FACULTY OF PSYCHOLOGY AND EDUCATIONAL SCIENCES

DEPARTMENT OF DATA-ANALYSIS (PP01)

HOW TO SAFELY RE-ASSESS VARIABILITY AND ADAPT SAMPLE SIZE?

*A primer for the **independent samples t-test***

Presenter: Lara Vankelecom

Promotors: Beatrijs Moerkerke, Tom Loeys

GHENT UNIVERSITY

INTRODUCTION

! POWER !

→ File-drawer problem
→ Replication crisis

Traditional fixed design approach:

A-priori sample size calculation

Adaptive design approach:

Internal pilot study design

GHENT UNIVERSITY

A-PRIORI SAMPLE SIZE CALCULATION

$$N = \frac{(z_{1-\alpha} + z_{1-\beta})^2}{\delta^2} \cdot 2 \cdot \sigma^2$$

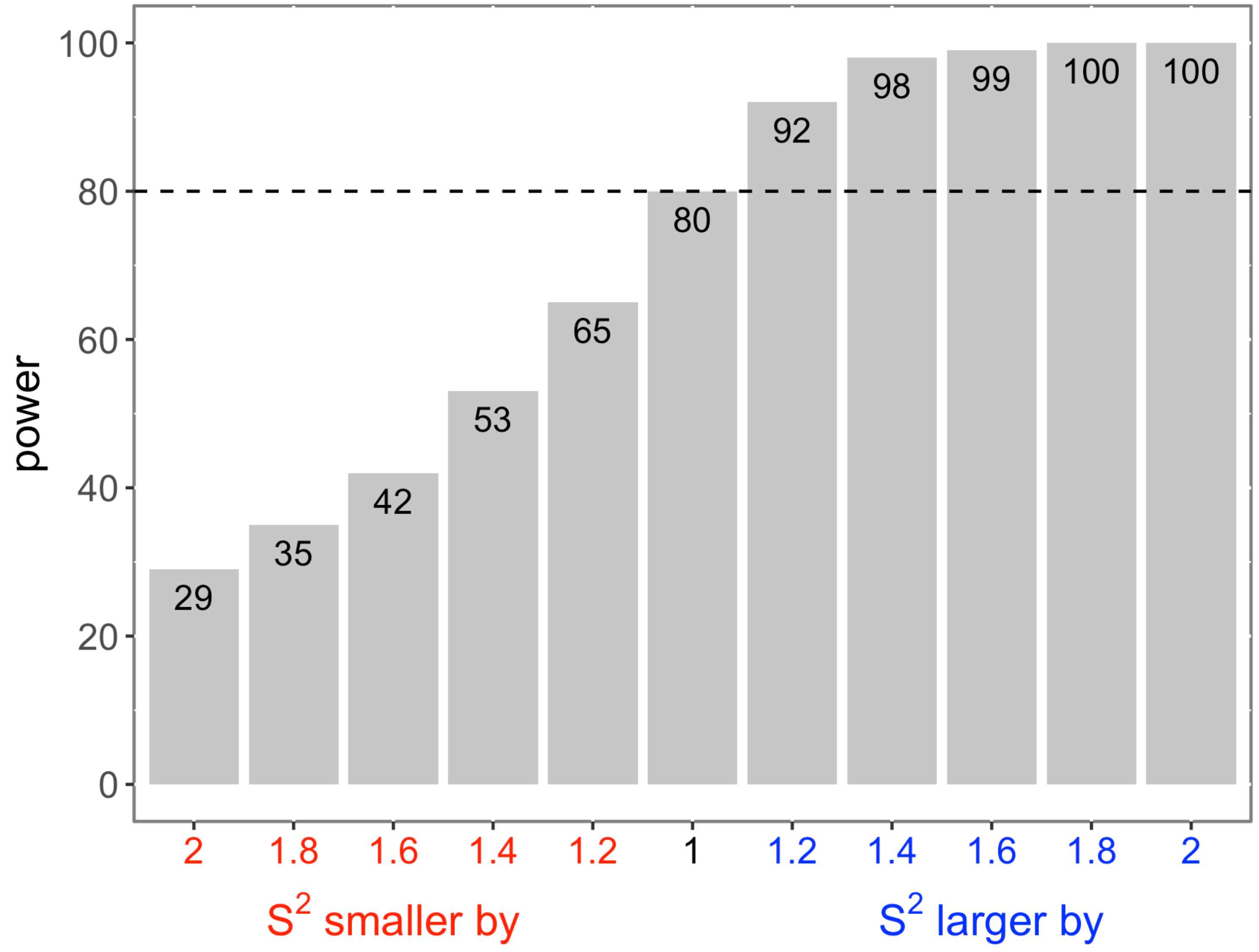
Two population parameters:

1) Population difference in means δ

   → Specify as smallest effect size of interest (SESOI)

2) Population variance σ^2

   → Estimate by S^2



INTERNAL PILOT STUDY DESIGN

Reasoning:

→ Collect pilot data to obtain estimate of S^2

→ Instead of discarding the pilot data, include them in the final dataset

Or in other words:

