Supplementary materials for *Small sample methods for cluster-robust variance estimation and hypothesis testing in fixed effects models*

November 10, 2022

Contents

S1 Proof of Theorem 1

Consider the matrix $\mathbf{B}_i$ as defined in Equation 8 of the main text:

$$\mathbf{B}_i = \mathbf{D}_i \left(\mathbf{I} - \mathbf{H_X}\right)_i \boldsymbol{\Phi} \left(\mathbf{I} - \mathbf{H_X}\right)_i' \mathbf{D}_i',$$

The Moore-Penrose inverse of $\mathbf{B}_i$ can be computed from its eigen-decomposition. Let $b \leq n_i$ denote the rank of $\mathbf{B}_i$. Let $\boldsymbol{\Lambda}$ be the $b \times b$ diagonal matrix of the positive eigenvalues of $\mathbf{B}_i$ and $\mathbf{V}$ be the $n_i \times b$ matrix of corresponding eigen-vectors, so that $\mathbf{B}_i = \mathbf{V}\boldsymbol{\Lambda}\mathbf{V}'$. Then $\mathbf{B}_i^+ = \mathbf{V}\boldsymbol{\Lambda}^{-1}\mathbf{V}'$ and $\mathbf{B}_i^{+1/2} = \mathbf{V}\boldsymbol{\Lambda}^{-1/2}\mathbf{V}'$. Because the adjustment matrices taken to be $\mathbf{A}_i = \mathbf{D}_i'\mathbf{B}_i^{+1/2}\mathbf{D}_i$, we have that

$$
\begin{aligned}
\ddot{\mathbf{R}}_i'\mathbf{W}_i\mathbf{A}_i \left(\mathbf{I} - \mathbf{H_X}\right)_i \boldsymbol{\Phi} \left(\mathbf{I} - \mathbf{H_X}\right)_i' \mathbf{A}_i'\mathbf{W}_i\ddot{\mathbf{R}}_i &= \ddot{\mathbf{R}}_i'\mathbf{W}_i\mathbf{D}_i'\mathbf{B}_i^{+1/2}\mathbf{B}_i\mathbf{B}_i^{+1/2}\mathbf{D}_i\mathbf{W}_i\ddot{\mathbf{R}}_i \\
&= \ddot{\mathbf{R}}_i'\mathbf{W}_i\mathbf{D}_i'\mathbf{V}\mathbf{V}'\mathbf{D}_i\mathbf{W}_i\ddot{\mathbf{R}}_i.
\end{aligned}
\tag{S1}
$$

Thus, it will suffice to show that $\mathbf{V}'\mathbf{D}_i\mathbf{W}_i\ddot{\mathbf{R}}_i = \mathbf{D}_i\mathbf{W}_i\ddot{\mathbf{R}}_i$.

Now, because $\mathbf{D}_i$ and $\boldsymbol{\Phi}$ are positive definite and $\mathbf{B}_i$ is symmetric, the eigen-vectors $\mathbf{V}$ define an orthonormal basis for the column span of $\left(\mathbf{I} - \mathbf{H_X}\right)_i$. We now show that $\ddot{\mathbf{U}}_i$ is in the column space of $\left(\mathbf{I} - \mathbf{H_X}\right)_i$. Let $\mathbf{Z}_i$ be an $n_i \times (r+s)$ matrix of zeros. With $\mathbf{L}_i$ as defined in Theorem 1, take $\mathbf{Z}_k = -\ddot{\mathbf{U}}_k\mathbf{L}_i^{-1}\mathbf{M}_{\ddot{\mathbf{U}}}^{-1}$, for $k \neq i$ and $\mathbf{Z} = \left(\mathbf{Z}_1', ..., \mathbf{Z}_m'\right)'$. Observe that $\left(\mathbf{I} - \mathbf{H_T}\right)\mathbf{Z} = \mathbf{Z}$. It follows that

$$
\begin{aligned}
\left(\mathbf{I} - \mathbf{H_X}\right)_i \mathbf{Z} &= \left(\mathbf{I} - \mathbf{H_{\ddot{U}}}\right)_i \left(\mathbf{I} - \mathbf{H_T}\right)\mathbf{Z} \\
&= \left(\mathbf{I} - \mathbf{H_{\ddot{U}}}\right)_i \mathbf{Z} \\
&= \mathbf{Z}_i - \ddot{\mathbf{U}}_i\mathbf{M}_{\ddot{\mathbf{U}}} \sum_{k=1}^m \ddot{\mathbf{U}}_k'\mathbf{W}_k\mathbf{Z}_k \\
&= \ddot{\mathbf{U}}_i\mathbf{M}_{\ddot{\mathbf{U}}} \left(\sum_{k\neq i} \ddot{\mathbf{U}}_k'\mathbf{W}_k\ddot{\mathbf{U}}\right) \mathbf{L}_i^{-1}\mathbf{M}_{\ddot{\mathbf{U}}}^{-1} \\
&= \ddot{\mathbf{U}}_i\mathbf{M}_{\ddot{\mathbf{U}}}\mathbf{L}_i\mathbf{L}_i^{-1}\mathbf{M}_{\ddot{\mathbf{U}}}^{-1} \\
&= \ddot{\mathbf{U}}_i.
\end{aligned}
$$

Thus, there exists an $N \times (r+s)$ matrix $\mathbf{Z}$ such that $\left(\mathbf{I} - \mathbf{H_{\ddot{X}}}\right)_i \mathbf{Z} = \ddot{\mathbf{U}}_i$, i.e., $\ddot{\mathbf{U}}_i$ is in the column span of $\left(\mathbf{I} - \mathbf{H_X}\right)_i$. Because $\mathbf{D}_i\mathbf{W}_i$ is positive definite and $\ddot{\mathbf{R}}_i$ is a sub-matrix of $\ddot{\mathbf{U}}_i$, $\mathbf{D}_i\mathbf{W}_i\ddot{\mathbf{R}}_i$ is also in the column span of $\left(\mathbf{I} - \mathbf{H_X}\right)_i$. It follows that

$$\ddot{\mathbf{R}}_i'\mathbf{W}_i\mathbf{D}_i'\mathbf{V}\mathbf{V}'\mathbf{D}_i\mathbf{W}_i\ddot{\mathbf{R}}_i = \ddot{\mathbf{R}}_i'\mathbf{W}_i\boldsymbol{\Phi}_i\mathbf{W}_i\ddot{\mathbf{R}}_i. \tag{S2}$$

Substituting (S2) into (S1) demonstrates that $\mathbf{A}_i$ satisfies the generalized BRL criterion (Equation 6 of the main text).

Under the working model, the residuals from cluster i have mean $\mathbf{0}$ and variance

$$\text{Var}\left(\ddot{\mathbf{e}}_i\right) = \left(\mathbf{I} - \mathbf{H_X}\right)_i \boldsymbol{\Phi} \left(\mathbf{I} - \mathbf{H_X}\right)_i',$$

It follows that

$$
\begin{aligned}
\text{E}\left(\mathbf{V}^{CR2}\right) &= \mathbf{M}_{\ddot{\mathbf{R}}} \left[\sum_{i=1}^m \ddot{\mathbf{R}}_i'\mathbf{W}_i\mathbf{A}_i \left(\mathbf{I} - \mathbf{H_X}\right)_i \boldsymbol{\Phi} \left(\mathbf{I} - \mathbf{H_X}\right)_i' \mathbf{A}_i\mathbf{W}_i\ddot{\mathbf{R}}_i\right] \mathbf{M}_{\ddot{\mathbf{R}}} \\
&= \mathbf{M}_{\ddot{\mathbf{R}}} \left[\sum_{i=1}^m \ddot{\mathbf{R}}_i'\mathbf{W}_i\boldsymbol{\Phi}_i\mathbf{W}_i\ddot{\mathbf{R}}_i\right] \mathbf{M}_{\ddot{\mathbf{R}}} \\
&= \text{Var}\left(\hat{\boldsymbol{\beta}}\right)
\end{aligned}
$$

S2   Proof of Theorem 2

If $\mathbf{W}_i = \mathbf{\Phi}_i = \mathbf{I}_i$, then we can write $\mathbf{B}_i$ from Equation 8 of the main text as

$$
\begin{aligned}
\mathbf{B}_i &= \mathbf{D}_i \left(\mathbf{I} - \mathbf{H}_{\ddot{\mathbf{U}}}\right)_i \left(\mathbf{I} - \mathbf{H}_{\mathbf{T}}\right) \mathbf{\Phi} \left(\mathbf{I} - \mathbf{H}_{\mathbf{T}}\right)' \left(\mathbf{I} - \mathbf{H}_{\ddot{\mathbf{U}}}\right)'_i \mathbf{D}'_i \\
&= \left(\mathbf{I} - \mathbf{H}_{\ddot{\mathbf{U}}} - \mathbf{H}_{\mathbf{T}}\right)_i \left(\mathbf{I} - \mathbf{H}_{\ddot{\mathbf{U}}} - \mathbf{H}_{\mathbf{T}}\right)'_i \\
&= \left(\mathbf{I}_i - \ddot{\mathbf{U}}_i \mathbf{M}_{\ddot{\mathbf{U}}} \ddot{\mathbf{U}}'_i - \mathbf{T}_i \mathbf{M}_{\mathbf{T}} \mathbf{T}'_i\right),
\end{aligned}
\tag{S3}
$$

where the last equality follows from the fact that $\ddot{\mathbf{U}}'_i \mathbf{T}_i = \mathbf{0}$ for $i = 1, ..., m$. Similarly, we can write

$$
\tilde{\mathbf{B}}_i = \left(\mathbf{I}_i - \ddot{\mathbf{U}}_i \mathbf{M}_{\ddot{\mathbf{U}}} \ddot{\mathbf{U}}'_i\right).
\tag{S4}
$$

We now show that $\tilde{\mathbf{A}}_i \mathbf{T}_i = \mathbf{T}_i$. Denote the rank of $\ddot{\mathbf{U}}_i$ as $u_i \leq \min\{n_i, r+s\}$ and take the thin QR decomposition of $\ddot{\mathbf{U}}_i$ as $\ddot{\mathbf{U}}_i = \mathbf{Q}_i \mathbf{R}_i$, where $\mathbf{Q}_i$ is an $n_i \times u_i$ semi-orthonormal matrix and $\mathbf{R}_i$ is a $u_i \times r+s$ matrix of rank u_i, with $\mathbf{Q}'_i \mathbf{Q}_i = \mathbf{I}$. Note that $\mathbf{Q}'_i \mathbf{T}_i = \mathbf{0}$. From the observation that $\tilde{\mathbf{B}}_i$ can be written as

$$
\tilde{\mathbf{B}}_i = \mathbf{I}_i - \mathbf{Q}_i \mathbf{Q}'_i + \mathbf{Q}_i \left(\mathbf{I} - \mathbf{R}_i \mathbf{M}_{\ddot{\mathbf{U}}} \mathbf{R}'_i\right) \mathbf{Q}'_i,
$$

it can be seen that

$$
\tilde{\mathbf{A}}_i = \tilde{\mathbf{B}}_i^{+1/2} = \mathbf{I}_i - \mathbf{Q}_i \mathbf{Q}'_i + \mathbf{Q}_i \left(\mathbf{I} - \mathbf{R}_i \mathbf{M}_{\ddot{\mathbf{U}}} \mathbf{R}'_i\right)^{+1/2} \mathbf{Q}'_i.
\tag{S5}
$$

It follows that $\tilde{\mathbf{A}}_i \mathbf{T}_i = \mathbf{T}_i$.

Setting

$$
\mathbf{A}_i = \tilde{\mathbf{A}}_i - \mathbf{T}_i \mathbf{M}_{\mathbf{T}} \mathbf{T}'_i,
\tag{S6}
$$

observe that

$$
\begin{aligned}
\mathbf{B}_i \mathbf{A}_i \mathbf{B}_i \mathbf{A}_i &= \left(\tilde{\mathbf{B}}_i - \mathbf{T}_i \mathbf{M}_{\mathbf{T}} \mathbf{T}'_i\right) \left(\tilde{\mathbf{A}}_i - \mathbf{T}_i \mathbf{M}_{\mathbf{T}} \mathbf{T}'_i\right) \left(\tilde{\mathbf{B}}_i - \mathbf{T}_i \mathbf{M}_{\mathbf{T}} \mathbf{T}'_i\right) \left(\tilde{\mathbf{A}}_i - \mathbf{T}_i \mathbf{M}_{\mathbf{T}} \mathbf{T}'_i\right) \\
&= \left(\tilde{\mathbf{B}}_i \tilde{\mathbf{A}}_i - \mathbf{T}_i \mathbf{M}_{\mathbf{T}} \mathbf{T}'_i\right) \left(\tilde{\mathbf{B}}_i \tilde{\mathbf{A}}_i - \mathbf{T}_i \mathbf{M}_{\mathbf{T}} \mathbf{T}'_i\right) \\
&= \left(\tilde{\mathbf{B}}_i \tilde{\mathbf{A}}_i \tilde{\mathbf{B}}_i \tilde{\mathbf{A}}_i - \mathbf{T}_i \mathbf{M}_{\mathbf{T}} \mathbf{T}'_i\right) \\
&= \left(\tilde{\mathbf{B}}_i - \mathbf{T}_i \mathbf{M}_{\mathbf{T}} \mathbf{T}'_i\right) \\
&= \mathbf{B}_i.
\end{aligned}
$$

It follows that $\mathbf{A}_i$ is the symmetric square root of the Moore-Penrose inverse of $\mathbf{B}_i$, i.e., $\mathbf{A}_i = \mathbf{B}_i^{+1/2}$. Because $\mathbf{T}'_i \ddot{\mathbf{R}}_i = \mathbf{0}$, it can be seen that $\mathbf{A}_i \ddot{\mathbf{R}}_i = \left(\tilde{\mathbf{A}}_i - \mathbf{T}_i \mathbf{M}_{\mathbf{T}} \mathbf{T}'_i\right) \ddot{\mathbf{R}}_i = \tilde{\mathbf{A}}_i \ddot{\mathbf{R}}_i$. Finally, equality of $\tilde{\mathbf{V}}^{CR}$ and $\mathbf{V}^{CR}$ follows by direct evaluation of Equation 4 from the main text.

