cascad

**May 17, 2023 - cascad#279**

# Execution Report

## Title: Non-Standard Errors


## Authors: ALBERT J. MENKVELD, ANNA DREBER, FELIX HOLZMEISTER, JUERGEN HUBER, MAGNUS JOHANNESSON, MICHAEL KIRCHLER, SEBASTIAN NEUSÜSS, MICHAEL RAZEN, UTZ WEITZEL et al.



**Full reference:** Menkveld et al. "Non-standard errors". May 3, 2023.


The structure and contents of this execution report provided by **cascad** for the certification are similar to those recommended by the AEA Data Editor.

## 1. DATA DESCRIPTION

This study is based on collected empirical data. 164 research teams from around the world received a dataset of transactions on the EuroStoxx 50 index futures on which they tested the same set of hypotheses. All teams then submitted a "short paper" summarizing their methodology and results. Those results were collected to analyze the variability of empirical finance results.

For a thorough description of the data, please refer to section I of the paper or to https://fincap.academy/.

## 2. CODE DESCRIPTION

The verification materials contain seven subfolders, each of which represents a stage of the study:

- a_create_balanced_sample_and_summary_statistics_plus/
- b_multiverse_1_convert_monthly_eurex_files_to_daily_sample/
- b_multiverse_2_conduct_multiverse_analysis_on_daily_sample/
- c_analysis_1_create_output_per_RT_Hi_for_report/
- c_analysis_2_standard-error-for-median-of-heteroskedastic-variables
- c_analysis_3_create_output_for_all_RT_Hi_for_report/
- d_latex_report/

Except for the latter, they are all divided into four self-explanatory subfolders: "code", "input", "output", and "temp". This structure follows advice from "Code and Data for the Social Sciences: A Practitioner's Guide," by Matthew Gentzkow and Jesse M. Shapiro. The idea is that the researcher opens a terminal, changes into the code directory of the stage at hand, and runs the python script in that code directory. This script then takes input data from the input directory, processes it, writes

temporary results to the temp directory, and writes the results to the output directory. The six folders also contain a bash script that will copy the various files from the output directory to the appropriate input directories of subsequent stages.

The final stage is the latex report, which requires latex to process the tex file.

## 3. VERIFICATION STEPS

The verification package was downloaded from the Dropbox repository on May 3rd and run as per readme, using python 3.8.10 on a computer with 256GB RAM, Intel Xeon Silver 4210R 2.4GHz (40 cores), NVIDIA RTX™ A5000, and Linux (Ubuntu 20.04).

According to the readme, "The multiverse was not rerun when preparing the final code as it requires substantial computing power". Indeed, this part of the code "was run on Snellius, which is a national supercomputer in the Netherlands built for academic use (https://servicedesk.surf.nl/wiki/display/WIKI/Snellius). Snellius dedicated 128 CPUs and an internal memory of 200G when running the code". Since none of our computer have that many CPUs, we did not run these sections of the code, and did not try to generate the results that rely on those multiverse analyses: that is, Panels (b) of Tables II and IA.7.

We received an error message when running "output_for_all_RT_Hi_for_report.py", located in the folder "c_analysis_3_create_output_for_all_RT_Hi_for_report":

"*Traceback (most recent call last):*
*File "__init__.pxd", line 943, in numpy.import_array*
*RuntimeError: module compiled against API version 0x10 but this version of numpy is 0xf . Check the section C-API incompatibility at the Troubleshooting ImportError section at https://numpy.org/devdocs/user/troubleshooting-importerror.html#c-api-incompatibility for indications on how to solve this problem .*

*During handling of the above exception, another exception occurred:*

*Traceback (most recent call last):*
*File "output_for_all_RT_Hi_for_report.py", line 29, in <module>*
*from pyqreg import QuantReg*
*File "/usr/local/lib/python3.8/dist-packages/pyqreg/__init__.py", line 1, in <module>*
*from .formula_api import quantreg*
*File "/usr/local/lib/python3.8/dist-packages/pyqreg/formula_api.py", line 4, in <module>*
*from .quantile_regression import QuantReg*
*File "/usr/local/lib/python3.8/dist-packages/pyqreg/quantile_regression.py", line 6, in <module>*
*from .c.blas_lapack import lapack_cholesky_inv*
*File "src/pyqreg/c/blas_lapack.pyx", line 10, in init pyqreg.c.blas_lapack*
*File "__init__.pxd", line 945, in numpy.import_array*
*ImportError: numpy.core.multiarray failed to import*"

We resolved this issue by running the command "python -m pip uninstall numpy". The script then worked as intended. We assume there was a conflict between two different installed versions of this package.