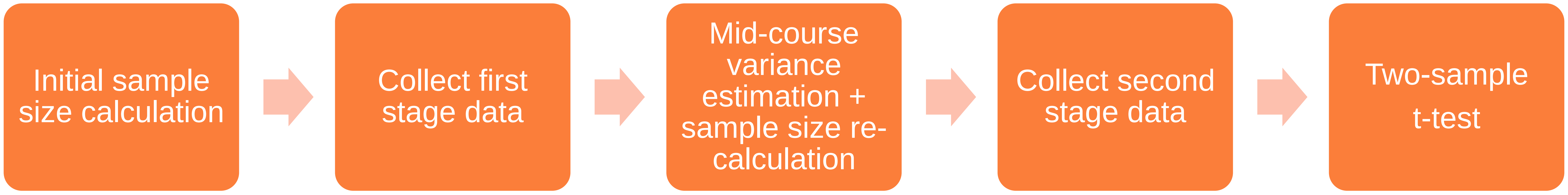
→ Estimate variance after first collected data and re-estimate sample size during course of the study

(Two-stage) adaptive design

Blinded sample size re-estimation

GHENT UNIVERSITY

BLINDED SAMPLE SIZE RE-ESTIMATION



Purely blinded:

- Naive one-sample estimate

  → Biased when effect ≠ 0

- Type I error rate inflation = negligible

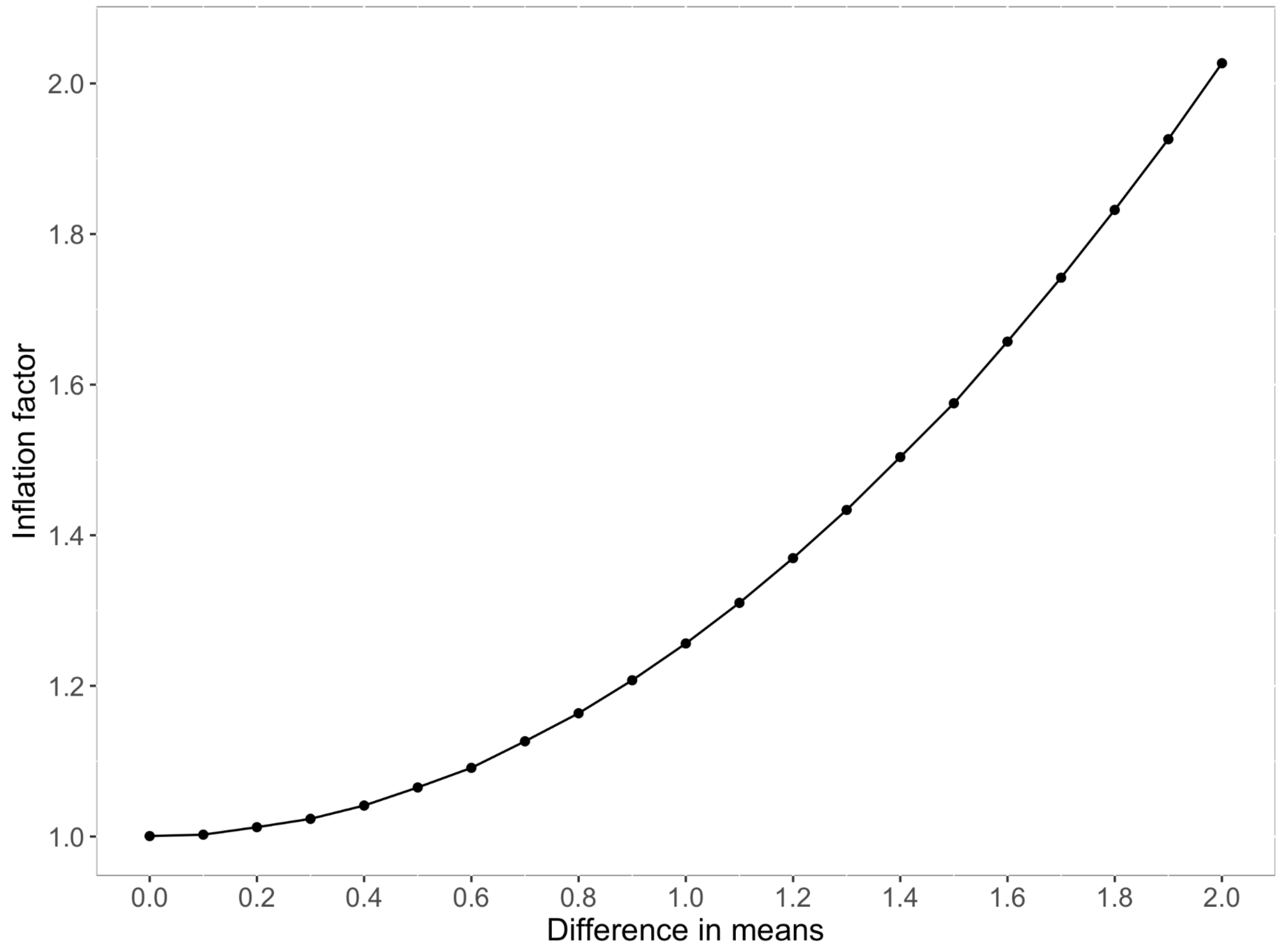
HOW BIASED IS NAIVE ESTIMATOR?