S3 Details of simulation study

We provide further details regarding the design of the simulations reported in Section 4 of the main text. Table S1 summarizes the factors manipulated in the simulation.

Table S1: Simulation design parameters

Parameter	levels
Design	RB, CR, DD
Balance	Balanced, Unbalanced
Outcome missingness	Complete data, 15% missing
Clusters (m)	15, 30, 50
Units per cluster (n)	12, 18, 30
Intra-class correlation (τ^2)	.05, .15, .25
Treatment effect variability (σ_δ^2)	.00, .04, .09
Correlation among outcomes (ρ)	.2, .8

The simulations examined six distinct study designs. Outcomes are measured for n units (which may be individuals, as in a cluster-randomized or block-randomized design, or time-points, as in a difference-in-differences panel) in each of m clusters under one of three treatment conditions. Suppose that there are G sets of clusters, each of size m_g, where the clusters in each set have a distinct configuration of treatment assignments. Let n_{ghi} denote the number of units at which cluster i in configuration g is observed under condition h, for $i = 1, ..., m$, $g = 1, ..., G$, and $h = 1, 2, 3$. Table S2 summarizes the cluster-level sample sizes and unit-level patterns of treatment allocation for each of the six designs. The simulated designs included the following:

1. A balanced, block-randomized design, with an un-equal allocation within each block. In the balanced design, the treatment allocation is identical for each block, so $G = 1$.

2. An unbalanced, block-randomized design, with two different patterns of treatment allocation ($G = 2$).

3. A balanced, cluster-randomized design, in which units are nested within clusters and an equal number of clusters are assigned to each treatment condition.

4. An unbalanced, cluster-randomized design, in which units are nested within clusters but the number of clusters assigned to each condition is not equal.

5. A balanced difference-in-differences design with two patterns of treatment allocation ($G = 2$), in which half of the clusters are observed under the first treatment condition only and the remaining half are observed under all three conditions.

6. An unbalanced difference-in-differences design, again with two patterns of treatment allocation ($G = 2$), but where 2/3 of the clusters are observed under the first treatment condition only and the remaining 1/3 of clusters are observed under all three conditions.

Table S2: Study designs used for simulation

Study design	Balance	Configuration	Clusters	Treatment allocation
Randomized Block	Balanced	1	$m_1 = m$	$n_{11i} = n/2, n_{12i} = n/3, n_{13i} = n/6$
Randomized Block	Unbalanced	1	$m_1 = m/2$	$n_{11i} = n/2, n_{12i} = n/3, n_{13i} = n/6$
		2	$m_2 = m/2$	$n_{21i} = n/3, n_{22i} = 5n/9, n_{23i} = n/9$
Cluster-Randomized	Balanced	1	$m_1 = m/3$	$n_{11i} = n$
		2	$m_2 = m/3$	$n_{22i} = n$
		3	$m_3 = m/3$	$n_{33i} = n$
Cluster-Randomized	Unbalanced	1	$m_1 = m/2$	$n_{11i} = n$
		2	$m_2 = 3m/10$	$n_{22i} = n$
		3	$m_3 = m/5$	$n_{33i} = n$
Difference-in-Differences	Balanced	1	$m_1 = m/2$	$n_{11i} = n$
		2	$m_2 = m/2$	$n_{21i} = n/2, n_{22i} = n/3, n_{23i} = n/6$
Difference-in-Differences	Unbalanced	1	$m_1 = 2m/3$	$n_{11i} = n$
		2	$m_2 = m/3$	$n_{21i} = n/2, n_{22i} = n/3, n_{23i} = n/6$

S4   Additional simulation results

S4.1   Rejection rates of AHT and standard tests
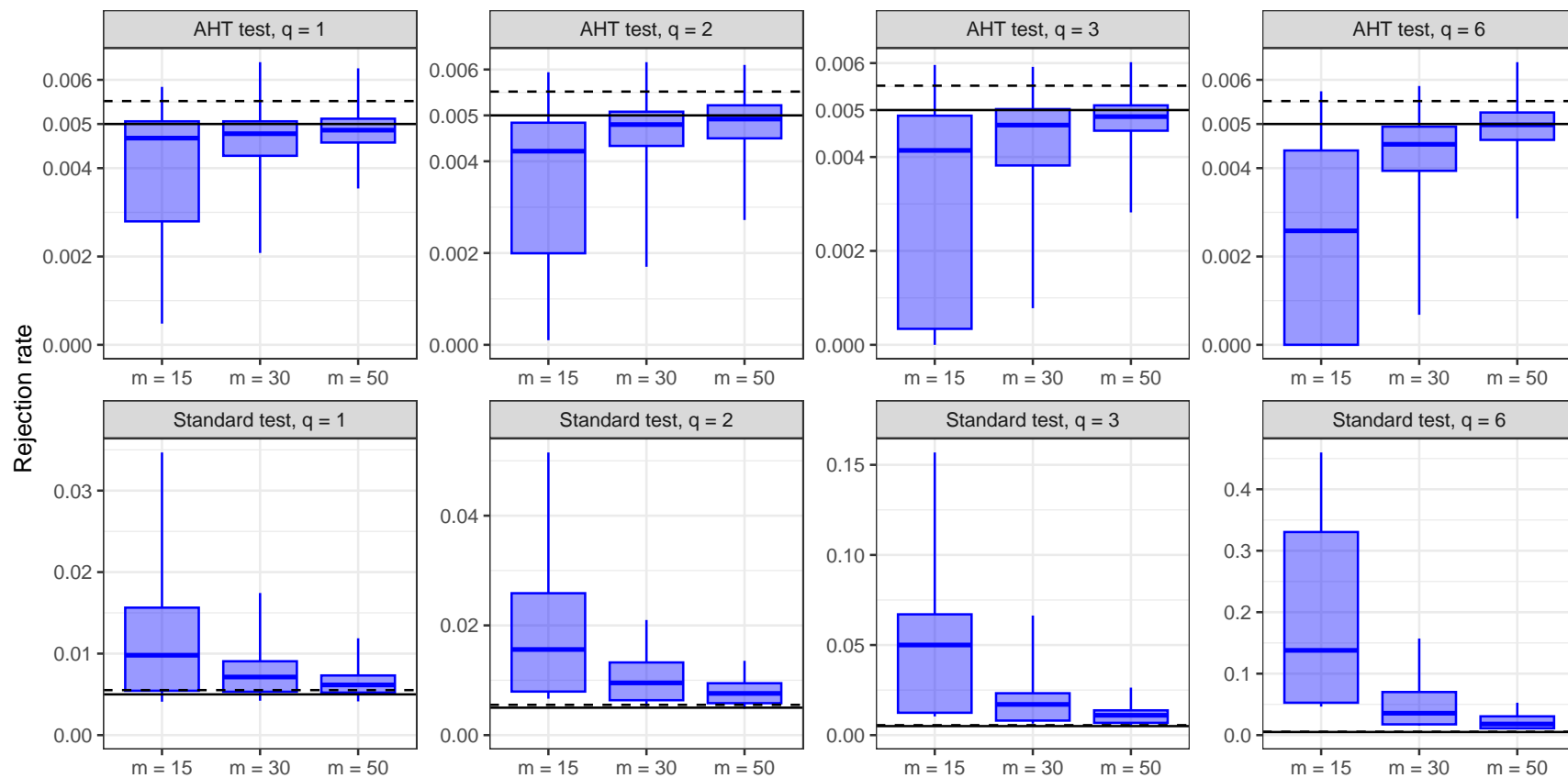
Figure S1: Rejection rates of AHT and standard tests for $\alpha = .005$, by dimension of hypothesis (q) and sample size (m).
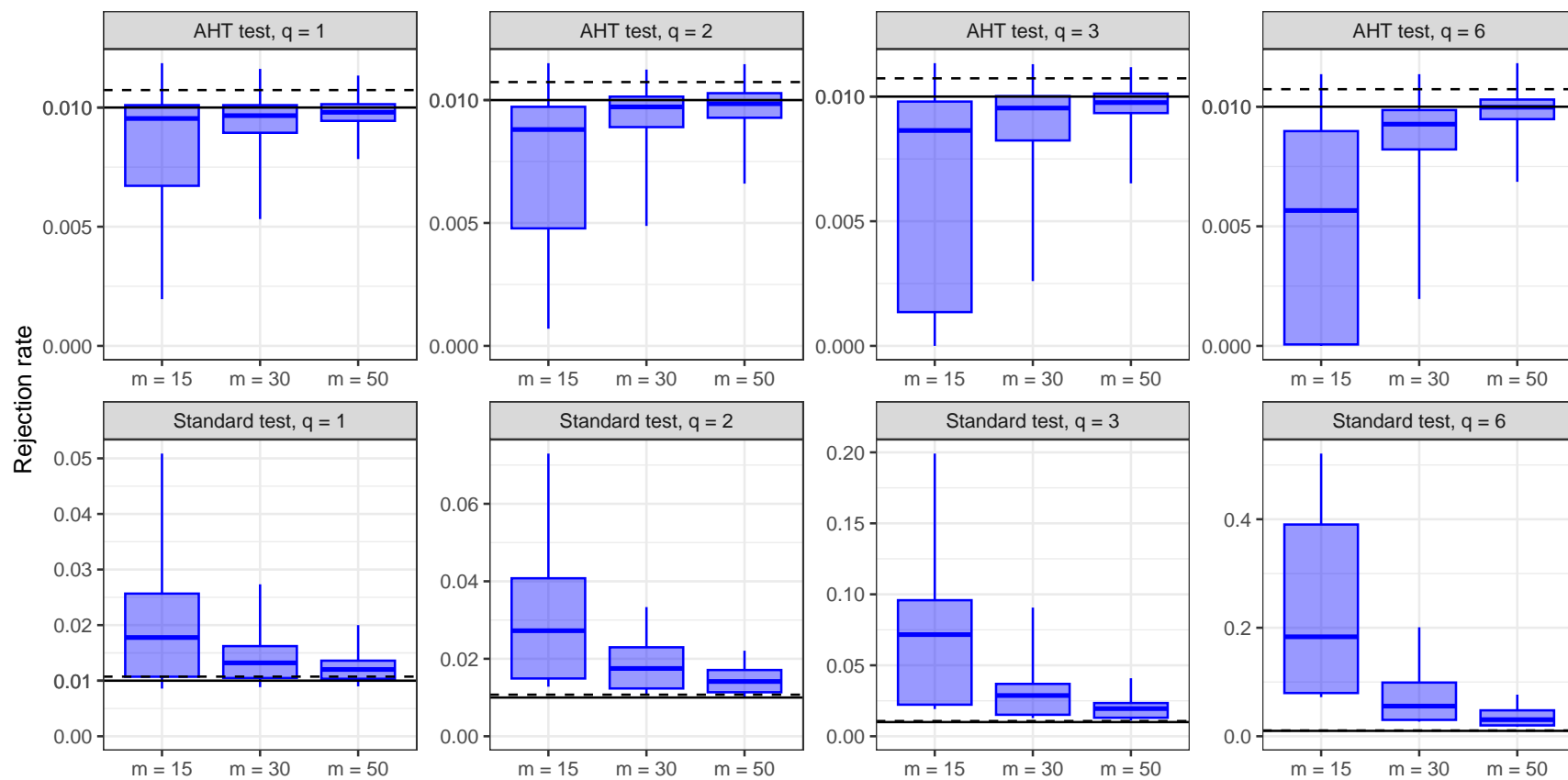
Figure S2: Rejection rates of AHT and standard tests for $\alpha = .01$, by dimension of hypothesis (q) and sample size (m).
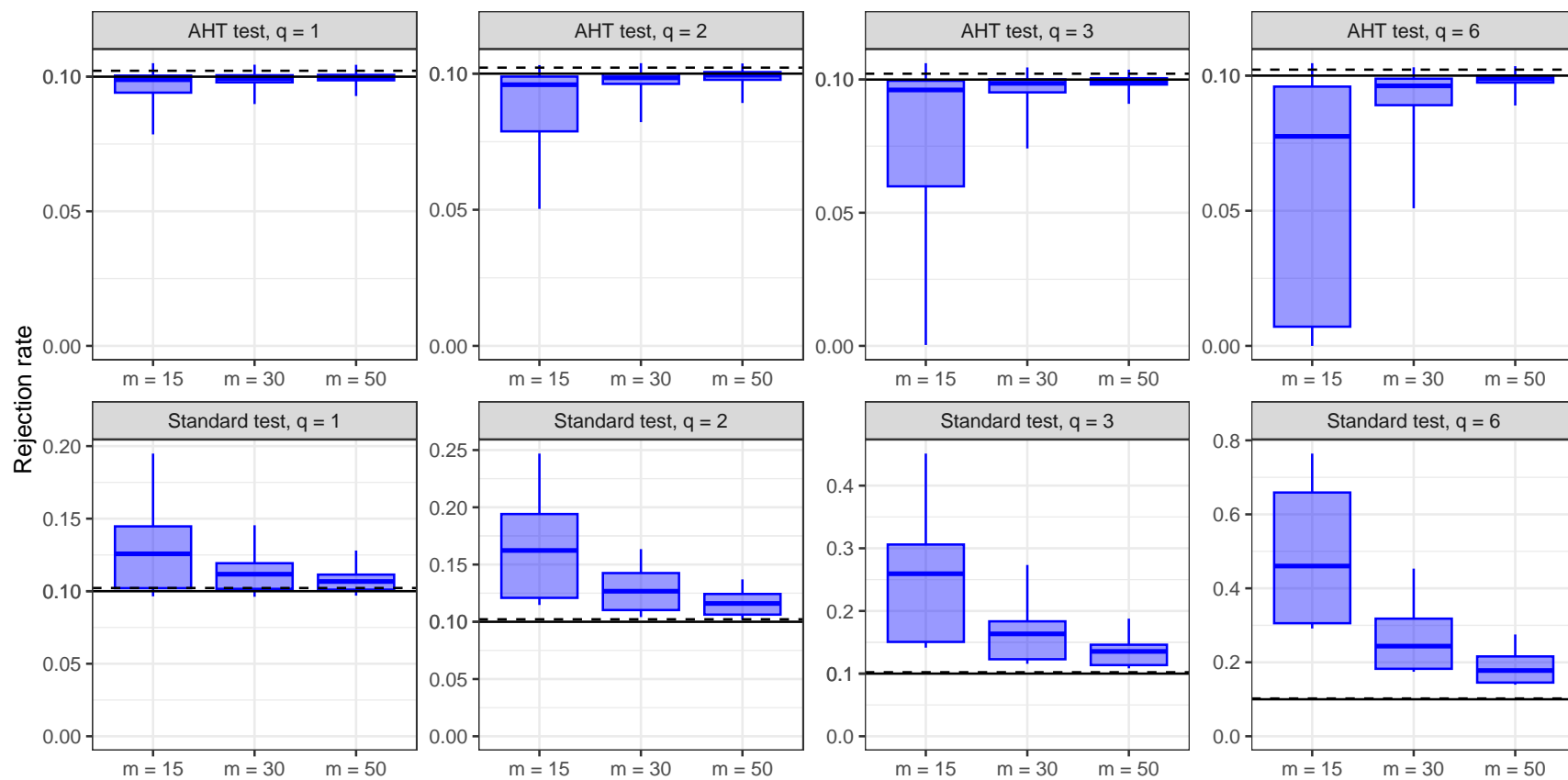
Figure S3: Rejection rates of AHT and standard tests for $\alpha = .10$, by dimension of hypothesis (q) and sample size (m).