## 4. REPRODUCED RESULTS

We did not try to reproduce Panels (b) of Tables 2 and IA.7 since they rely on a multiverse analysis that requires a supercomputer.
As shown below, we reproduced all Figures and the rest of the Tables with accuracy.

### 4.1 TABLE I. SUMMARY STATISTICS

Original:

Panel (a): Quality of the #fincap community

| | Research teams | Peer evaluators |
|---|---|---|
| Fraction with top finance/econ publications (see footnote 6) | 0.31 | 0.85 |
| Fraction including at least associate/full professor | 0.52 | 0.88 |
| Experience empirical-finance research (low-high, 1-10) | 8.1 (1.7) | 8.4 (1.8) |
| Experience market-liquidity research (low-high, 1-10) | 6.9 (2.4) | 7.8 (2.3) |
| Relevant experience (average of the above two items) | 7.5 (1.3) | 8.1 (1.7) |
| Fraction with "big data" experience (>#fincap sample) | 0.65 | 0.88 |
| Fraction teams consisting of two members (maximum team size) | 0.79 | |
| Number of observations | 164 | 34 |

Panel (b): Quality of the analysis of research teams

| | Research teams |
|---|---|
| Reproducibility score according to Cascad (low-high, 0-100) | 64.5 (43.7) |
| Paper quality as judged by peer evaluators (low-high, 0-10) | 6.2 (2.0) |

Panel (c): Dispersion across teams of stage-1 results: Estimates, SEs, and $t$-values

| | RT-H1 Efficiency | RT-H2 RSpread | RT-H3 Client Volume | RT-H4 Client RSpread | RT-H5 Client MOrders | RT-H6 Client GTR |
|---|---|---|---|---|---|---|
| *Estimate (yearly change, %)* | | | | | | |
| Mean | 446.3 | -1,093.4 | -3.5 | -38,276.1 | -3.5 | -87.1 |
| SD | 5,817.5 | 14,537.2 | 9.4 | 490,024.2 | 37.6 | 728.5 |
| Min | -171.1 | -186,074.5 | -117.5 | -6,275,383.0 | -452.9 | -8,254.5 |
| Q(0.10) | -23.7 | -6.9 | -3.8 | -6.7 | -1.6 | -192.1 |
| Q(0.25) | -6.2 | -3.6 | -3.5 | -2.1 | -0.6 | -18.2 |
| Median | -1.1 | -0.0 | -3.3 | 0.1 | -0.0 | 0.0 |
| Q(0.75) | 0.5 | 3.9 | -2.4 | 3.8 | 0.2 | 3.2 |
| Q(0.90) | 3.7 | 21.5 | -0.1 | 20.4 | 1.0 | 56.5 |
| black!10 IQR (i.e., NSE) | 6.7 | 7.5 | 1.2 | 5.9 | 0.8 | 21.4 |
| IDR | 27.3 | 28.4 | 3.7 | 27.1 | 2.5 | 248.5 |
| Max | 74,491.1 | 4,124.0 | 8.7 | 870.2 | 69.5 | 1,119.0 |
| *Standard error* | | | | | | |
| Mean | 468.7 | 1,195.3 | 3.7 | 38,302.0 | 6.2 | 148.2 |
| SD | 5,810.6 | 14,711.9 | 29.5 | 489,929.5 | 40.1 | 526.0 |
| Min | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| Q(0.10) | 0.1 | 0.2 | 0.1 | 0.2 | 0.1 | 0.0 |
| Q(0.25) | 0.5 | 1.1 | 0.3 | 1.2 | 0.2 | 0.7 |
| Median | 2.5 | 5.0 | 1.4 | 4.4 | 1.0 | 9.7 |
| Q(0.75) | 9.3 | 13.9 | 2.0 | 14.3 | 2.4 | 77.1 |
| Q(0.90) | 44.7 | 39.6 | 2.2 | 31.2 | 3.1 | 235.4 |
| IQR | 8.8 | 12.8 | 1.7 | 13.1 | 2.2 | 76.4 |
| IDR | 44.6 | 39.4 | 2.1 | 31.0 | 3.1 | 235.4 |
| Max | 74,425.5 | 188,404.1 | 378.8 | 6,274,203.0 | 463.7 | 4,836.2 |
| *t-value* | | | | | | |
| Mean | -3.6 | 35.3 | -47.1 | 24.3 | -5.7 | -2.0 |
| SD | 28.4 | 541.2 | 269.9 | 406.0 | 60.1 | 21.2 |
| Min | -322.3 | -764.6 | -2,770.6 | -852.6 | -631.6 | -191.7 |
| Q(0.10) | -4.7 | -5.7 | -37.4 | -3.5 | -2.3 | -1.7 |
| Q(0.25) | -1.9 | -1.5 | -11.5 | -1.0 | -0.6 | -1.0 |
| Median | -0.7 | -0.1 | -1.8 | 0.1 | 0.0 | 0.0 |
| Q(0.75) | 0.3 | 0.8 | -1.6 | 1.0 | 0.8 | 0.7 |
| Q(0.90) | 1.7 | 1.5 | -0.3 | 1.6 | 1.7 | 1.2 |
| IQR | 2.2 | 2.3 | 9.9 | 1.9 | 1.3 | 1.7 |
| IDR | 6.4 | 7.2 | 37.1 | 5.2 | 3.9 | 2.9 |
| Max | 51.6 | 6,880.5 | 29.5 | 5,119.5 | 89.6 | 100.6 |

