambda) &= \det(\lambda I_{n+m} - DAD) \\
&= -\det(D^2)\det(-\lambda D^{-2} + A).
\end{aligned}$$

By Proposition 1 of Bakirov (1998), this evaluates to

$$\begin{aligned}
f(\lambda) &= -\left(\prod_{i=1}^{n}(\mathrm{cv}\,\frac{m}{n-1}\sigma_{1,i}^2 - \lambda)\right)\left(\prod_{i=1}^{m}(\mathrm{cv}\,\frac{n}{m-1}\sigma_{2,i}^2 - \lambda)\right)\Bigg[1 + \\
&\qquad r\left(\sum_{i=1}^{n}\frac{1}{\mathrm{cv}\,\frac{m}{n-1} - \lambda\sigma_{1,i}^{-2}}\right) + d\left(\sum_{i=1}^{m}\frac{1}{\mathrm{cv}\,\frac{n}{m-1} - \lambda\sigma_{2,i}^{-2}}\right) + \\
&\qquad\qquad (rd-1)\left(\sum_{i=1}^{n}\frac{1}{\mathrm{cv}\,\frac{m}{n-1} - \lambda\sigma_{1,i}^{-2}}\right)\left(\sum_{i=1}^{m}\frac{1}{\mathrm{cv}\,\frac{n}{m-1} - \lambda\sigma_{2,i}^{-2}}\right)\Bigg] \\
&= \prod_{i=1}^{n}(x_i - \lambda)\prod_{i=1}^{m}(y_i - \lambda)\Bigg[a\sum_{i=1}^{n}\frac{x_i}{x_i - \lambda} + b\sum_{i=1}^{m}\frac{y_i}{y_i - \lambda} - 1 - \\
&\qquad\qquad C\left(\sum_{i=1}^{n}\frac{x_i}{x_i - \lambda}\right)\sum_{i=1}^{m}\frac{y_i}{y_i - \lambda}\Bigg]
\end{aligned}$$

where

$$
\begin{aligned}
r &= -\left(\frac{m}{n} + \text{cv}\,\frac{m}{n(n-1)}\right) \\
d &= -\left(\frac{n}{m} + \text{cv}\,\frac{n}{m(m-1)}\right) \\
x_i &= \text{cv}\,\frac{m}{n-1}\sigma_{1,i}^2 \\
y_i &= \text{cv}\,\frac{n}{m-1}\sigma_{2,i}^2 \\
a &= \frac{\frac{m}{n} + \text{cv}\,\frac{m}{n(n-1)}}{\text{cv}\,\frac{m}{n-1}} = \frac{n-1+\text{cv}}{n\,\text{cv}} \\
b &= \frac{\frac{n}{m} + \text{cv}\,\frac{n}{m(m-1)}}{\text{cv}\,\frac{n}{m-1}} = \frac{m-1+\text{cv}}{m\,\text{cv}} \\
C &= ab - \frac{1}{\text{cv}\,\frac{m}{n-1}\,\text{cv}\,\frac{n}{m-1}} = \frac{\text{cv}-2+m+n}{\text{cv}\,mn}
\end{aligned}
$$

Thus f, viewed as a function of $\{x_i\}_{i=1}^n$ only, has the same form as in B&S(05), and one can argue analogously that the largest rejection probability is obtained with all variances $\sigma_{1,i}$ to be the same or zero. More specifically, note that f has a single negative root. Without loss of generality, normalize $\sigma_{1,i}$ and $\sigma_{2,i}$ such that this root takes on the value -1. This amounts to

$$
a\sum_{i=1}^{n}\frac{x_i}{x_i+1} + b\sum_{i=1}^{m}\frac{y_i}{y_i+1} - C\left(\sum_{i=1}^{n}\frac{x_i}{x_i+1}\right)\sum_{i=1}^{m}\frac{y_i}{y_i+1} = 1. \tag{B1}
$$

Apply Lemma 2 and the development on page 10 of B&S(05) to obtain

$$
P(|t| > \text{cv}) = \frac{1}{\pi}\int_0^1 \frac{s^{\frac{n+m}{2}-1}}{\sqrt{f(-s)}}ds.
$$