→ When effect ≠ 0, the one-sample estimator overestimates the true variance in the data

→ Sample size calculation based on one-sample estimator will overestimate the true required sample size

→ How much?

$$\mathrm{IF} = \frac{\text{sample size naive estimator}}{\text{sample size true variance}}$$



GHENT UNIVERSITY

BLINDED SAMPLE SIZE RE-ESTIMATION

Initial sample size calculation → Collect first stage data → Mid-course variance estimation + sample size re-calculation → Collect second stage data → Two-sample t-test

Purely blinded:

- Naive one-sample estimate

  → Biased when effect ≠ 0

- Type I error rate inflation = negligible

Partially-blinded:

- Pooled variance estimate

  → Not biased

- Slight type I error rate inflation

SO FAR...

Previous research on the IPS design:

→ Solely focused on the Student t-test (assumed common variance between the two groups)

  → Remember: naive or pooled interim variance estimation

→ Only described for equal group sizes

Goals current study:

→ How to implement the design when researcher cannot assume homoscedasticity and/or is not able to collect equally-sized groups?

→ What is the T1ER and the power of the Welch t-test in this design?

IMPLEMENTATION IPS DESIGN

→ Procedure remains the same, sample size formula needs to be modified

Previous:

$$N = \frac{(z_{1-\alpha} + z_{1-\beta})^2}{\delta^2} \cdot 2 \cdot \sigma^2$$

Current study:

$$N_1 = \frac{(z_{1-\alpha} + z_{1-\beta})^2}{\delta^2 \cdot \pi_2} . \sigma_1^2 . (\pi_1 + \pi_2)$$

$$\pi_1 = \frac{\text{variance group 2}}{\text{variance group 1}}$$

$$\pi_2 = \frac{\text{size group 2}}{\text{size group 1}}$$

SIMULATION STUDIES – ERROR RATES

→ Simulate **type I error rate** and **power** of final Welch t-test when implementing this design

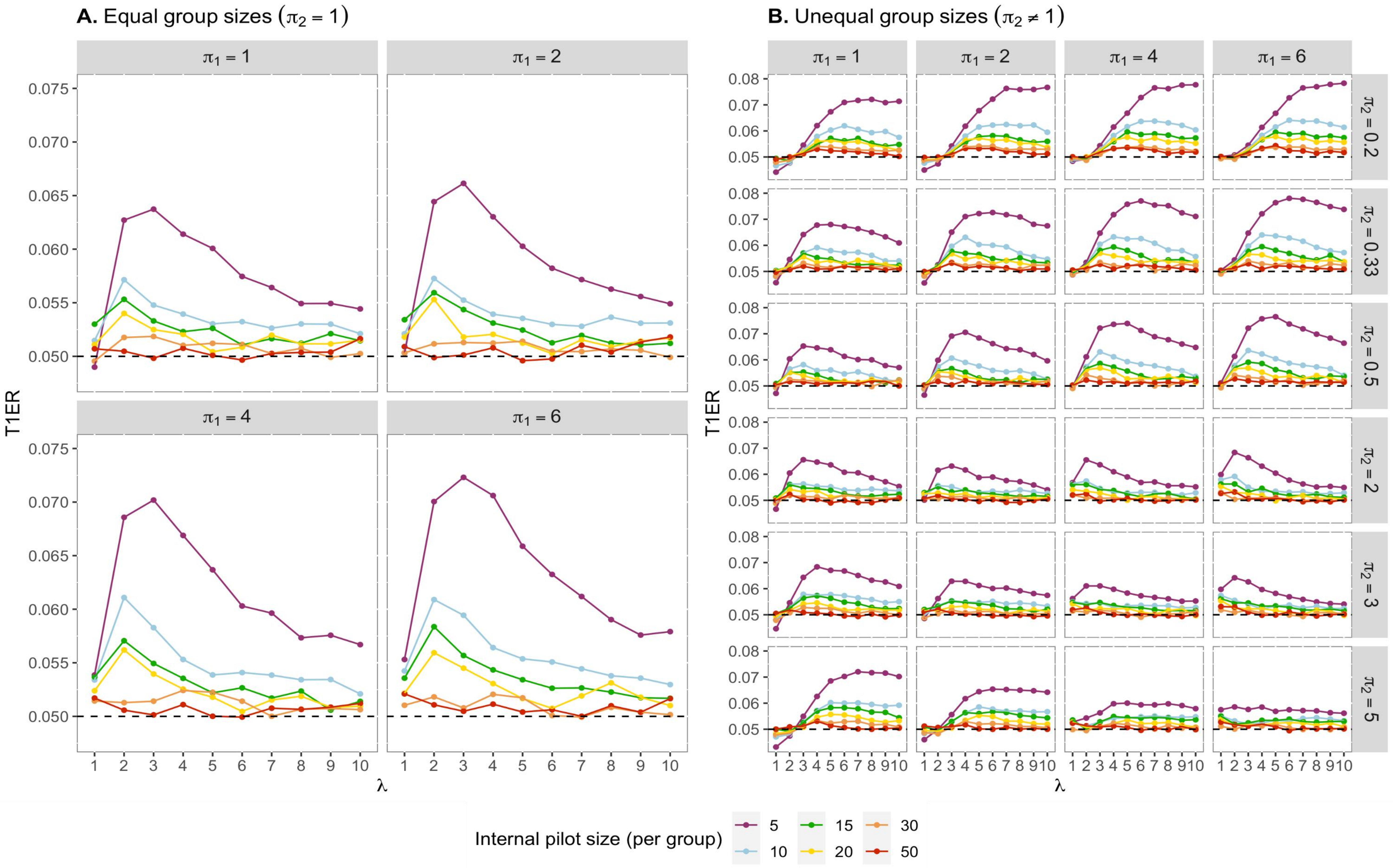
→ For different experimental scenarios: 2x2 of (un)equal variances – (un)equal group sizes