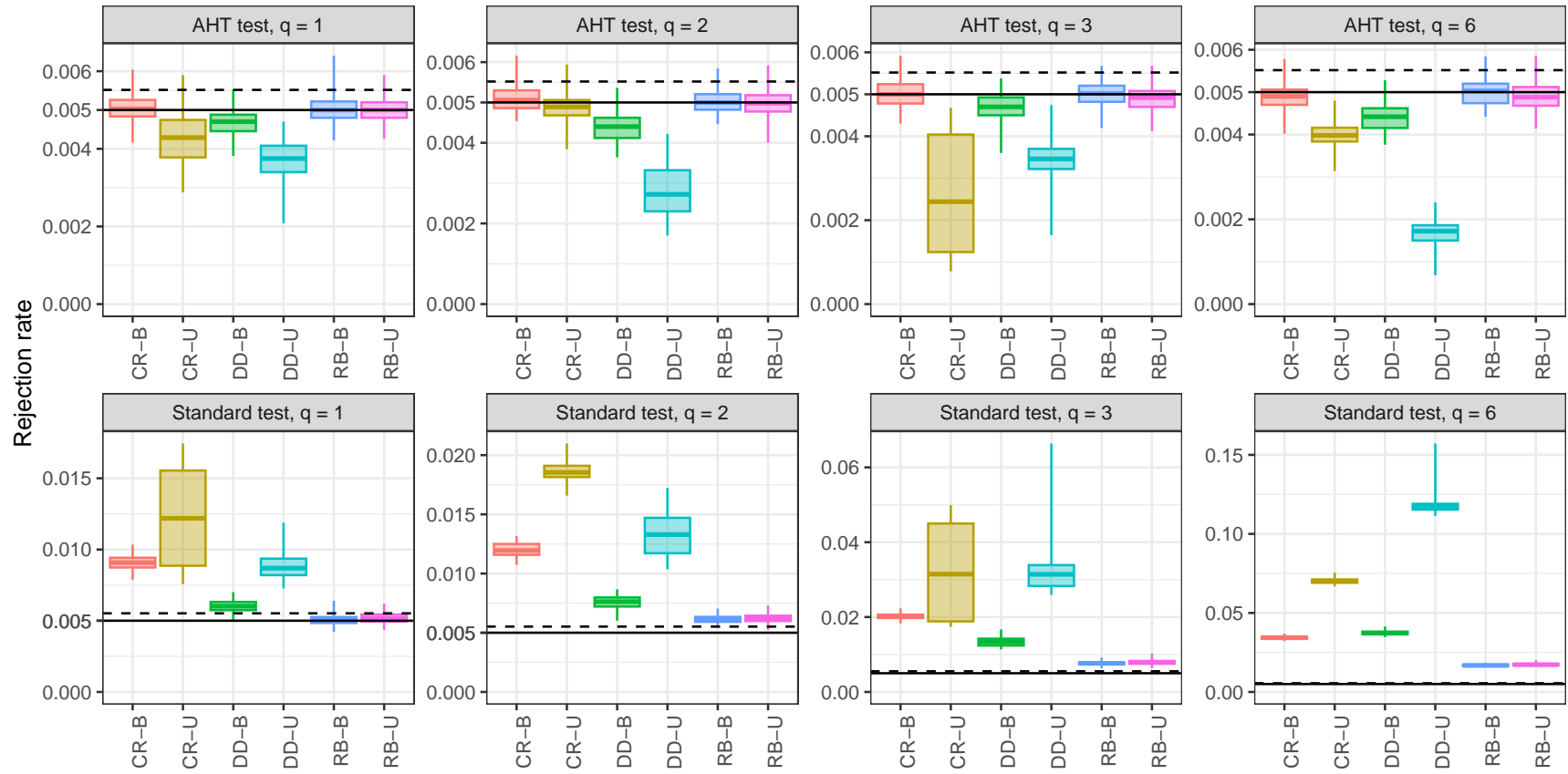## S4.2 Rejection rates of AHT and standard tests by study design
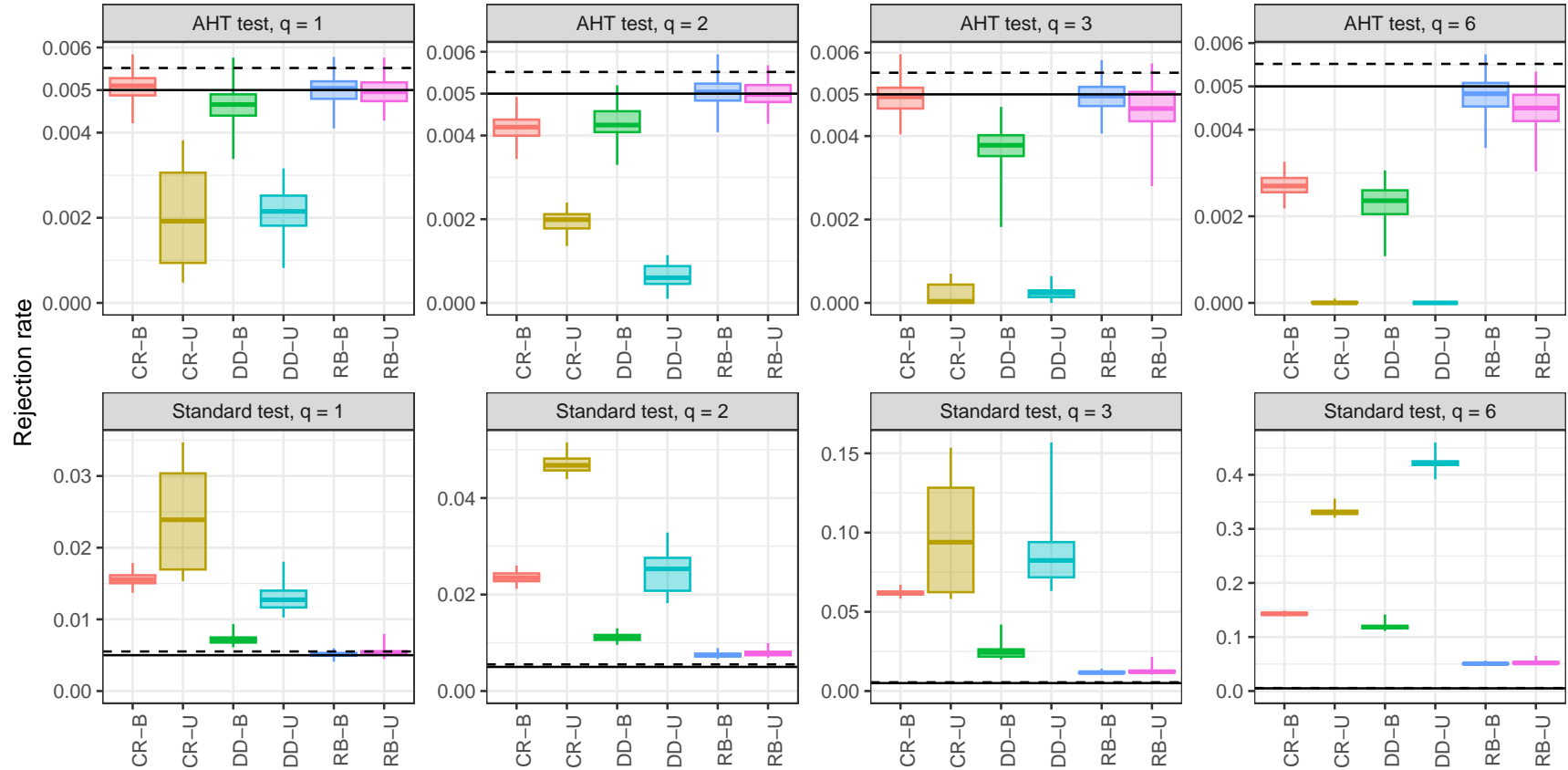
Figure S4: Rejection rates of AHT and standard tests, by study design and dimension of hypothesis (q) for $\alpha = .005$ and $m = 15$. CR = cluster-randomized design; DD = difference-in-differences design; RB = randomized block design; B = balanced; U = unbalanced.
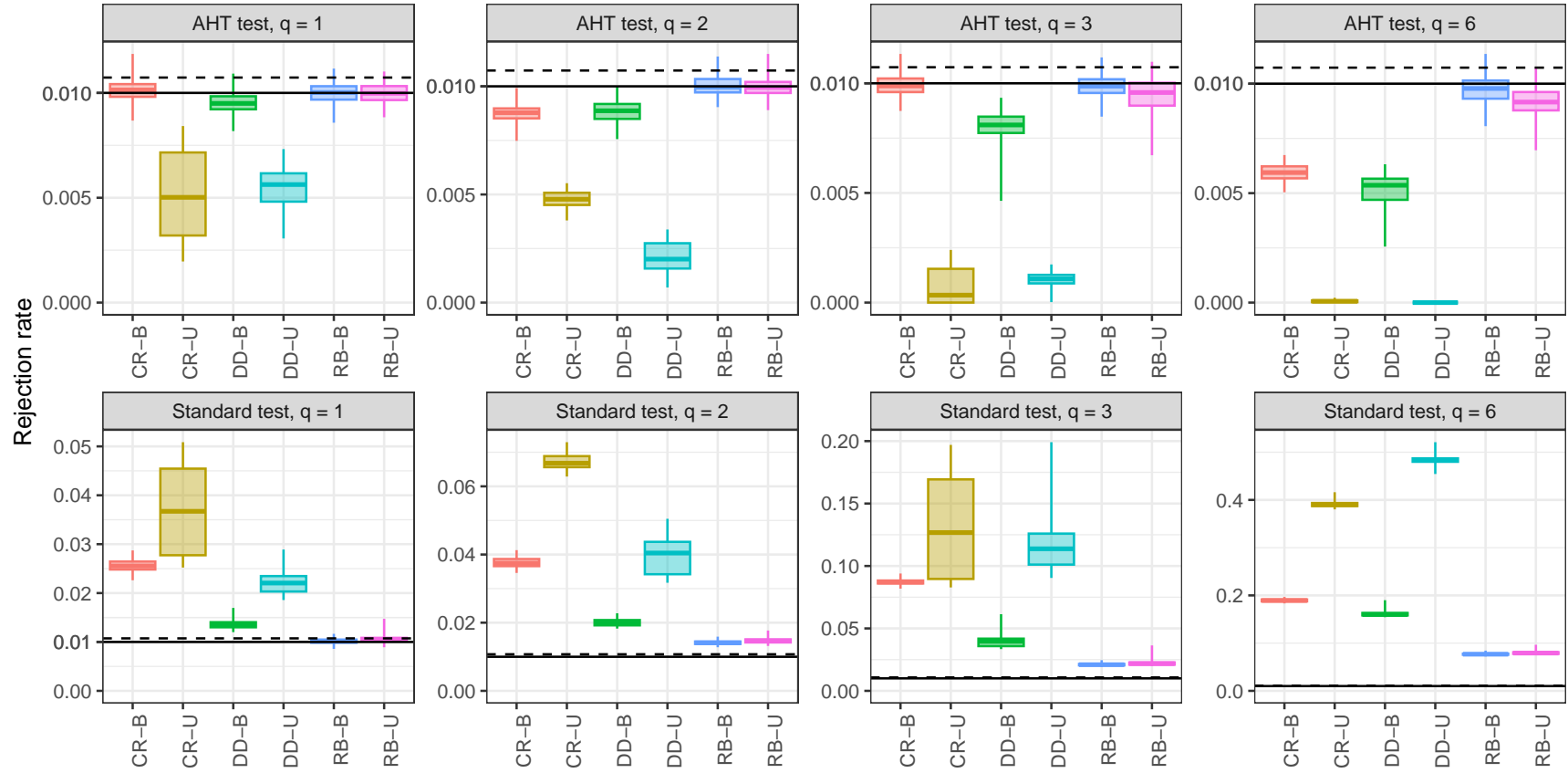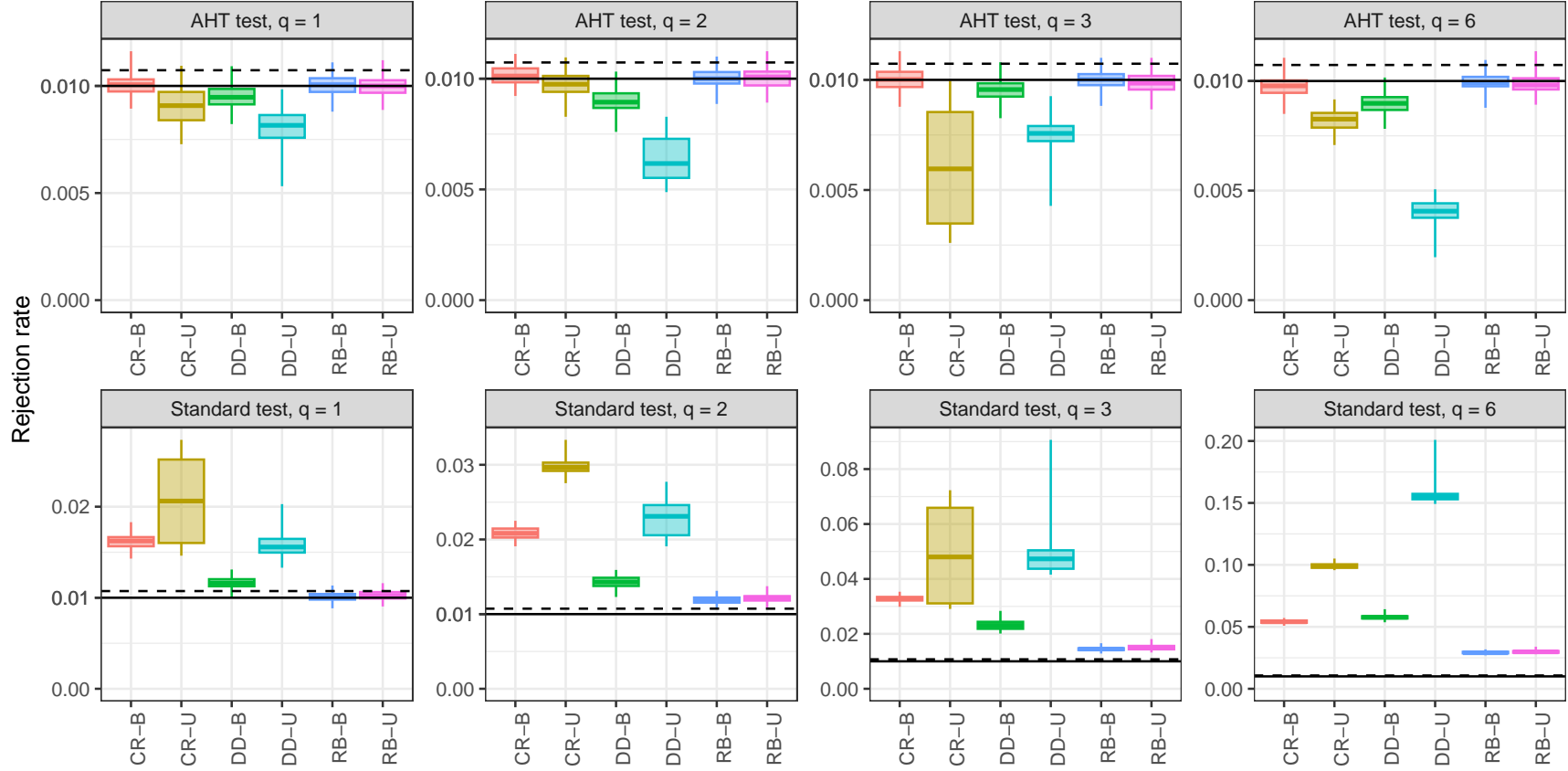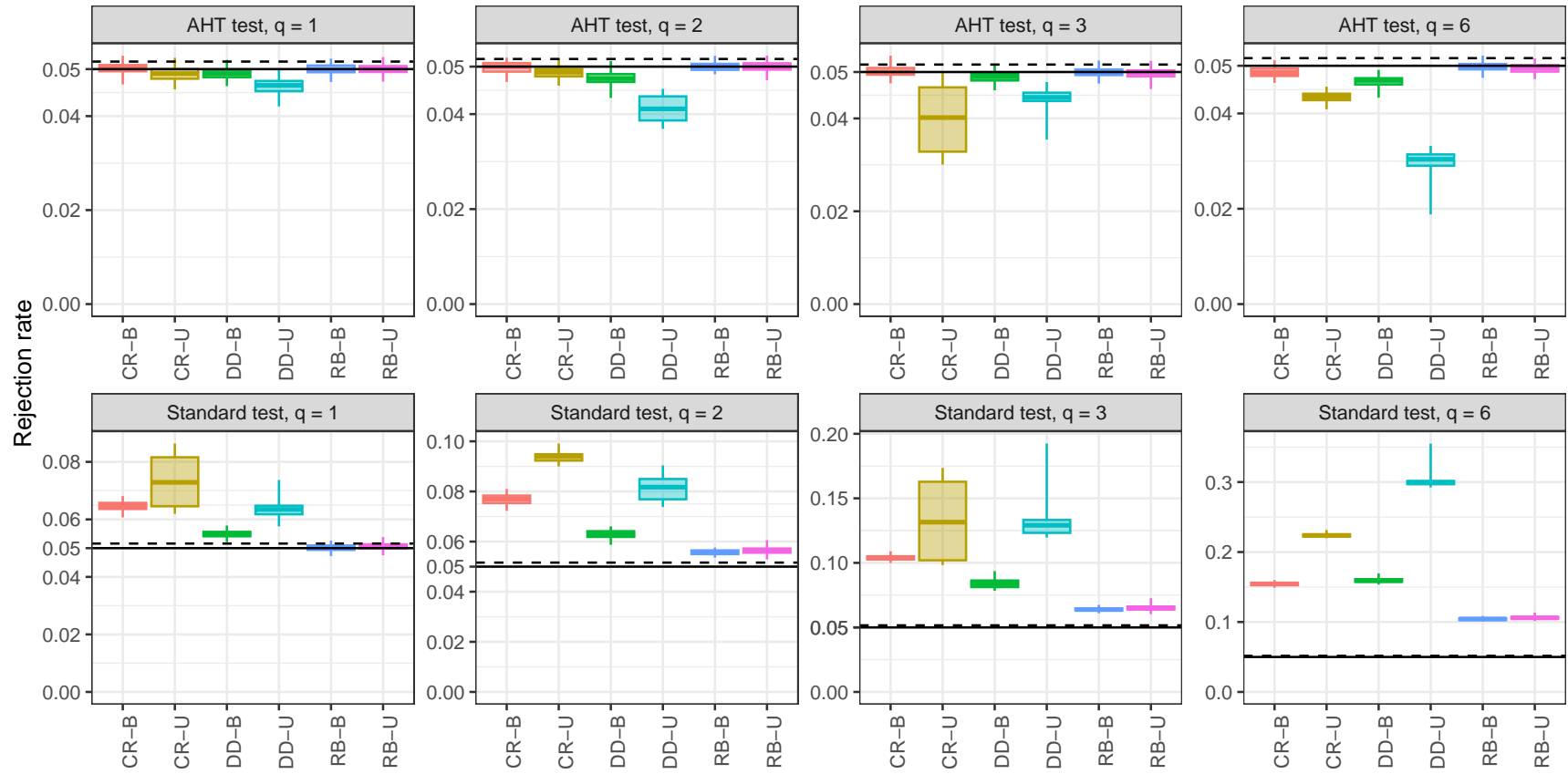
Figure S5: Rejection rates of AHT and standard tests, by study design and dimension of hypothesis (q) for $\alpha = .01$ and $m = 15$. CR = cluster-randomized design; DD = difference-in-differences design; RB = randomized block design; B = balanced; U = unbalanced.
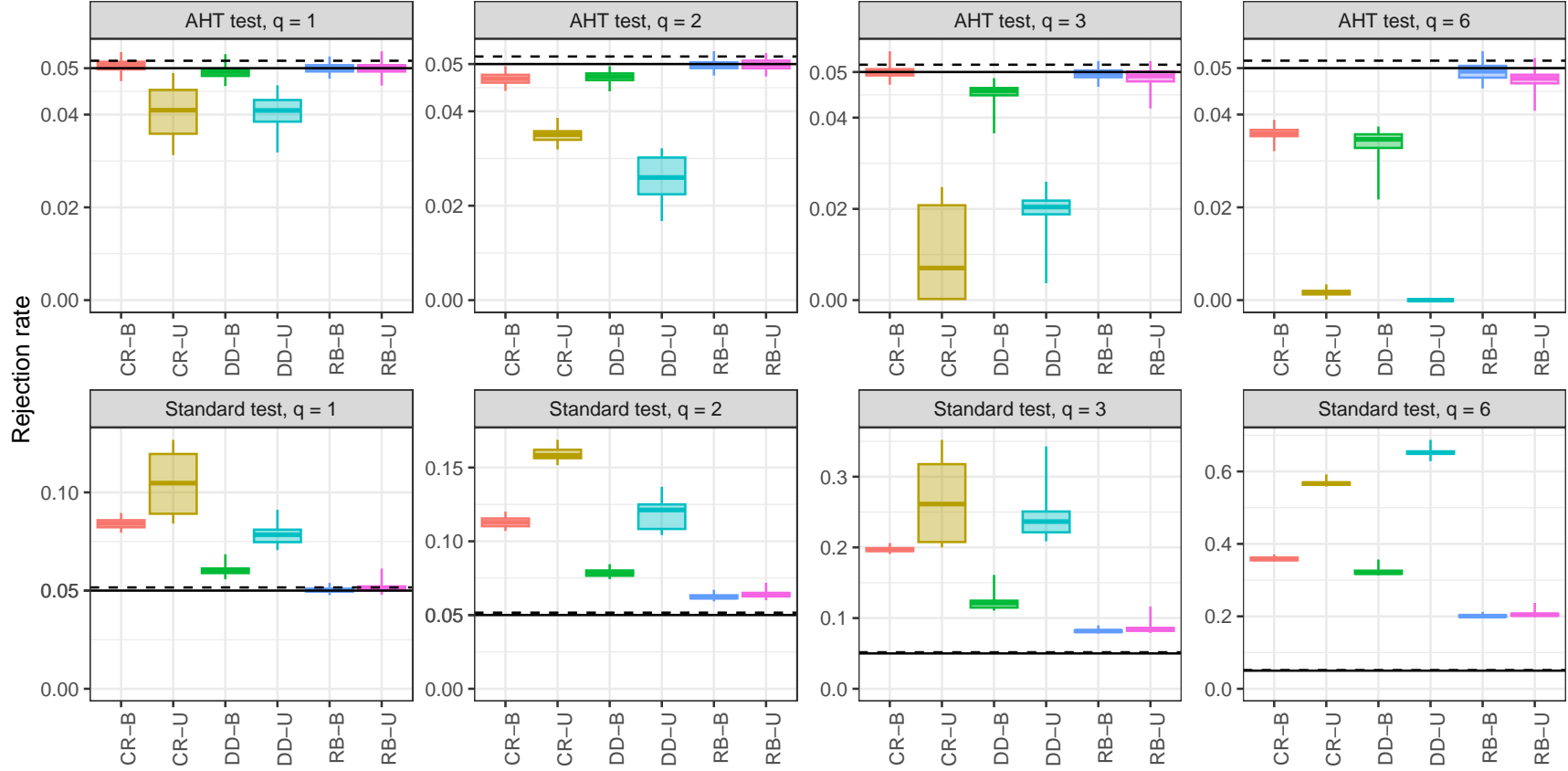
Figure S6: Rejection rates of AHT and standard tests, by study design and dimension of hypothesis (q) for $\alpha = .05$ and $m = 15$. CR = cluster-randomized design; DD = difference-in-differences design; RB = randomized block design; B = balanced; U = unbalanced.