Now fix all x_j, $j \geq 3$ and y_j, $j \geq 1$, and consider x_2 as a function of x_1 via (B1). Since as a function of x_1 and x_2, $f(-s)$ and (B1) have the same functional form as in B&S (05), their reasoning in (9)-(12) and their Lemma 3 still holds. We can conclude that at the point of maximum, either $x_1 = x_2$, or $x_1 x_2 = 0$. Applying the same argument to arbitrary x_i, x_j, and also to arbitrary pairs of y_i, y_j, we conclude that $P(|t| > \text{cv})$ takes on the largest value if $n^* \leq n$ values of $\sigma_{1,j}^2$ are equal to 1, the remaining values of $\sigma_{1,j}^2$ are zero, $m^* \leq m$ values of $\sigma_{2,j}^2$ are equal to ρ^2, $\rho \geq 0$, and the remaining ones are equal to zero.

Step 2: For $\min(m^*, n^*) = 1$ the result follows from B&S(05), since the t-statistic then reduces to a one-sample t-statistic. Therefore, assume $\min(m^*, n^*) > 1$ in the following.

Further, assume $\rho \leq 1$ without loss of generality, as we do not put any constraint on whether $m \geq n$ or $n \leq m$. Finally, again without loss of generality, let $X_i = Y_j = 0$ almost surely for $i > n^*$, $j > m^*$. The remainder of the proof considers all possible values for cv, n, m, n^* and m^* that satisfy these conditions. The only continuous variable that we have to consider is ρ.

With $\tilde{X} = n^{*-1} \sum_{i=1}^{n^*} X_i$, we have

$$\begin{aligned}
\sum_{i=1}^{n} (X_i - \bar{X})^2 &= \sum_{i=1}^{n^*} (X_i - \bar{X})^2 + (n - n^*)\bar{X}^2 \\
&= \sum_{i=1}^{n^*} (X_i - \tilde{X})^2 + \frac{(n - n^*)n}{n^*} \bar{X}^2
\end{aligned}$$

and also

$$\sum_{i=1}^{m} (Y_i - \bar{Y})^2 = \sum_{i=1}^{m^*} (Y_i - \tilde{Y})^2 + \frac{(m - m^*)m}{m^*} \bar{Y}^2.$$

Note that $\sum_{i=1}^{n^*} (X_i - \tilde{X})^2$ is independent of $\bar{X}$, and also $\sum_{i=1}^{m^*} (Y_i - \tilde{Y})^2$ is independent of $\bar{Y}$. Now

$$\begin{aligned}
P(|t| > \mathrm{cv}) &= P\left(\frac{(\bar{X} - \bar{Y})^2}{\frac{1}{n(n-1)} \sum_{i=1}^{n} (X_i - \bar{X})^2 + \frac{1}{m(m-1)} \sum_{i=1}^{m} (Y_i - \bar{Y})^2} \geq \mathrm{cv} \right) \\
&= P\left(\frac{(\bar{X} - \bar{Y})^2 - \mathrm{cv}\, \frac{(n-n^*)}{n^*(n-1)} \bar{X}^2 - \mathrm{cv}\, \frac{(m-m^*)}{m^*(m-1)} \bar{Y}^2}{\frac{1}{n(n-1)} \sum_{i=1}^{n^*} (X_i - \tilde{X})^2 + \frac{1}{m(m-1)} \sum_{i=1}^{m^*} (Y_i - \tilde{Y})^2} \geq \mathrm{cv} \right)
\end{aligned}$$

and, with $Z^* = (Z_1, \ldots, Z_{n^*}, Z_n, \cdots, Z_{n+m^*})' \sim \mathcal{N}(0, I_{n^*+m^*})$,