- Equal group size (π_2 = 1)
  - Equal variances (π_1 = 1)
  - Unequal variances (π_1 = 2, 4, 6)
- Unequal group size (π_2 = 2, 3, 5)
  - Equal variances (π_1 = 1)
  - Unequal variances
    - Large group – large variance (π_1 = 2, 4, 6)
    - Large group – small variance (π_1 = 1/2, 1/4, 1/6)

GHENT UNIVERSITY

TYPE I ERROR RATE

$$\lambda = \frac{true\ required\ sample\ size}{internal\ pilot\ size}$$



IN PRACTICE?

Problem?

→ True required sample size not known (= estimated)

→ $\boldsymbol{\lambda}$ is not known

→ Exact inflation of the type I error rate for a particular study cannot be identified

Solution?

→ Internal pilot size and $\boldsymbol{\pi}_2$ are known, $\boldsymbol{\pi}_1$ can be estimated

→ Search for the maximized inflation (over $\boldsymbol{\lambda}$ values)

→ Based on maximized inflation, an adjusted $\boldsymbol{\alpha}$-level can be determined to use in Welch t-test

GHENT UNIVERSITY

R FUNCTION

```
find_adjalpha(n1.1 = n1.1, est.pi1 = est.pi1, pi2 = pi2, alpha_nom = alpha_nom,
              upper = 0.80, range = 200, step = 10, range2 = 15, step2= 1,
              diff = 1, eps = 0.001, asim = 1000000)
```

1. Maximized type I error rate inflation

→ by calculating the actual type I error rate (with simulation) for range of different values of true sample size

→ Important that correct true sample sizes are considered in range

2. Adjusted **α**-level

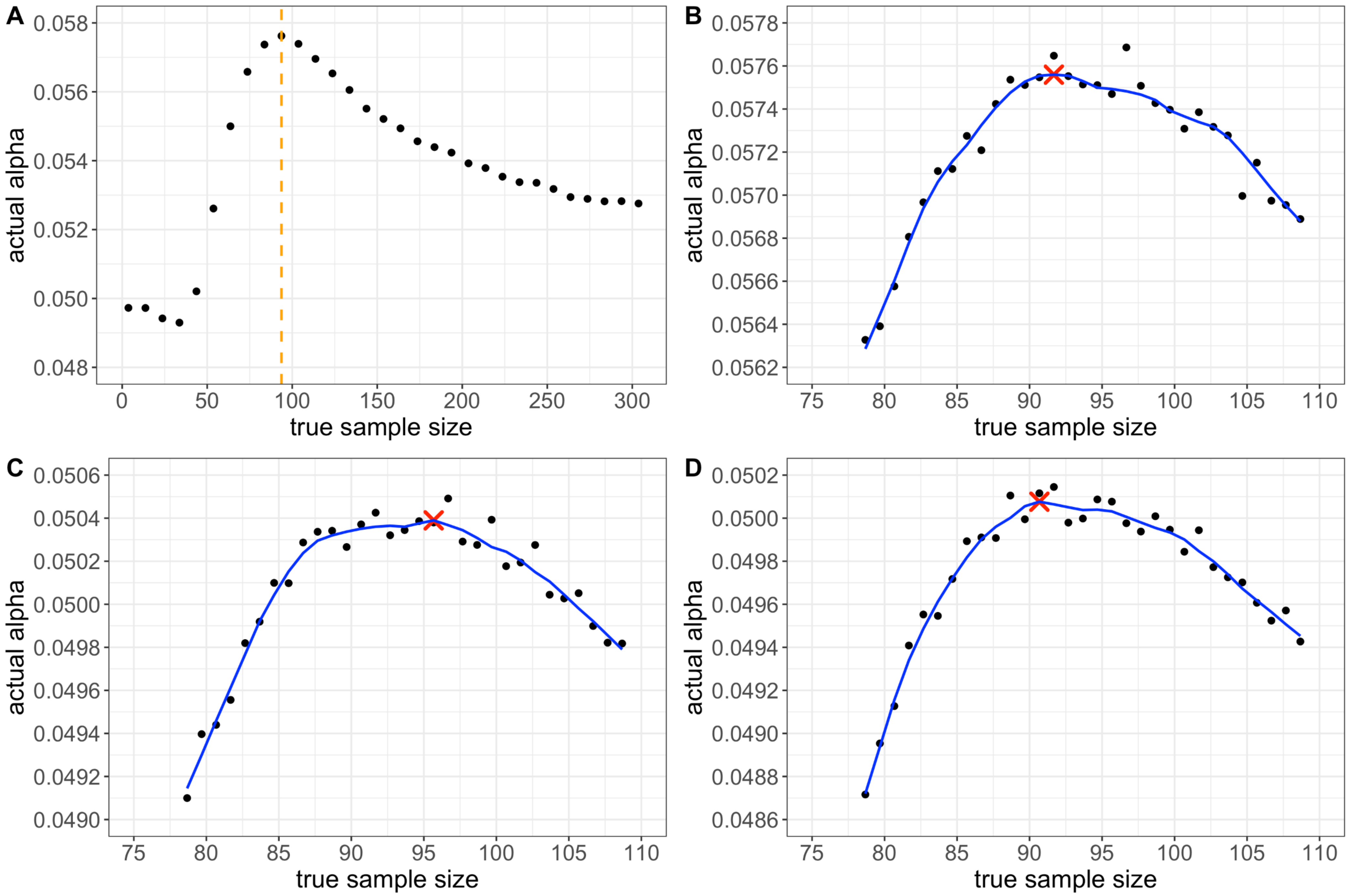
→ by applying iterative algorithm of Kieser & Friede (2000) to this range