Reproduced:

Panel (a): Quality of the #fincap community

| | Research teams | Peer evaluators |
|---|---|---|
| Fraction with top finance/econ publications (see footnote 6) | 0.31 | 0.85 |
| Fraction including at least associate/full professor | 0.52 | 0.88 |
| Experience empirical-finance research (low-high, 1-10) | 8.1 (1.7) | 8.4 (1.8) |
| Experience market-liquidity research (low-high, 1-10) | 6.9 (2.4) | 7.8 (2.3) |
| Relevant experience (average of the above two items) | 7.5 (1.3) | 8.1 (1.7) |
| Fraction with "big data" experience (>#fincap sample) | 0.65 | 0.88 |
| Fraction teams consisting of two members (maximum team size) | 0.79 | |
| Number of observations | 164 | 34 |

Panel (b): Quality of the analysis of research teams

| | Research teams |
|---|---|
| Reproducibility score according to Cascad (low-high, 0-100) | 64.5 (43.7) |
| Paper quality as judged by peer evaluators (low-high, 0-10) | 6.2 (2.0) |

| | RT-H1 Efficiency | RT-H2 RSpread | RT-H3 Client Volume | RT-H4 Client RSpread | RT-H5 Client MOrders | RT-H6 Client GTR |
|---|---|---|---|---|---|---|
| *Estimate (yearly change, %)* | | | | | | |
| Mean | 446.3 | -1,093.4 | -3.5 | -38,276.1 | -3.5 | -87.1 |
| SD | 5,817.5 | 14,537.2 | 9.4 | 490,024.2 | 37.6 | 728.5 |
| Min | -171.1 | -186,074.5 | -117.5 | -6,275,383.0 | -452.9 | -8,254.5 |
| Q(0.10) | -23.7 | -6.9 | -3.8 | -6.7 | -1.6 | -192.1 |
| Q(0.25) | -6.2 | -3.6 | -3.5 | -2.1 | -0.6 | -18.2 |
| Median | -1.1 | -0.0 | -3.3 | 0.1 | -0.0 | 0.0 |
| Q(0.75) | 0.5 | 3.9 | -2.4 | 3.8 | 0.2 | 3.2 |
| Q(0.90) | 3.7 | 21.5 | -0.1 | 20.4 | 1.0 | 56.5 |
| black!10 IQR (i.e., NSE) | 6.7 | 7.5 | 1.2 | 5.9 | 0.8 | 21.4 |
| IDR | 27.3 | 28.4 | 3.7 | 27.1 | 2.5 | 248.5 |
| Max | 74,491.1 | 4,124.0 | 8.7 | 870.2 | 69.5 | 1,119.0 |
| *Standard error* | | | | | | |
| Mean | 468.7 | 1,195.3 | 3.7 | 38,302.0 | 6.2 | 148.2 |
| SD | 5,810.6 | 14,711.9 | 29.5 | 489,929.5 | 40.1 | 526.0 |
| Min | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| Q(0.10) | 0.1 | 0.2 | 0.1 | 0.2 | 0.1 | 0.0 |
| Q(0.25) | 0.5 | 1.1 | 0.3 | 1.2 | 0.2 | 0.7 |
| Median | 2.5 | 5.0 | 1.4 | 4.4 | 1.0 | 9.7 |
| Q(0.75) | 9.3 | 13.9 | 2.0 | 14.3 | 2.4 | 77.1 |
| Q(0.90) | 44.7 | 39.6 | 2.2 | 31.2 | 3.1 | 235.4 |
| IQR | 8.8 | 12.8 | 1.7 | 13.1 | 2.2 | 76.4 |
| IDR | 44.6 | 39.4 | 2.1 | 31.0 | 3.1 | 235.4 |
| Max | 74,425.5 | 188,404.1 | 378.8 | 6,274,203.0 | 463.7 | 4,836.2 |
| *t-value* | | | | | | |
| Mean | -3.6 | 35.3 | -47.1 | 24.3 | -5.7 | -2.0 |
| SD | 28.4 | 541.2 | 269.9 | 406.0 | 60.1 | 21.2 |
| Min | -322.3 | -764.6 | -2,770.6 | -852.6 | -631.6 | -191.7 |
| Q(0.10) | -4.7 | -5.7 | -37.4 | -3.5 | -2.3 | -1.7 |
| Q(0.25) | -1.9 | -1.5 | -11.5 | -1.0 | -0.6 | -1.0 |
| Median | -0.7 | -0.1 | -1.8 | 0.1 | 0.0 | 0.0 |
| Q(0.75) | 0.3 | 0.8 | -1.6 | 1.0 | 0.8 | 0.7 |
| Q(0.90) | 1.7 | 1.5 | -0.3 | 1.6 | 1.7 | 1.2 |
| IQR | 2.2 | 2.3 | 9.9 | 1.9 | 1.3 | 1.7 |
| IDR | 6.4 | 7.2 | 37.1 | 5.2 | 3.9 | 2.9 |
| Max | 51.6 | 6,880.5 | 29.5 | 5,119.5 | 89.6 | 100.6 |