$$\begin{aligned}
&(\bar{X} - \bar{Y})^2 - \mathrm{cv}\, \frac{(n - n^*)}{n^*(n-1)} \bar{X}^2 - \mathrm{cv}\, \frac{(m - m^*)}{m^*(m-1)} \bar{Y}^2 \\
= \ & Z'D' \begin{pmatrix} (\frac{1}{n^2} - \mathrm{cv}\, \frac{n-n^*}{n^* n^2 (n-1)}) e_n e_n' & -e_n e_m'/mn \\ -e_m e_n'/mn & (\frac{1}{m^2} - \mathrm{cv}\, \frac{m-m^*}{m^* m^2 (m-1)}) e_m e_m' \end{pmatrix} DZ \\
= \ & Z^{*\prime} \begin{pmatrix} (\frac{1}{n^2} - \mathrm{cv}\, \frac{n-n^*}{n^* n^2 (n-1)}) e_{n^*} e_{n^*}' & -\rho e_{n^*} e_{m^*}'/mn \\ -\rho e_{m^*} e_{n^*}'/mn & \rho^2 (\frac{1}{m^2} - \mathrm{cv}\, \frac{m-m^*}{m^* m^2 (m-1)}) e_{m^*} e_{m^*}' \end{pmatrix} Z^* \\
= \ & Z^{*\prime} \begin{pmatrix} \frac{d_n}{n^{*2}} e_{n^*} e_{n^*}' & -\rho e_{n^*} e_{m^*}'/mn \\ -\rho e_{m^*} e_{n^*}'/mn & \rho^2 \frac{d_m}{m^{*2}} e_{m^*} e_{m^*}' \end{pmatrix} Z^* \\
= \ & Z^{*\prime} H(\rho) Z^*
\end{aligned}$$

where

$$d_n = n^{*2} \left(\frac{1}{n^2} - \mathrm{cv}\, \frac{n - n^*}{n^* n^2 (n-1)} \right), \; d_m = m^{*2} \left(\frac{1}{m^2} - \mathrm{cv}\, \frac{m - m^*}{m^* m^2 (m-1)} \right).$$

$H(\rho)$ has at most two non-zero eigenvalues. If both of them are (weakly) negative, then we have nothing to prove (since then $P(|t| > \text{cv}) = 0$). The largest eigenvalue is positive if for some $v(s) = (e'_{n^*}, -se'_{m^*})'$ with $s \in \mathbb{R}$,

$$v(s)'H(\rho)v(s) > 0.$$

This evaluates to

$$v(s)'H(\rho)v(s) = d_n + s\frac{2m^*n^*}{mn}\rho + s^2 d_m \rho^2 > 0 \tag{B2}$$

so that either $d_m \geq 0$, or with

$$s^* = \frac{\frac{m^*n^*}{mn}}{-d_m}\frac{1}{\rho} = -\frac{m^*n^*}{mnd_m}\frac{1}{\rho}$$

condition (B2) becomes

$$\begin{aligned} d_n - 2\frac{(m^*n^*)^2}{(mn)^2 d_m} + \frac{(m^*n^*)^2}{(mn)^2 d_m^2} d_m &> 0 \\ sign(d_m)(d_n d_m - \frac{(m^*n^*)^2}{(mn)^2}) &> 0. \end{aligned}$$

Thus, if for a given set of $(\text{cv}, m, m^*, n, n^*)$,

$$d_m < 0 \text{ and } sign(d_m)(d_n d_m - \frac{(m^*n^*)^2}{(mn)^2}) \leq 0 \tag{B3}$$

then we have nothing to prove.

Otherwise, let $\nu(\rho)$ be the non-negative eigenvalue, and $-\eta(\rho)$ be the weakly negative eigenvalue of $H(\rho)$, so that

$$Z^{*\prime}H(\rho)Z^* \sim \nu(\rho)Z_0^2 - \eta(\rho)Z_{-1}^2$$

with Z_0, Z_{-1} independent $\mathcal{N}(0,1)$, and independent of Z and Z^*.