R FUNCTION - EXAMPLE

```
find_adjalpha(n1.1 = 15, est.pi1 = 1.67, pi2 = 0.5, alpha_nom = 0.05)
```

```
## [1] "Refine sample size value for this parameter configuration"
## [1] "Searching for the adjusted alpha-level..."
```

```
## $`Adjusted alpha`
## [1] 0.04343347
##
## $`Maximized T1ER`
## [1] 0.0575593
```

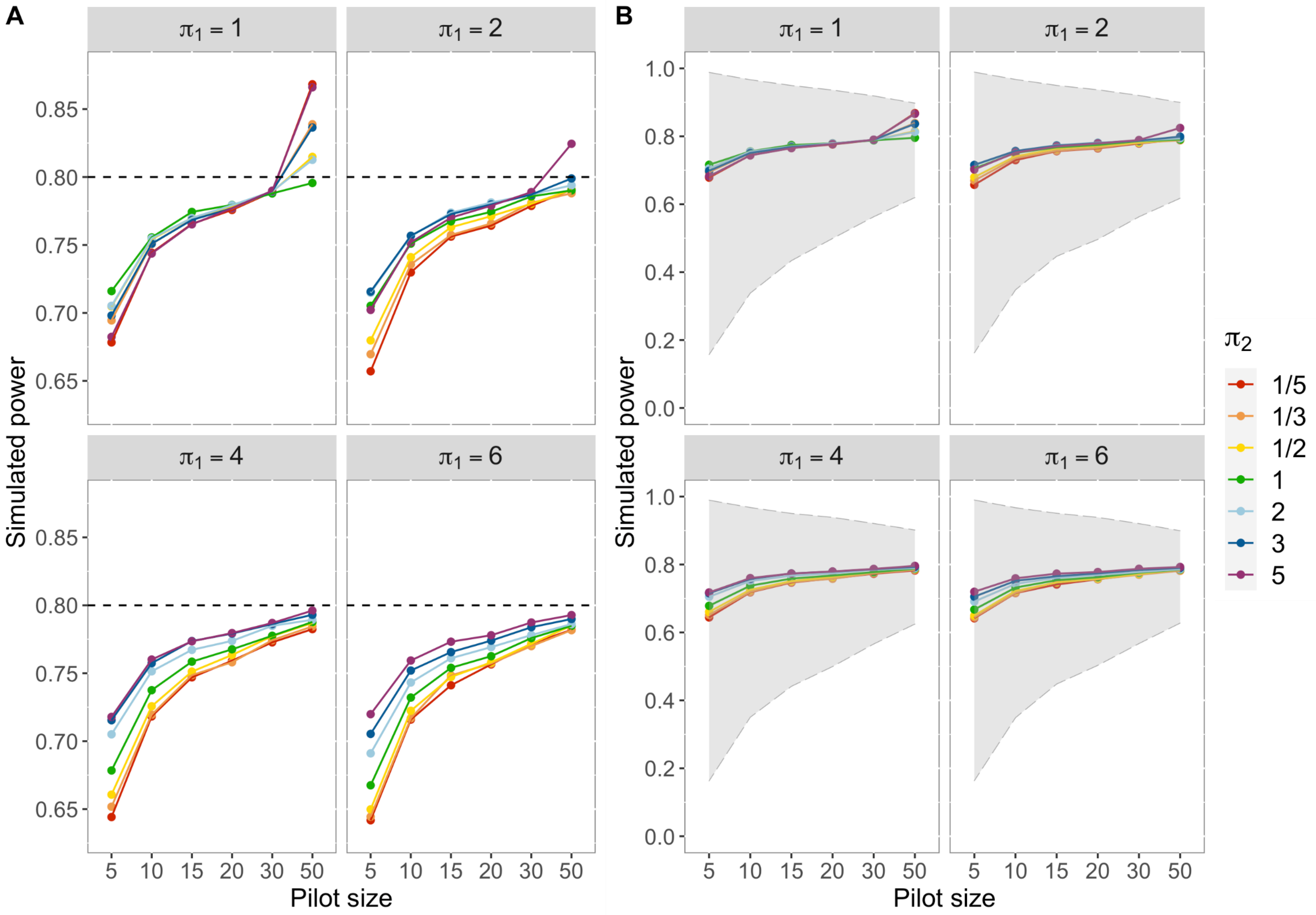


GHENT UNIVERSITY

POWER

- Demonstration of power for $\boldsymbol{\delta}$ = 0.5 (same as SESOI used in sample size re-calculation)
- Variance for group 1 set to 1, for group 2 set to the values of $\boldsymbol{\pi}_1$
- $\boldsymbol{\alpha}_{adj}$ used as rejection boundary

- On average, power approaches the pre-defined level as pilot size increases
- The range of power over simulations decreases with larger internal pilots



DISCUSSION

Advantages:

- This design improves chance of obtaining well-powered study
- Type I error rate inflation is minimal when re-estimating sample size after large internal pilot and can be corrected for when large internal pilot is not feasible