## 4.2. TABLE II. NON-STANDARD ERROR TEST (PANEL A ONLY)

Original:

Panel (a): Multiple tests (Bonferroni)

| | Reject no-NSE at 0.5%? | $p$-value of family test | Mean (SD) correlation test statistics | Effective number of tests |
|---|---|---|---|---|
| RT-H1 | Yes (8, 25) | $<0.0001$ | 0.00 (0.00) | 164 |
| RT-H2 | Yes (24, 10) | $<0.0001$ | 0.00 (0.00) | 164 |
| RT-H3 | Yes (13, 25) | $<0.0001$ | 0.00 (0.00) | 164 |
| RT-H4 | Yes (22, 4) | $<0.0001$ | 0.00 (0.00) | 164 |
| RT-H5 | Yes (13, 10) | $<0.0001$ | 0.00 (0.00) | 164 |
| RT-H6 | Yes (8, 3) | $<0.0001$ | 0.00 (0.00) | 164 |

Reproduced:

Panel (a): Multiple tests (Bonferroni)

| | Reject no-NSE at 0.5%? | $p$-value of family test | Mean (SD) correlation test statistics | Effective number of tests |
|---|---|---|---|---|
| RT-H1 | Yes (8, 25) | $<0.0001$ | 0.00 (0.00) | 164 |
| RT-H2 | Yes (24, 10) | $<0.0001$ | 0.00 (0.00) | 164 |
| RT-H3 | Yes (13, 25) | $<0.0001$ | 0.00 (0.00) | 164 |
| RT-H4 | Yes (22, 4) | $<0.0001$ | 0.00 (0.00) | 164 |
| RT-H5 | Yes (13, 10) | $<0.0001$ | 0.00 (0.00) | 164 |
| RT-H6 | Yes (8, 3) | $<0.0001$ | 0.00 (0.00) | 164 |