We first show that η and ν are (weakly) increasing in ρ. We have

$$\begin{aligned} \nu(\rho) &= \max_s \frac{v(s)'H(\rho)v(s)}{v(s)'v(s)} \\ &= \max_s \frac{d_n + s\frac{2m^*n^*}{mn}\rho + s^2 d_m \rho^2}{n^* + s^2 m^*}. \end{aligned}$$

Let $s_0^* = \arg\max_s \frac{v(s)'H(\rho_0)v(s)}{v(s)'v(s)}$. Then, for $0 \leq \rho_0 < \rho_1$, with $s_1 = s_0^*\rho_0/\rho_1$

$$\nu(\rho_0) = \max_s \frac{v(s)'H(\rho_0)v(s)}{v(s)'v(s)} = \frac{d_n + \frac{2m^*n^*}{mn}s_0^*\rho_0 + d_m s_0^{*2}\rho_0^2}{n^* + s_0^{*2}m^*}$$

$$
\begin{aligned}
&= \frac{d_n + \frac{2m^* n^*}{mn} s_1 \rho_1 + d_m s_1^2 \rho_1^2}{n^* + m^* s_1^2 \rho_1^2 / \rho_0^2} \\
&\leq \frac{d_n + \frac{2m^* n^*}{mn} s_1 \rho_1 + d_m s_1^2 \rho_1^2}{n^* + m^* s_1^2} \\
&\leq \max_s \frac{v(s)' H(\rho_1) v(s)}{v(s)' v(s)} = \nu(\rho_1)
\end{aligned}
$$

and, using the same argument,

$$
\eta(\rho_0) = \max_s \frac{v(s)'(-H(\rho_0))v(s)}{v(s)'v(s)} \leq \max_s \frac{v(s)'(-H(\rho_1))v(s)}{v(s)'v(s)} = \eta(\rho_1).
$$

Step 3: Suppose that $d_n < 0$. We can write

$$
\begin{aligned}
P(|t| > \text{cv}) &= P\left(\frac{(\bar{X} - \bar{Y})^2 - \text{cv} \frac{(n-n^*)}{n^*(n-1)} \bar{X} - \text{cv} \frac{(m-m^*)}{m^*(m-1)} \bar{Y}}{\frac{1}{n(n-1)} \sum_{i=1}^{n^*} (X_i - \tilde{X})^2 + \frac{1}{m(m-1)} \sum_{i=1}^{m^*} (Y_i - \tilde{Y})^2} \geq \text{cv} \right) \\
&= P\left(\frac{\nu(\rho) Z_0^2 - \eta(\rho) Z_{-1}^2}{\frac{1}{n(n-1)} \sum_{i=1}^{n^*} (X_i - \tilde{X})^2 + \frac{1}{m(m-1)} \rho^2 \sum_{i=1}^{m^*} (Z_i - \bar{Z})^2} \geq \text{cv} \right).
\end{aligned}
$$

Since η and ν are increasing in $0 \leq \rho \leq 1$, for any M, and $i = 1, \ldots, M$,

$$
P(|t| > \text{cv}) \leq P\left(\left(\frac{\nu(i/M) Z_0^2 - \eta((i-1)/M) Z_{-1}^2}{\frac{\text{cv}}{n(n-1)} \sum_{i=1}^{n^*} (X_i - \tilde{X})^2 + \frac{\text{cv}}{m(m-1)} (\frac{i-1}{M})^2 \sum_{i=1}^{m^*} (Z_i - \bar{Z})^2} \geq 1 \right) \right).
$$

A direct computation using Lemma 1 now shows that for $N = 100$, the right-hand side is smaller than α for all $i = 1, \ldots, M$ with sufficiently large M. (It is numerically convenient to first try smaller M, and only increase it if the inequality does not hold. $M = 10,000$ yields the required inequality in all cases, though.)