- Researcher is only concerned with a-priori specifying an unstandardized SESOI

Disadvantages:

- Defining such SESOI is not straightforward
- Risk of severely under- or overpowered study when using (very) small internal pilot

FACULTY OF PSYCHOLOGY
AND EDUCATIONAL SCIENCES

Lara Vankelecom

Lara.Vankelecom@ugent.be
+32 9 264 63 51

Department of Data Analysis
Faculty of Psychology and Educational Sciences

Ghent University
Henri Dunantlaan 1
B-9000 Ghent
Belgium

GHENT
UNIVERSITY

ITERATIVE ALGORITHM

Kieser & Friede (2000):

$$\alpha_{act}^{max}(\alpha_{adj}) = \alpha$$

→ The maximized actual T1ER is identified for each newly obtained adjusted **α**-level

→ The final adjusted **α**-level is that value for which the corresponding maximized actual T1ER does not exceed the originally planned nominal level

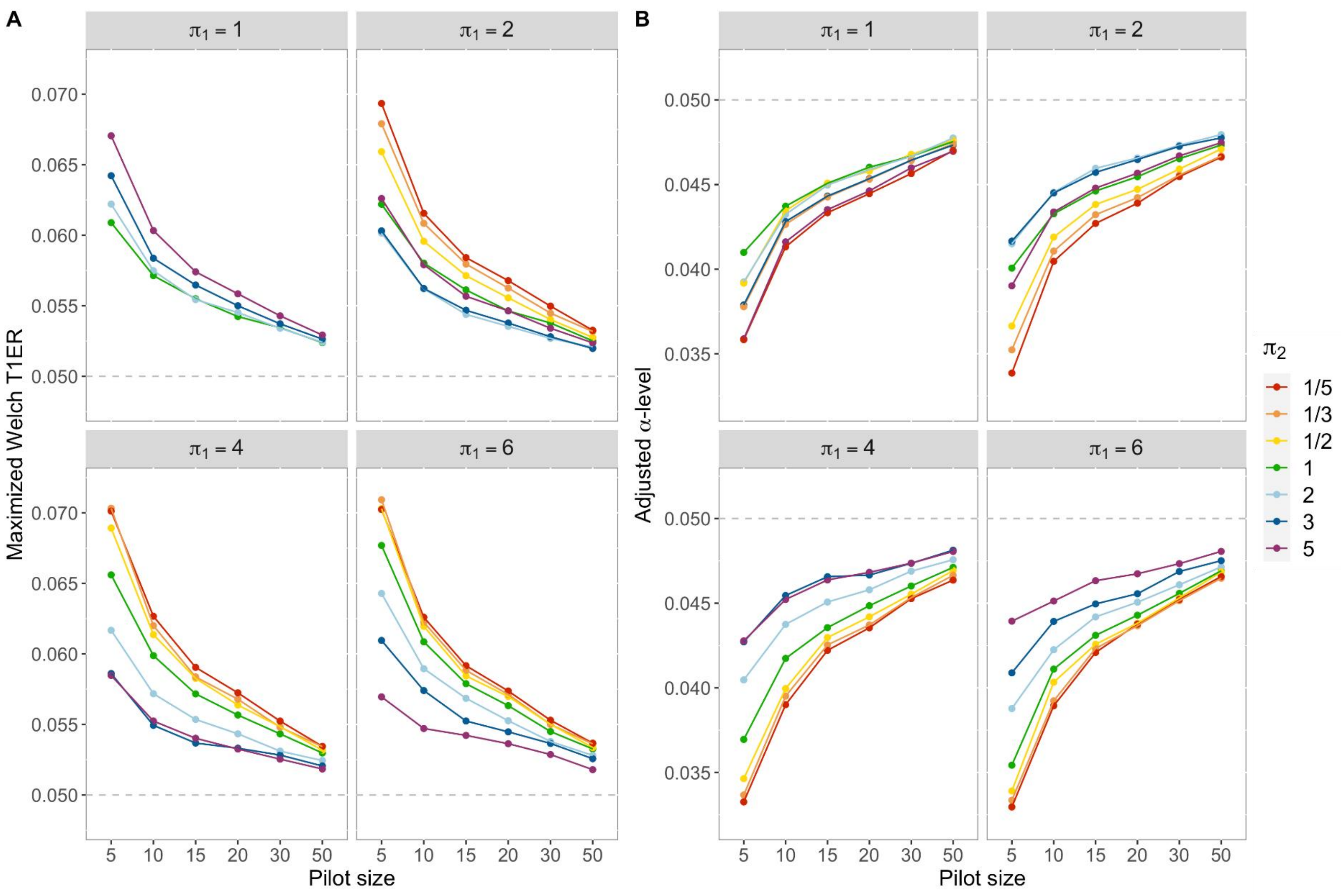
Initial adjusted **α**-level:

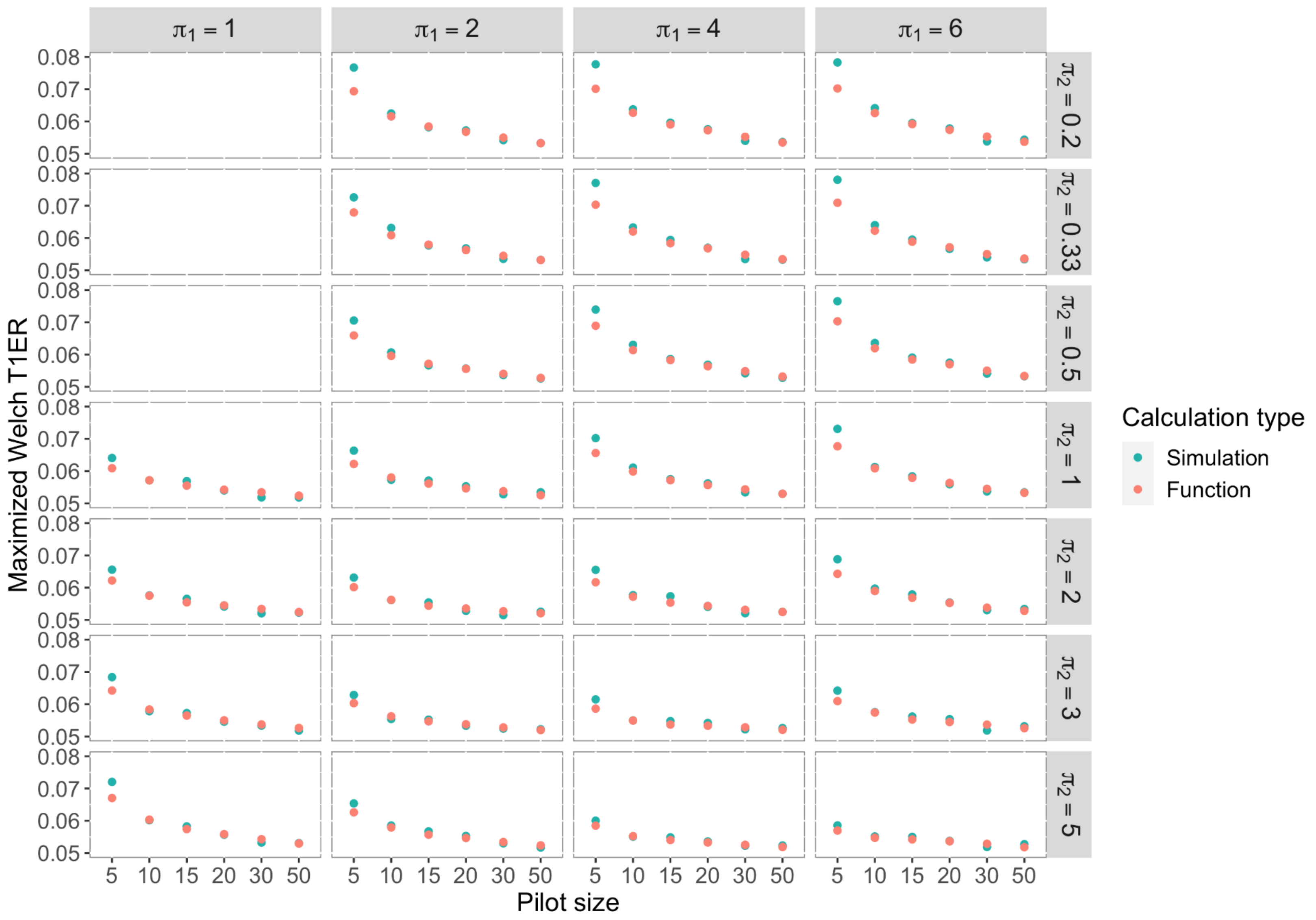
$$\alpha_{adj}^{(0)} = \alpha \cdot \frac{\alpha}{\alpha_{act}^{max}(\alpha)}$$