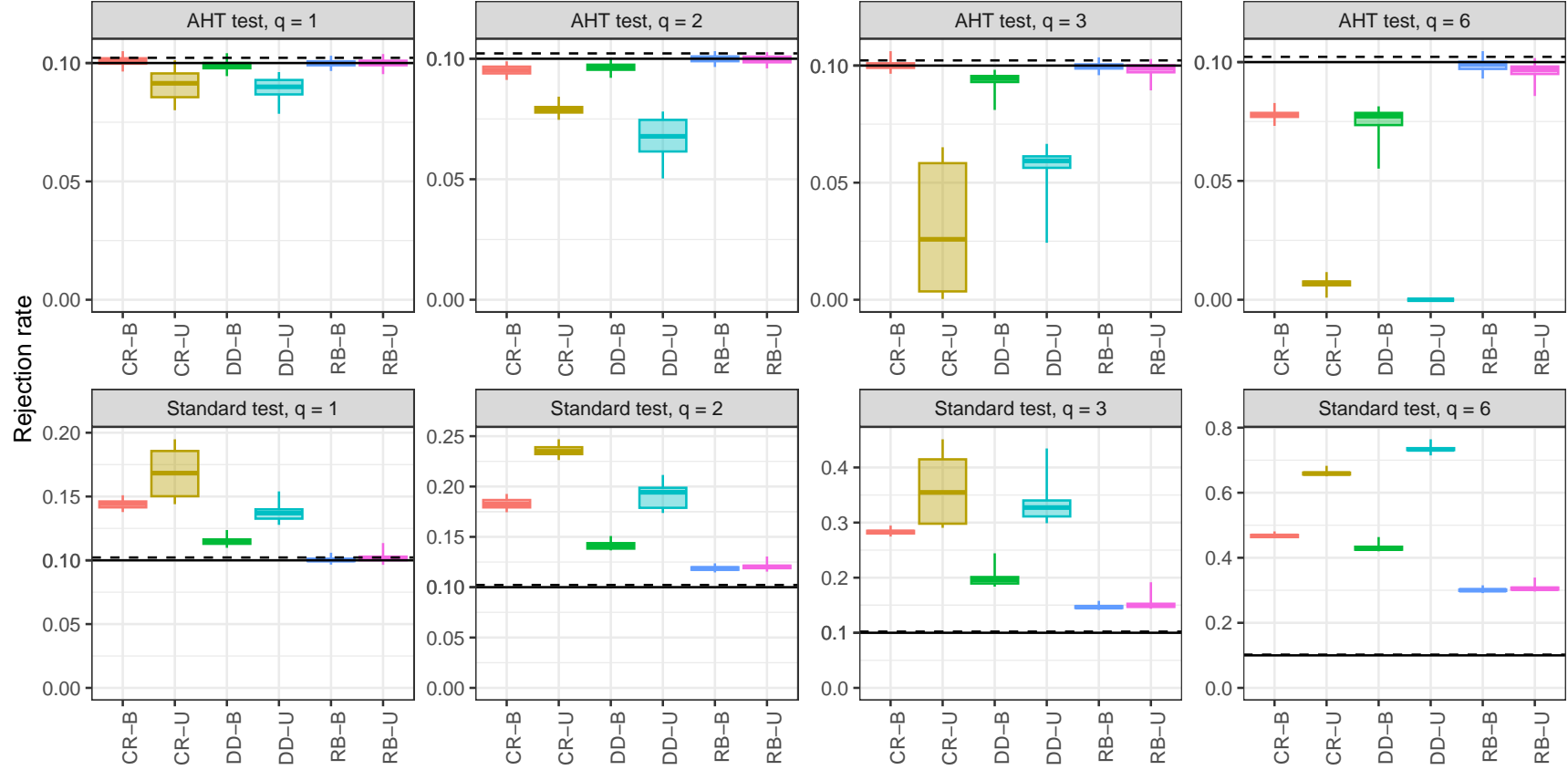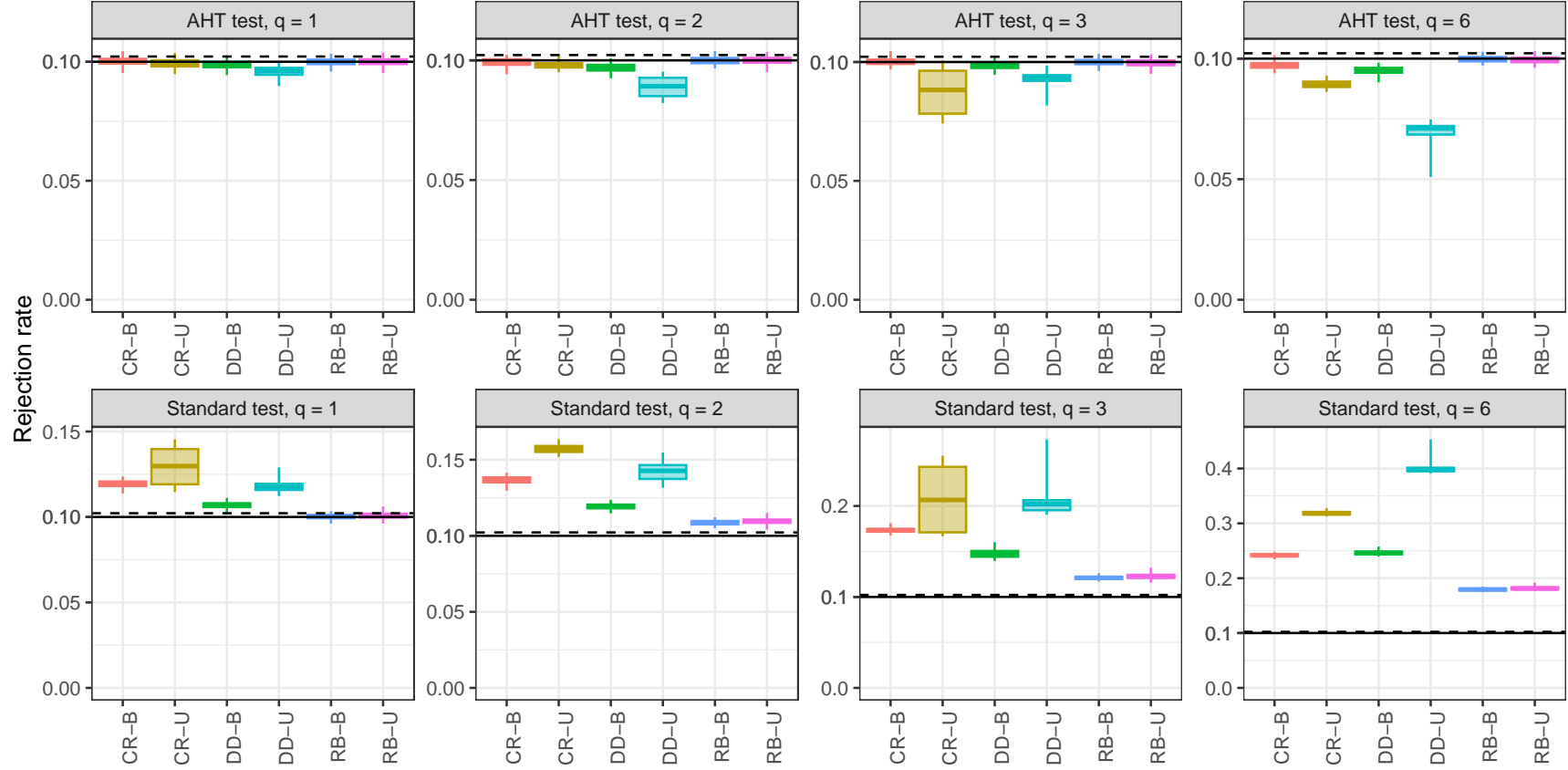
Figure S7: Rejection rates of AHT and standard tests, by study design and dimension of hypothesis (q) for $\alpha = .10$ and $m = 15$. CR = cluster-randomized design; DD = difference-in-differences design; RB = randomized block design; B = balanced; U = unbalanced.
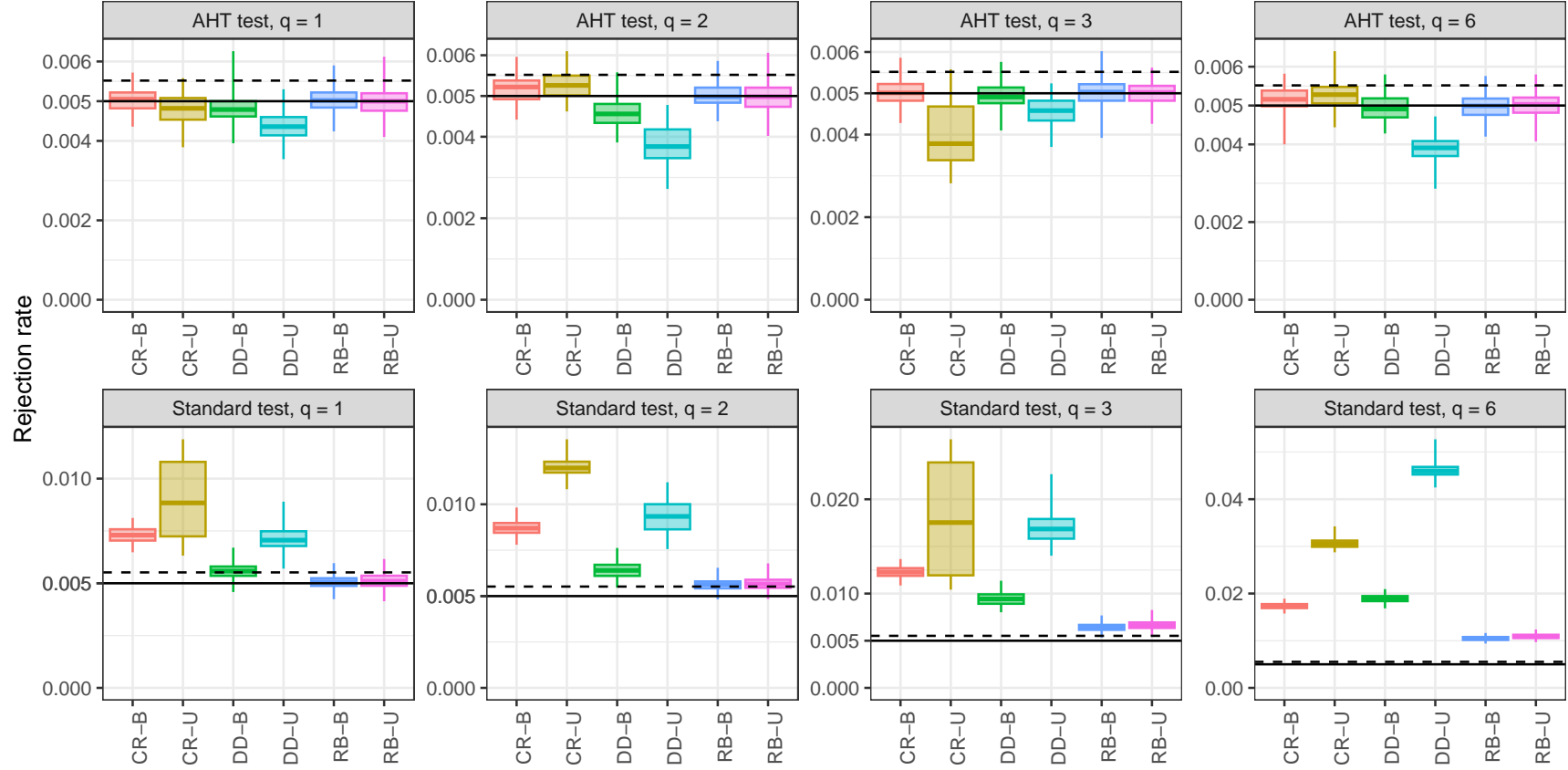
Figure S8: Rejection rates of AHT and standard tests, by study design and dimension of hypothesis (q) for $\alpha = .005$ and $m = 30$. CR = cluster-randomized design; DD = difference-in-differences design; RB = randomized block design; B = balanced; U = unbalanced.
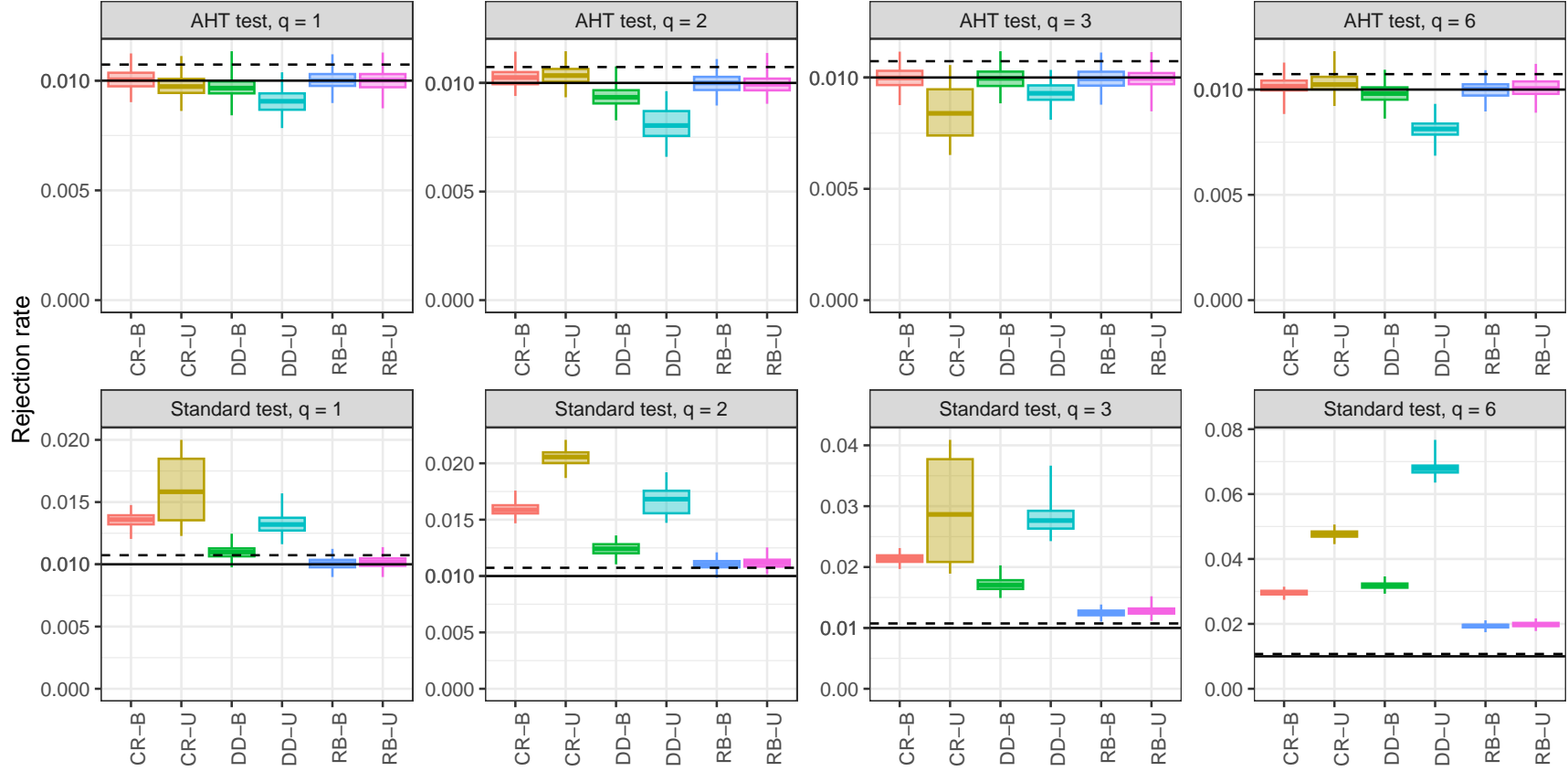
Figure S9: Rejection rates of AHT and standard tests, by study design and dimension of hypothesis (q) for $\alpha = .01$ and $m = 30$. CR = cluster-randomized design; DD = difference-in-differences design; RB = randomized block design; B = balanced; U = unbalanced.