Step 4: We are left two show the result for $d_n > 0$. Let

$$
\kappa_0 = \frac{d_n}{n^*} \quad \text{and} \quad \kappa_1 = \max\left(\frac{d_m}{m^*}, \frac{m^* n^{*2}}{d_n m^2 n^2} \right).
$$

We first show that

$$
\nu(\rho) \leq \kappa_0 + \kappa_1 \rho^2
$$

uniformly in $0 \leq \rho \leq 1$. With $v(s) = (e'_{n^*}, -s e'_{m^*})'$ for $s \in \mathbb{R}$, this is equivalent to

$$
(\kappa_0 + \kappa_1 \rho^2) v(s)' v(s) - v(s)' H(\rho) v(s) \geq 0
$$

for all $s \in \mathbb{R}$, which evaluates to

$$(\kappa_0 + \kappa_1 \rho^2)(n^* + s^2 m^*) - d_n - 2s\rho \frac{m^* n^*}{mn} - s^2 \rho^2 d_m \geq 0. \tag{B4}$$

The coefficient on s^2 then is

$$\kappa_0 n^* + \rho^2 (m^* \kappa_1 - d_m). \tag{B5}$$

By definition of κ_0 and κ_1, (B5) is positive for all $0 \leq \rho \leq 1$. The smallest value of (B4) is hence taken on at

$$s = \frac{\rho \frac{m^* n^*}{mn}}{\kappa_0 m^* + \rho^2 (m^* \kappa_1 - d_m)}.$$

Plugging this into (B4) and some rearranging yields a non-negative number for all $0 \leq \rho \leq 1$, so that $\nu(\rho) \leq \kappa_0 + \kappa_1 \rho^2$.

Furthermore we will show that

$$\eta(\rho) \geq \psi_0 \frac{\rho^2}{\kappa_0 + \kappa_1 \rho^2}.$$

To this end, define

$$h^2 = \frac{m^* \kappa_1}{n^* \kappa_0}$$

and consider $v = (\rho h e_{n^*}', e_{m^*}')'$. Then

$$\begin{aligned} \eta(\rho) &\geq \frac{v'(-H(\rho))v}{v'v} \\ &= -\rho^2 \frac{d_n h^2 - 2h \frac{m^* n^*}{mn} + d_m}{\frac{m^*}{\kappa_0}(\kappa_0 + \kappa_1 \rho^2)} \\ &= \psi_0 \frac{\rho^2}{\kappa_0 + \kappa_1 \rho^2} \end{aligned}$$

with

$$\psi_0 = -\kappa_0 \frac{d_n h^2 - 2h \frac{m^* n^*}{mn} + d_m}{m^*}.$$

Now due to these inequalities, we have that

$$\begin{aligned} P(|t| > \mathrm{cv}) &= P\left(\frac{(\bar{X} - \bar{Y})^2 - \mathrm{cv}\, \frac{(n-n^*)}{n^*(n-1)} \bar{X} - \mathrm{cv}\, \frac{(m-m^*)}{m^*(m-1)} \bar{Y}}{\frac{1}{n(n-1)} \sum_{i=1}^{n^*} (X_i - \tilde{X})^2 + \frac{1}{m(m-1)} \sum_{i=1}^{m^*} (Y_i - \tilde{Y})^2} \geq \mathrm{cv} \right) \\ &= P\left(\frac{\nu(\rho) Z_0^2 - \eta(\rho) Z_{-1}^2}{\frac{\mathrm{cv}}{n(n-1)} \sum_{i=1}^{n^*} (X_i - \tilde{X})^2 + \frac{\mathrm{cv}}{m(m-1)} \sum_{i=1}^{m^*} (Y_i - \tilde{Y})^2} \geq 1 \right) \end{aligned}$$

$$
\begin{aligned}
&\le P\left(\frac{(\kappa_0+\kappa_1\rho^2)Z_0^2 - \psi_0 \frac{\rho^2}{\kappa_0+\kappa_1\rho^2} Z_{-1}^2}{\frac{\text{cv}}{n(n-1)}\sum_{i=1}^{n^*}(X_i-\tilde{X})^2 + \frac{\text{cv}}{m(m-1)}\rho^2\sum_{i=1}^{m^*}(Z_i-\bar{Z})^2} \ge 1 \right) \\
&= P\left((\kappa_0+\kappa_1\rho^2)Z_0^2 \ge \psi_0 \frac{\rho^2}{\kappa_0+\kappa_1\rho^2} Z_{-1}^2 + \frac{\text{cv}}{n(n-1)}\sum_{i=1}^{n^*}(X_i-\tilde{X})^2 + \right. \\
&\qquad\qquad \left. \frac{\text{cv}}{m(m-1)}\rho^2\sum_{i=1}^{m^*}(Z_i-\bar{Z})^2 \right)
\end{aligned}
$$