In the next *k* steps of the algorithm (*k* ≥ 1):

$$\alpha_{adj}^{(k)} = \alpha_{adj}^{(k-1)} \cdot \frac{\alpha}{\alpha_{act}^{max}(\alpha_{adj}^{(k-1)})}$$

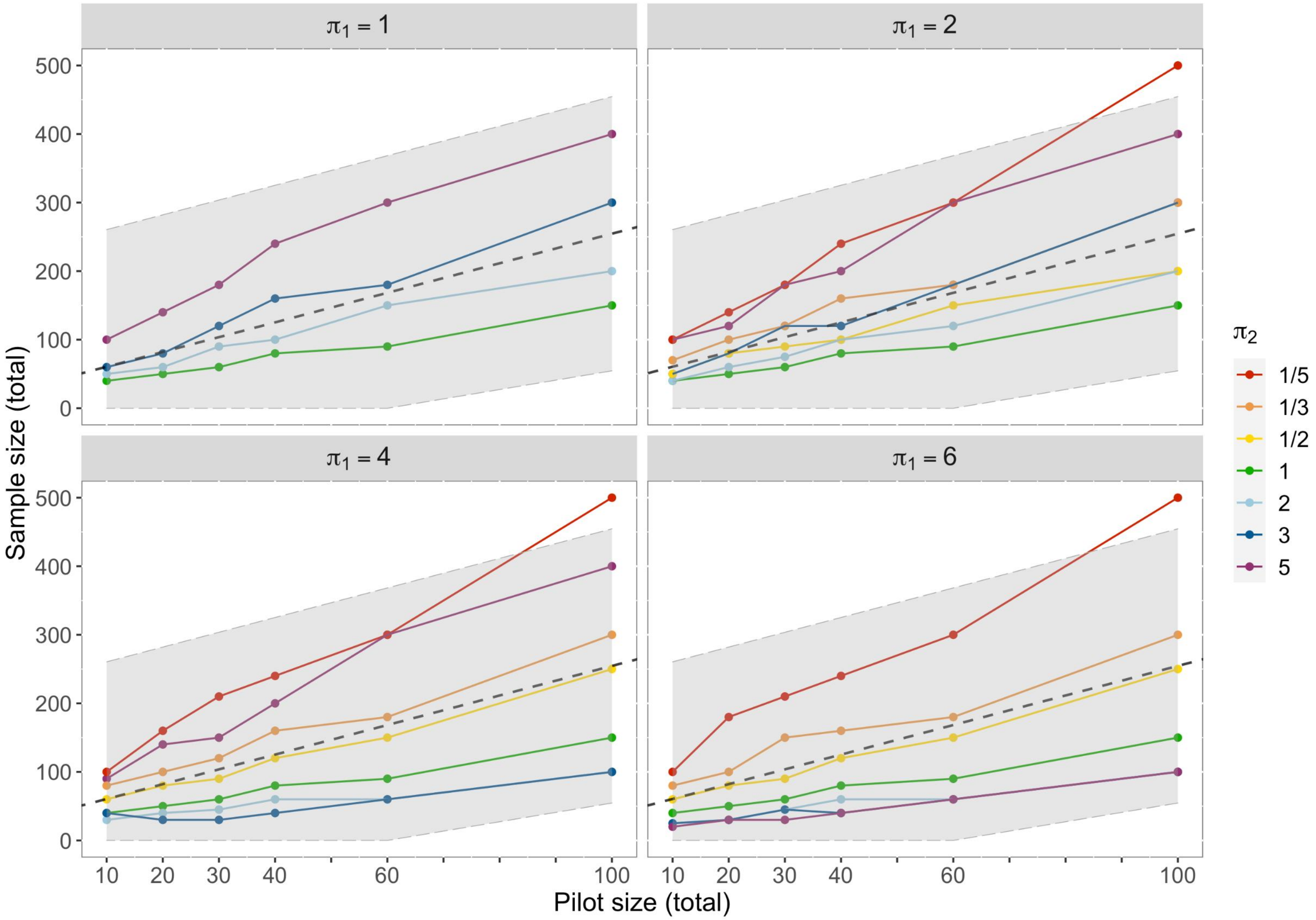
R FUNCTION - VALIDATION





GHENT UNIVERSITY

R FUNCTION – DEFAULT VALUE 'RANGE'



EQUAL VS. UNEQUAL PILOT SIZE

GHENT UNIVERSITY