Figure S10: Rejection rates of AHT and standard tests, by study design and dimension of hypothesis (q) for $\alpha = .05$ and $m = 30$. CR = cluster-randomized design; DD = difference-in-differences design; RB = randomized block design; B = balanced; U = unbalanced.
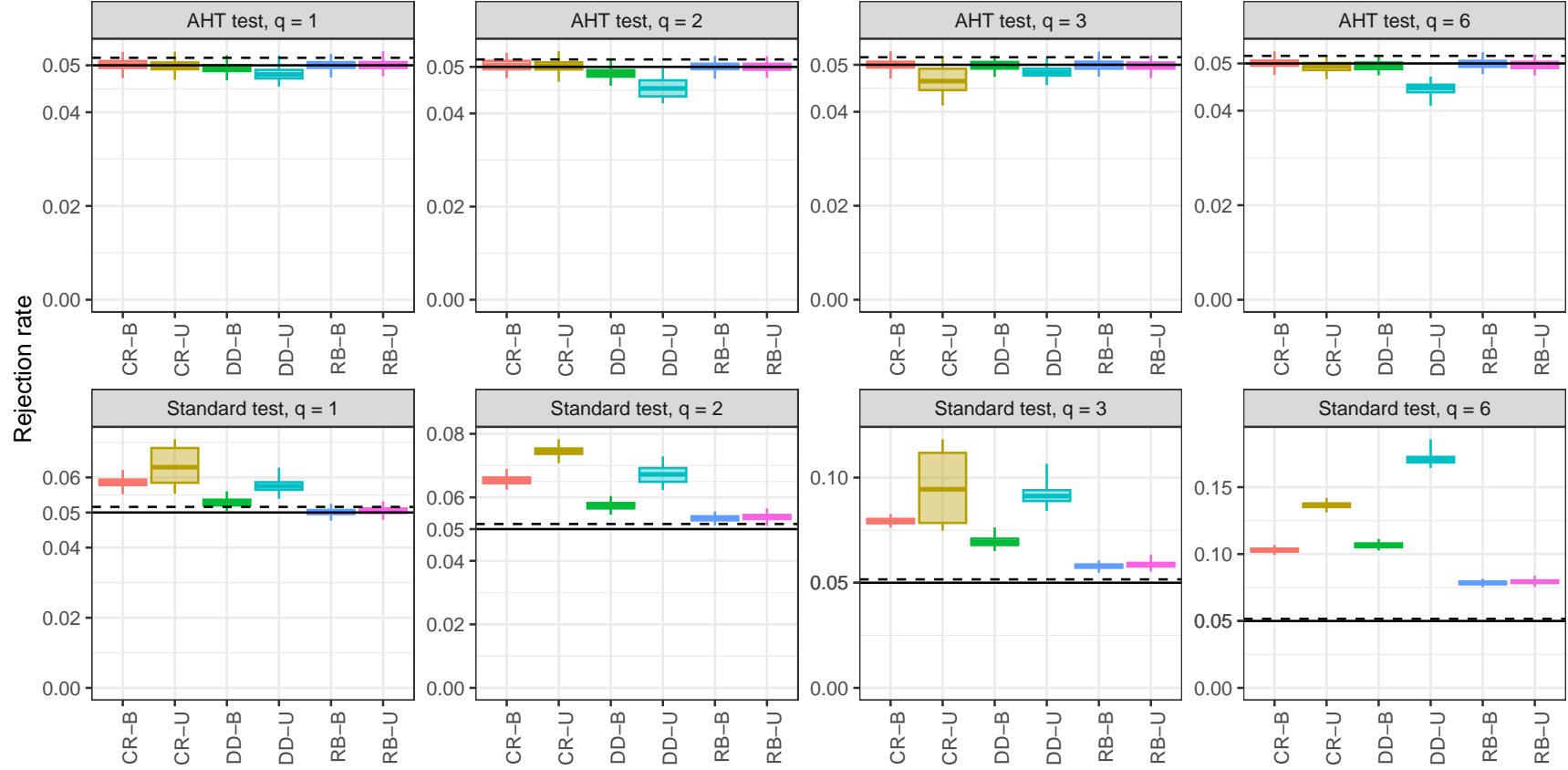
Figure S11: Rejection rates of AHT and standard tests, by study design and dimension of hypothesis (q) for $\alpha = .10$ and $m = 30$. CR = cluster-randomized design; DD = difference-in-differences design; RB = randomized block design; B = balanced; U = unbalanced.
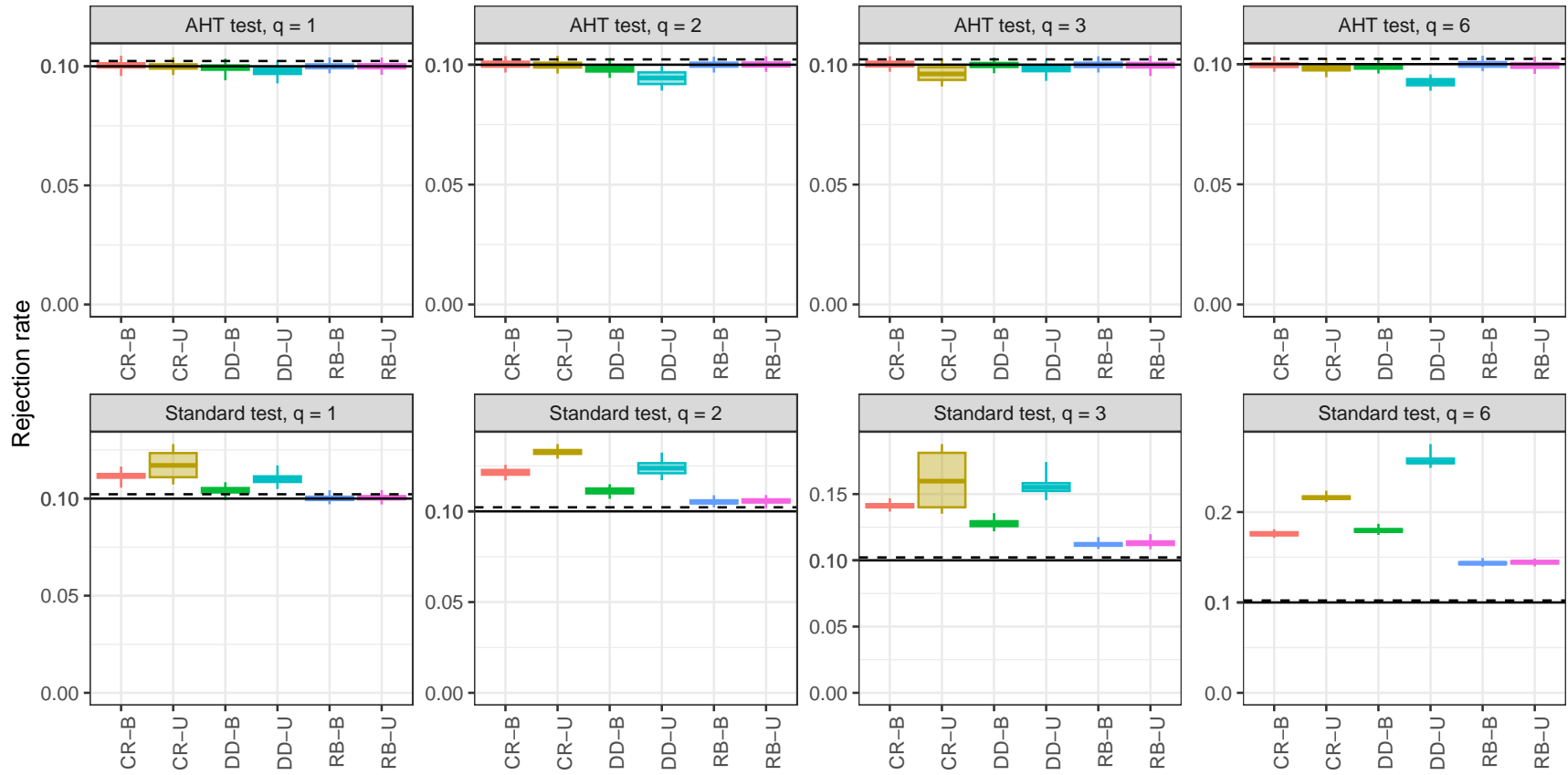
Figure S12: Rejection rates of AHT and standard tests, by study design and dimension of hypothesis (q) for $\alpha = .005$ and $m = 50$. CR = cluster-randomized design; DD = difference-in-differences design; RB = randomized block design; B = balanced; U = unbalanced.