and, with

$$
\begin{aligned}
\lambda_2 &= \frac{1}{\kappa_0+\kappa_1\rho^2}\frac{\text{cv}}{n(n-1)} \\
\lambda_1 &= \frac{\rho^2}{\kappa_0+\kappa_1\rho^2}\frac{\text{cv}}{m(m-1)} \\
\lambda_0 &= \frac{\rho^2}{(\kappa_0+\kappa_1\rho^2)^2}\psi_0
\end{aligned}
$$

by Lemma 2 and the development on page 10 of B&S(05), we find

$$
P(|t|>\text{cv}) \le \frac{1}{\pi}\int_0^1 (1-s)^{-1/2}s^{-1/2}s^{(n^*+m^*-1)/2}(s+\lambda_0)^{-1/2}(s+\lambda_2)^{-(n^*-1)/2}(s+\lambda_1)^{-(m^*-1)/2}ds. \tag{B6}
$$

Now define

$$
z = \frac{\rho^2}{\kappa_0+\kappa_1\rho^2}, \quad \text{so that} \quad \frac{1}{\kappa_0+\kappa_1\rho^2} = \frac{1-\kappa_1 z}{\kappa_0}
$$

so that

$$
\begin{aligned}
\lambda_2 &= \frac{\text{cv}}{n(n-1)}\frac{1-\kappa_1 z}{\kappa_0} \\
\lambda_1 &= \frac{\text{cv}}{m(m-1)}z \\
\lambda_0 &= \frac{z-\kappa_1 z^2}{\kappa_0}\psi_0.
\end{aligned}
$$

By inspection of the second derivative, the function

$$
z \mapsto (s+\frac{z-\kappa_1 z^2}{\kappa_0}\psi_0)^{-1/2}(s+\frac{\text{cv}}{m(m-1)}z)^{-(n^*-1)/2}(s+\frac{\text{cv}}{n(n-1)}\frac{1-\kappa_1 z}{\kappa_0})^{-(m^*-1)/2}
$$

is seen to be log-convex for all $s \in [0,1]$, so that the integral on the right-hand side of (B6) is also a convex function of z. Since $dz/d\rho > 0$ for all $0 < \rho \le 1$, the extreme values for

$P(|t|>\mathrm{cv})$ thus arise with $\rho=0$ or $\rho=1$. For $\rho=0$, we obtain

$$\begin{aligned}
P(|t| &> \mathrm{cv}) \leq \left(\frac{\kappa_0 Z_0^2}{\frac{\mathrm{cv}}{n(n-1)}\sum_{i=1}^{n^*}(X_i-\tilde{X})^2}\geq 1\right) \\
&= P\left(\frac{\frac{(\mathrm{cv}+n-1)n^*-\mathrm{cv}\,n}{n(n^*-1)n^*}Z_0^2}{\frac{1}{n^*(n^*-1)}\sum_{i=1}^{n^*}(X_i-\tilde{X})^2}\geq \mathrm{cv}\right) \\
&\leq P\left(\frac{Z_0^2}{\frac{1}{n^*(n^*-1)}\sum_{i=1}^{n^*}(X_i-\tilde{X})^2}\geq \mathrm{cv}\right)
\end{aligned}$$

and the inequality of the Theorem follows from B&S(05). For $\rho=1$, we obtain

$$P(|t|>\mathrm{cv})\leq P\left(\frac{(\kappa_0+\kappa_1)Z_0^2}{\psi_0\frac{1}{\kappa_0+\kappa_1}Z_{-1}^2+\frac{\mathrm{cv}}{n(n-1)}\sum_{i=1}^{n^*}(X_i-\tilde{X})^2+\frac{\mathrm{cv}}{m(m-1)}\sum_{i=1}^{m^*}(Z_i-\bar{Z})^2}\geq 1\right)$$

and a direct calculation using Lemma 1 with $N=100$ shows that the inequality of the Theorem holds.