Figure S13: Rejection rates of AHT and standard tests, by study design and dimension of hypothesis (q) for $\alpha = .01$ and $m = 50$. CR = cluster-randomized design; DD = difference-in-differences design; RB = randomized block design; B = balanced; U = unbalanced.

Figure S14: Rejection rates of AHT and standard tests, by study design and dimension of hypothesis (q) for $\alpha = .05$ and $m = 50$. CR = cluster-randomized design; DD = difference-in-differences design; RB = randomized block design; B = balanced; U = unbalanced.

Figure S15: Rejection rates of AHT and standard tests, by study design and dimension of hypothesis (q) for $\alpha = .10$ and $m = 50$. CR = cluster-randomized design; DD = difference-in-differences design; RB = randomized block design; B = balanced; U = unbalanced.

S4.3 Rejection rates of AHT test using CR1 or CR2, with and without accounting for absorption
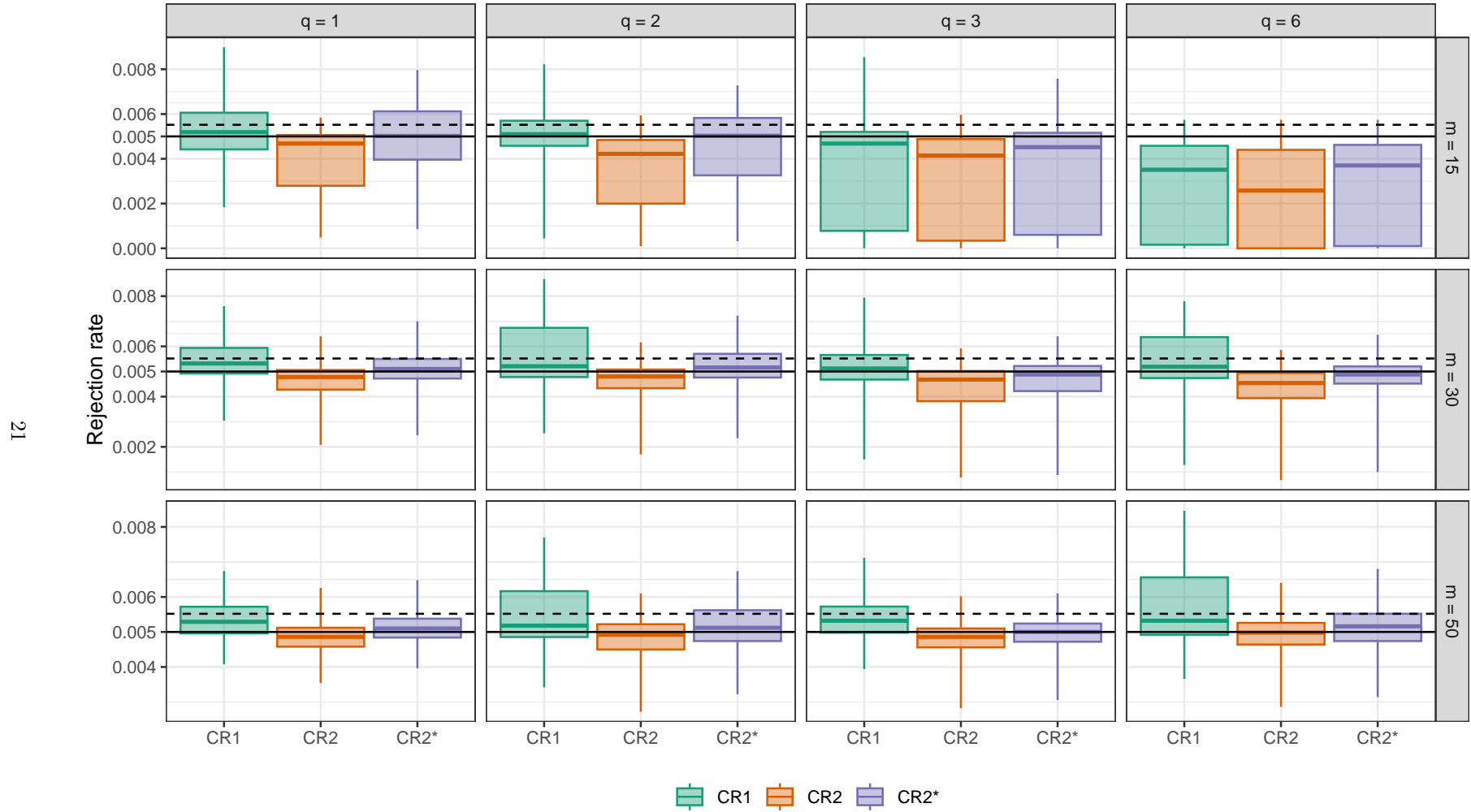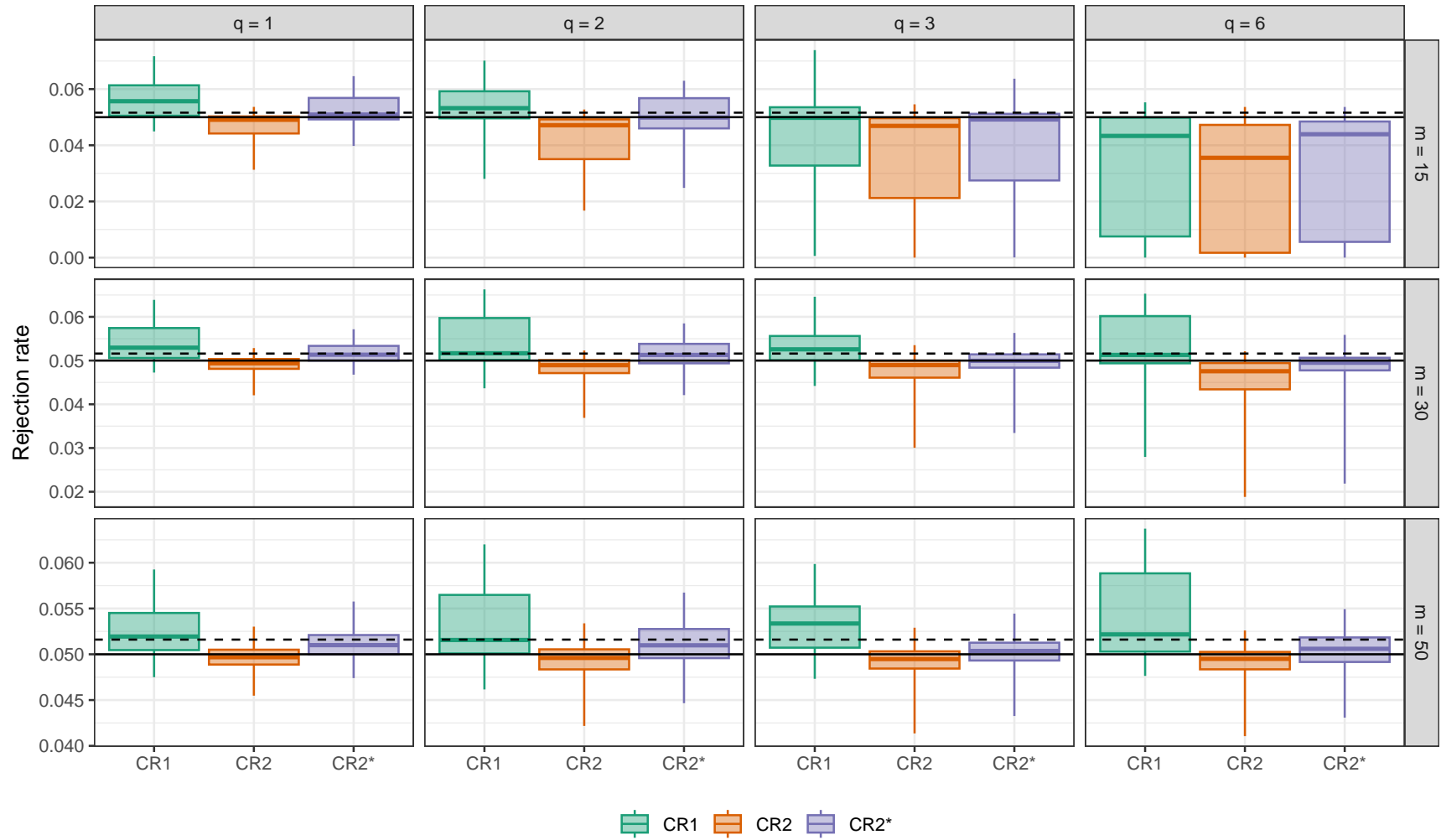

Figure S16: Rejection rates of AHT tests using CR1, CR2, or CR2 calculated without accounting for absorption of fixed effects (CR2*), by sample size (m) and dimension of hypothesis (q), for $\alpha = .005$.
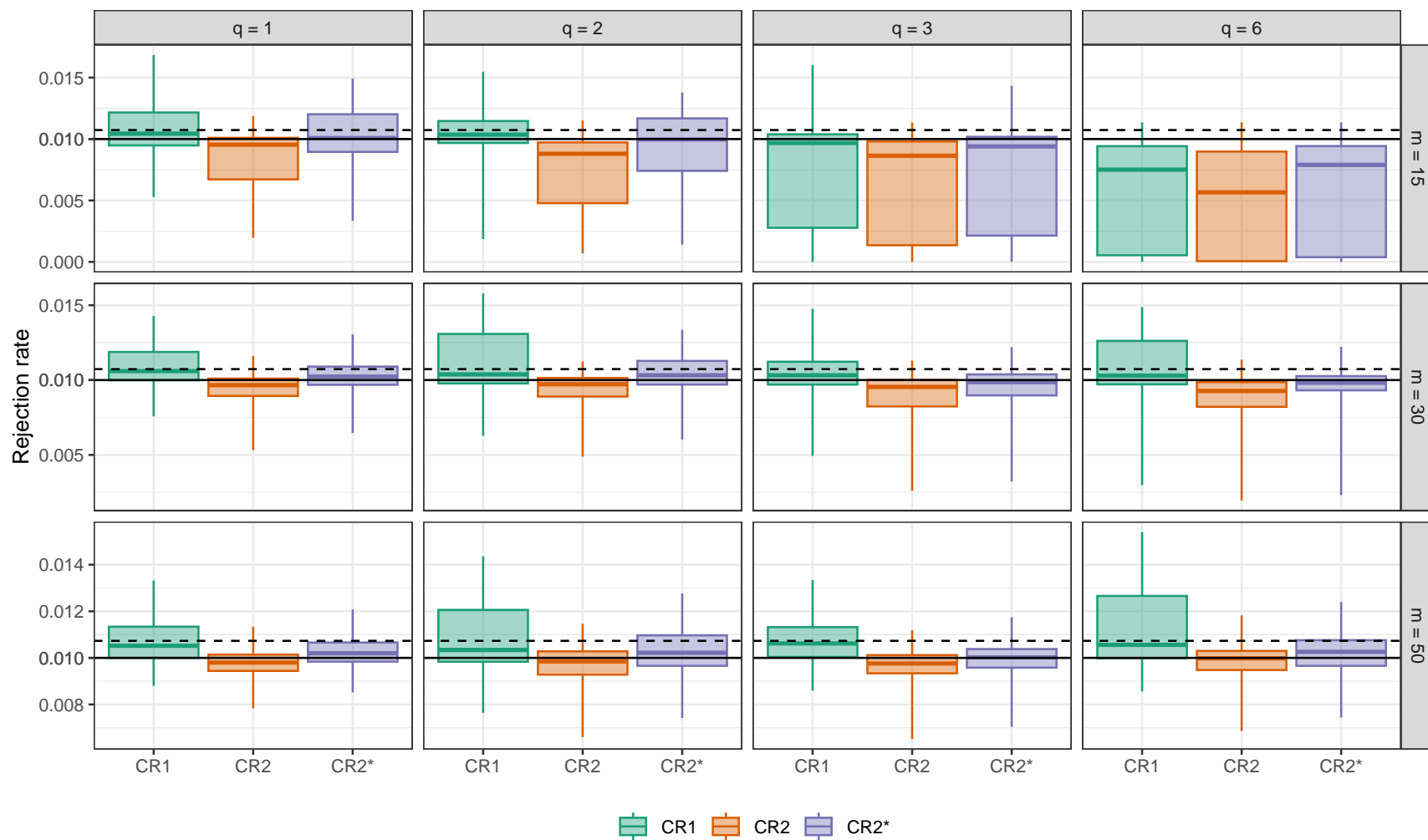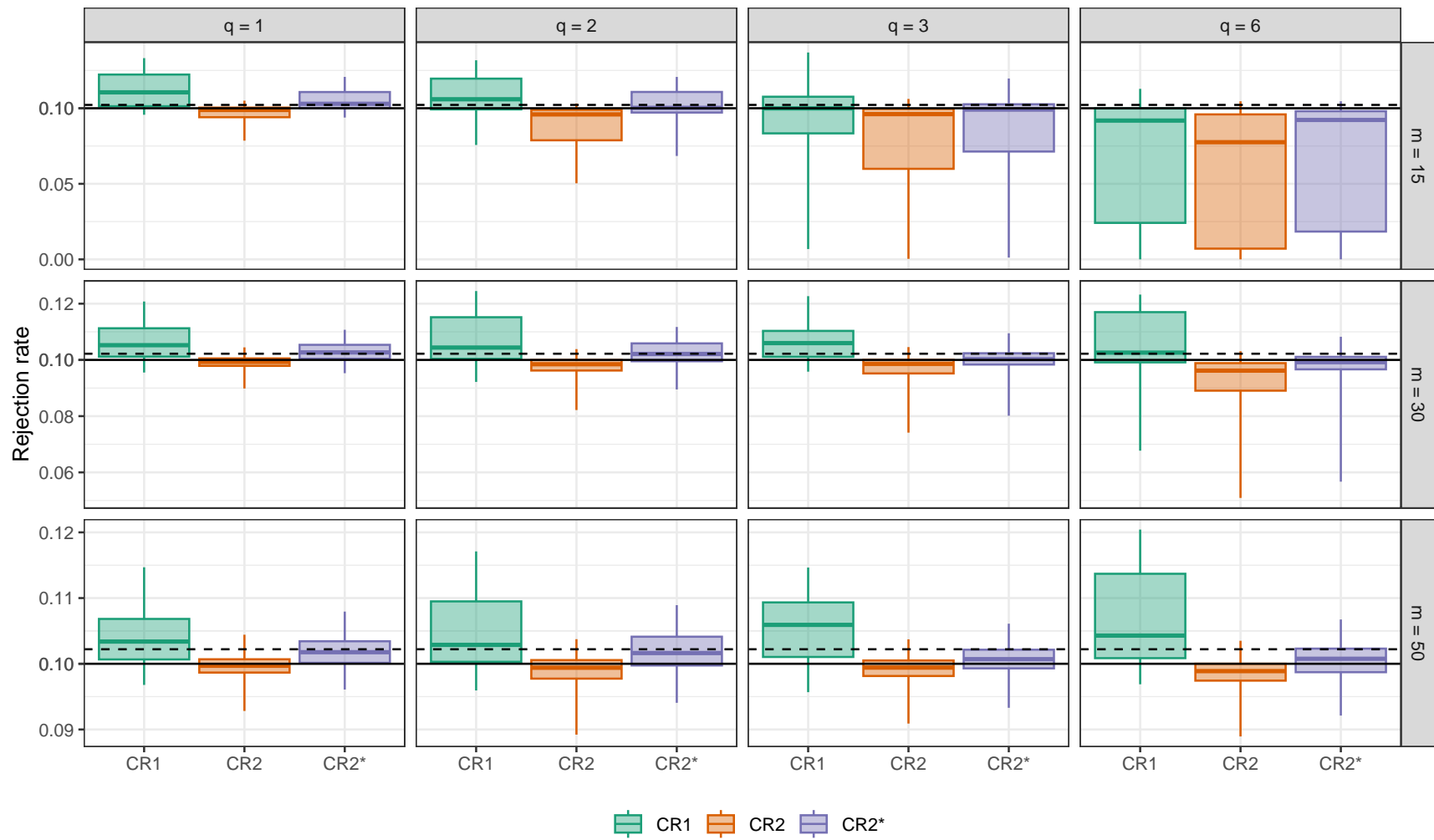
Figure S17: Rejection rates of AHT tests using CR1, CR2, or CR2 calculated without accounting for absorption of fixed effects (CR2*), by sample size (m) and dimension of hypothesis (q), for $\alpha = .01$.

Figure S18: Rejection rates of AHT tests using CR1, CR2, or CR2 calculated without accounting for absorption of fixed effects (CR2*), by sample size (m) and dimension of hypothesis (q), for $\alpha = .05$.

Figure S19: Rejection rates of AHT tests using CR1, CR2, or CR2 calculated without accounting for absorption of fixed effects (CR2*), by sample size (m) and dimension of hypothesis (q), for $\alpha = .10$.

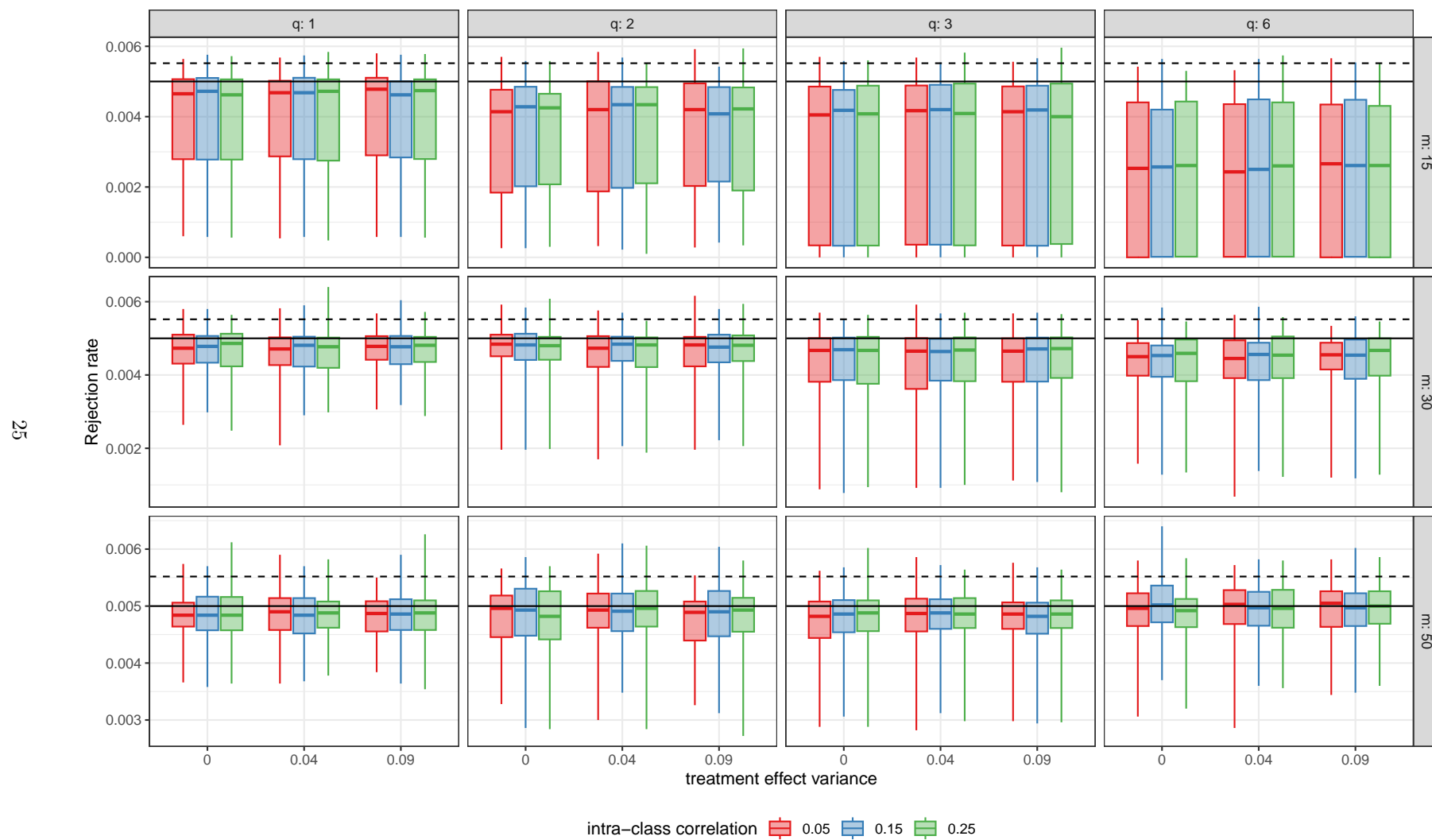## S4.4 Rejection rates of AHT test by degree of working model misspecification

Figure S20: Rejection rates of CR2 AHT test, by treatment effect variance and intra-class correlation for $\alpha = .005$.
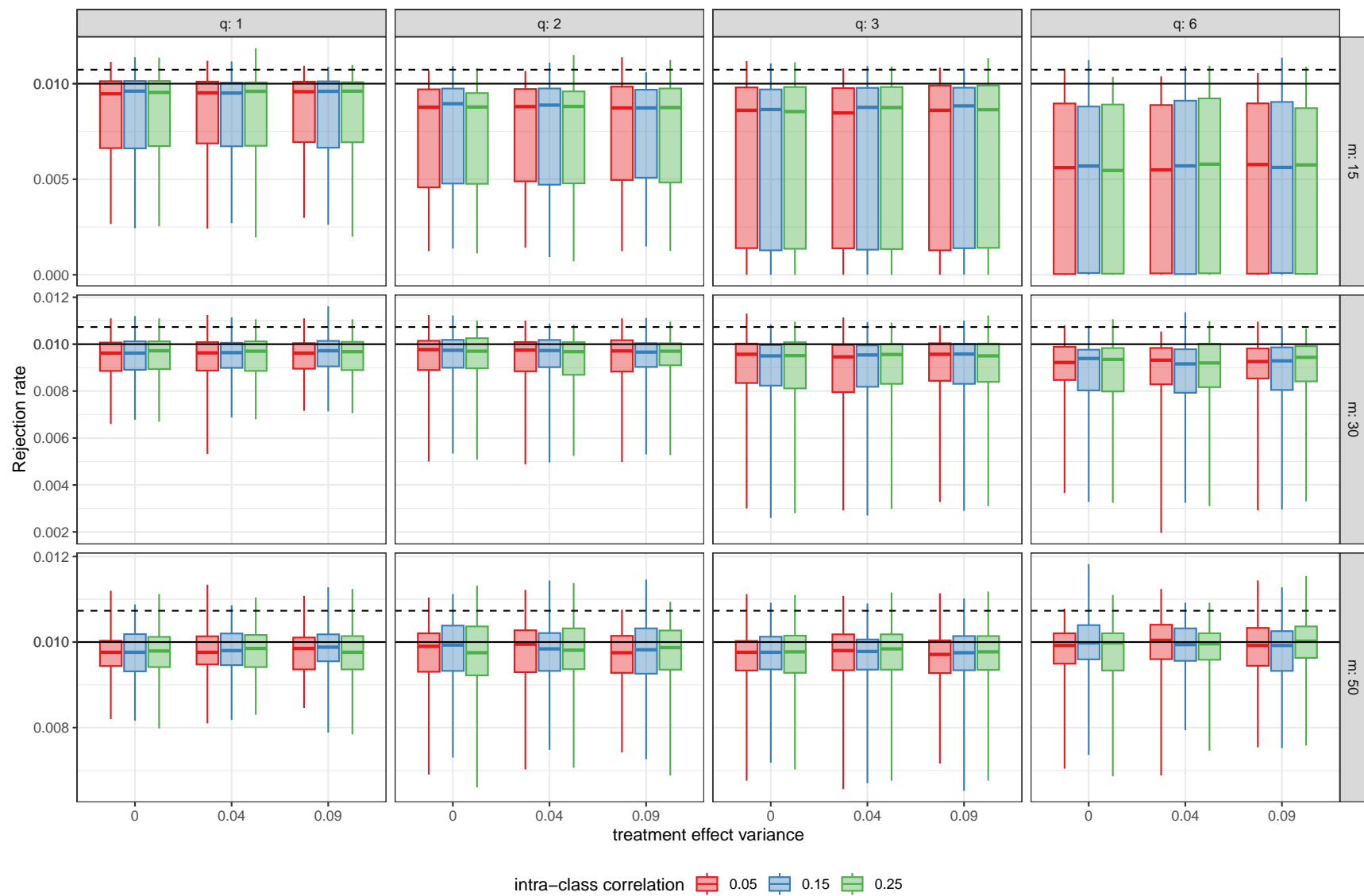
Figure S21: Rejection rates of CR2 AHT test, by treatment effect variance and intra-class correlation for $\alpha = .01$.
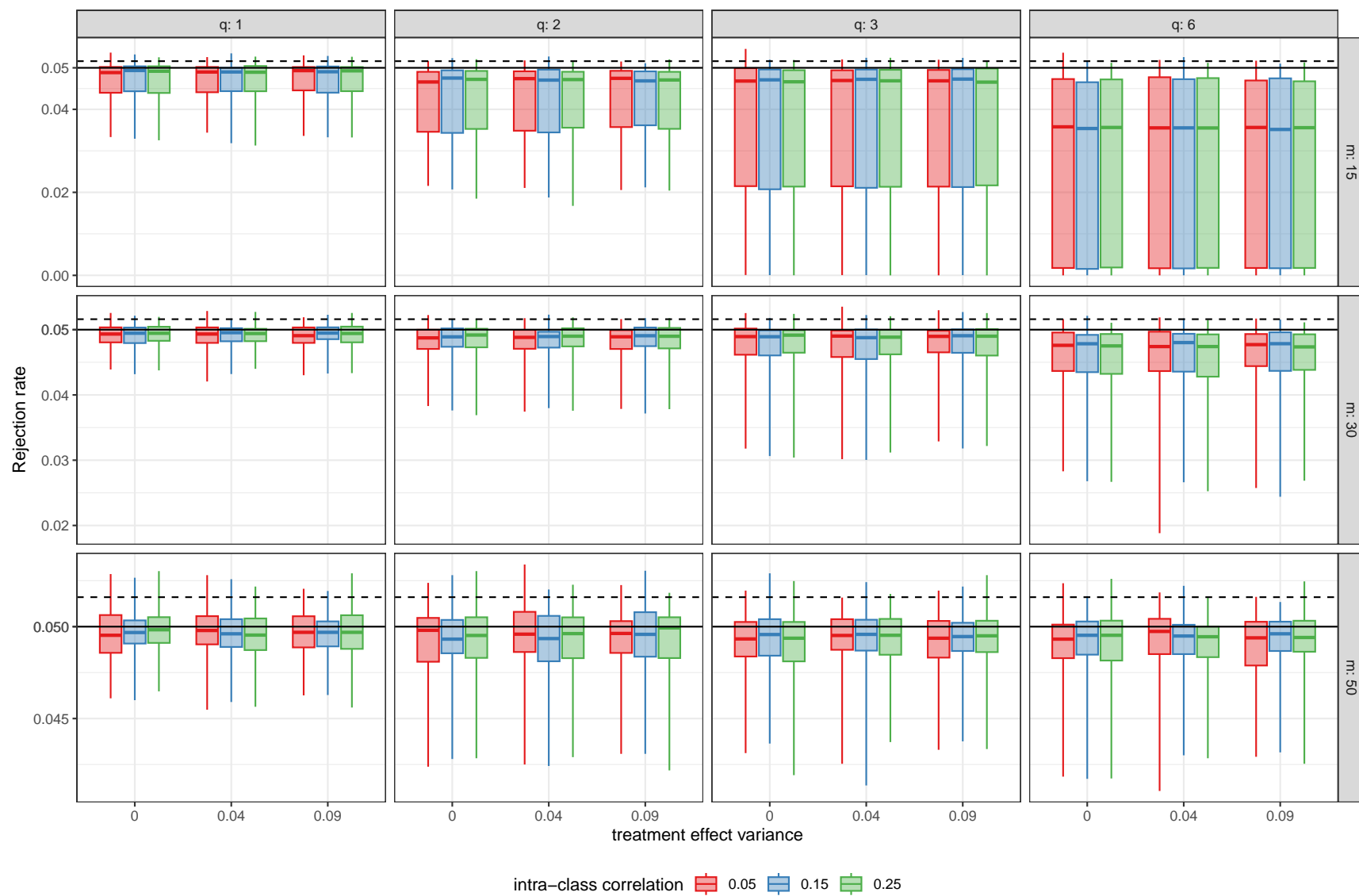
Figure S22: Rejection rates of CR2 AHT test, by treatment effect variance and intra-class correlation for $\alpha = .05$.
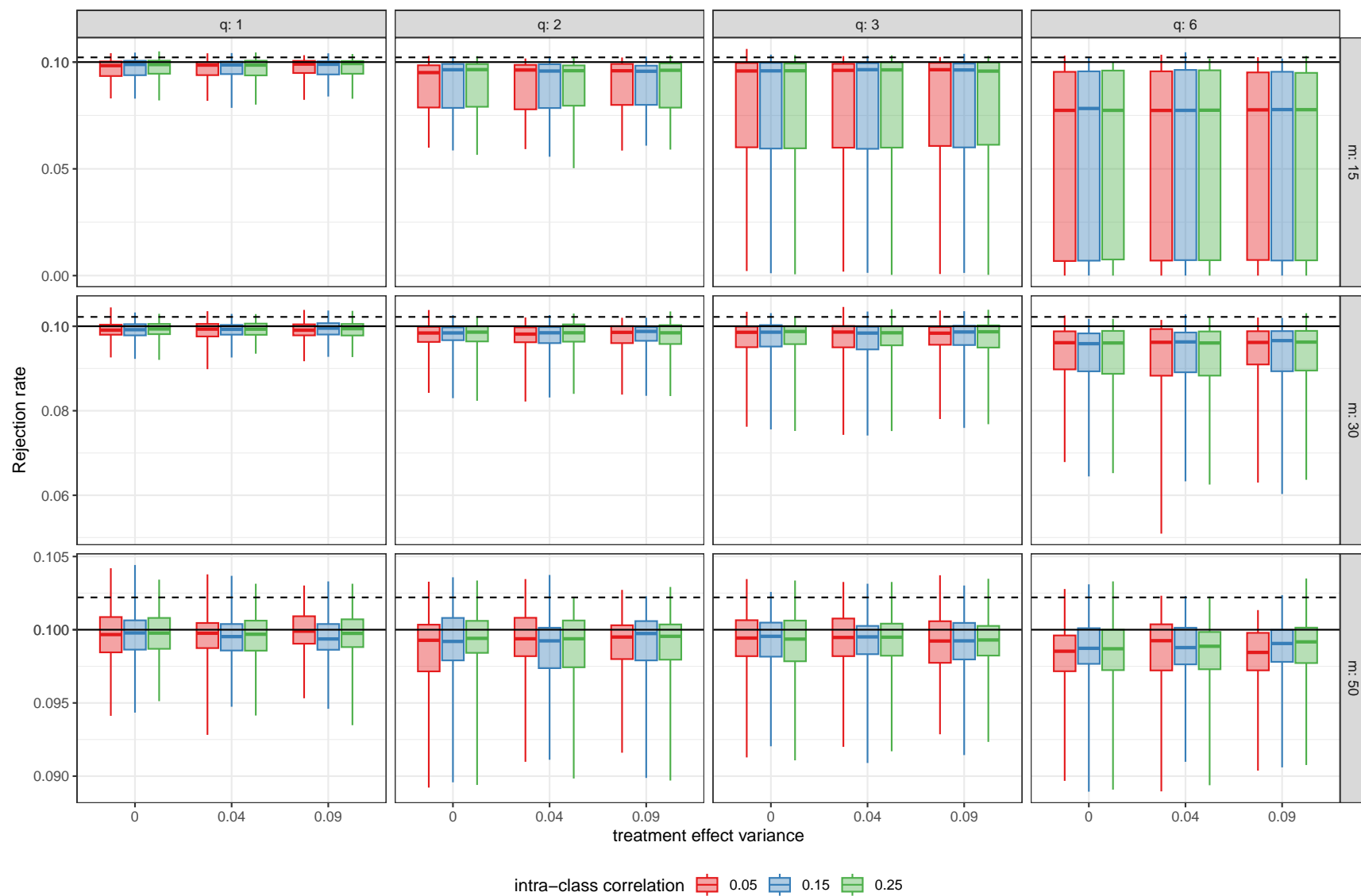
Figure S23: Rejection rates of CR2 AHT test, by treatment effect variance and intra-class correlation for $\alpha = .10$.