## 4.3. TABLE III. STAGE-1 QUANTILE REGRESSIONS

Original:

| | Q(0.10) | Q(0.25) | Q(0.50) | Q(0.75) | Q(0.90) |
|---|---|---|---|---|---|
| Team quality (standardized/scaled) | 0.597** (0.030) | 0.004 (0.014) | 0.002 (0.007) | 0.032** (0.012) | -0.325** (0.030) |
| Reproducibility score (standardized/scaled) | 0.473** (0.033) | 0.109** (0.014) | -0.001 (0.007) | -0.142** (0.011) | -0.555** (0.028) |
| Average rating (standardized/scaled) | 0.766** (0.034) | 0.230** (0.014) | -0.001 (0.007) | -0.097** (0.011) | -0.626** (0.028) |
| Dummy RT-H1 Efficiency | -29.592** (0.813) | -6.099** (0.340) | -1.132** (0.166) | 0.939** (0.269) | 9.057** (0.708) |
| Dummy RT-H2 RSpread | -15.933** (0.849) | -3.930** (0.342) | -0.017 (0.166) | 3.674** (0.268) | 22.451** (0.705) |
| Dummy RT-H3 Client Volume | -5.629** (0.836) | -3.789** (0.339) | -3.319** (0.166) | -2.386** (0.268) | 0.221 (0.721) |
| Dummy RT-H4 Client RSpread | -12.089** (0.837) | -2.437** (0.340) | 0.162 (0.166) | 4.161** (0.266) | 19.619** (0.704) |
| Dummy RT-H5 Client MOrders | -2.479** (0.837) | -0.744* (0.339) | -0.001 (0.166) | 0.297 (0.268) | 1.625* (0.721) |
| Dummy RT-H6 GTR | -194.457** (0.806) | -21.385** (0.337) | 0.022 (0.167) | 5.137** (0.268) | 65.203** (0.679) |
| #Observations | 984 | 984 | 984 | 984 | 984 |

Reproduced:

| | Q(0.10) | Q(0.25) | Q(0.50) | Q(0.75) | Q(0.90) |
|---|---|---|---|---|---|
| Team quality (standardized/scaled) | 0.597** (0.030) | 0.004 (0.014) | 0.002 (0.007) | 0.032** (0.012) | -0.325** (0.030) |
| Reproducibility score (standardized/scaled) | 0.473** (0.033) | 0.109** (0.014) | -0.001 (0.007) | -0.142** (0.011) | -0.555** (0.028) |
| Average rating (standardized/scaled) | 0.766** (0.034) | 0.230** (0.014) | -0.001 (0.007) | -0.097** (0.011) | -0.626** (0.028) |
| Dummy RT-H1 Efficiency | -29.592** (0.813) | -6.099** (0.340) | -1.132** (0.166) | 0.939** (0.269) | 9.057** (0.727) |
| Dummy RT-H2 RSpread | -15.933** (0.849) | -3.930** (0.342) | -0.017 (0.166) | 3.674** (0.268) | 22.451** (0.705) |
| Dummy RT-H3 Client Volume | -5.629** (0.836) | -3.789** (0.339) | -3.319** (0.166) | -2.386** (0.268) | 0.221 (0.702) |
| Dummy RT-H4 Client RSpread | -12.089** (0.838) | -2.437** (0.340) | 0.162 (0.166) | 4.161** (0.266) | 19.619** (0.704) |
| Dummy RT-H5 Client MOrders | -2.479** (0.837) | -0.744* (0.339) | -0.001 (0.166) | 0.297 (0.268) | 1.625* (0.702) |
| Dummy RT-H6 GTR | -194.457** (0.806) | -21.385** (0.337) | 0.022 (0.167) | 5.137** (0.268) | 65.203** (0.679) |
| #Observations | 984 | 984 | 984 | 984 | 984 |

## 4.4. TABLE IV. ALL-STAGES QUANTILE REGRESSIONS

Original:

| | Q(0.10) | Q(0.25) | Q(0.50) | Q(0.75) | Q(0.90) |
|---|---|---|---|---|---|
| Dummy Stage 2 - Dummy Stage 1 | 2.44* | 0.07 | -0.00 | -0.06 | -0.73 |
| | (1.18) | (0.14) | (0.01) | (0.06) | (0.64) |
| Dummy Stage 3 - Dummy Stage 2 | 0.94* | 0.15 | 0.00 | -0.09 | -0.73 |
| | (0.41) | (0.09) | (0.01) | (0.05) | (0.40) |
| Dummy Stage 4 - Dummy Stage 3 | 0.21* | 0.06* | 0.00 | -0.04 | -0.25* |
| | (0.09) | (0.03) | (0.01) | (0.03) | (0.11) |
| Dummy Stage 4 - Dummy Stage 1 | 3.59** | 0.28* | -0.00 | -0.19** | -1.71** |
| | (1.23) | (0.14) | (0.01) | (0.05) | (0.50) |
| RT-hypotheses dummies | Yes | Yes | Yes | Yes | Yes |
| #Observations | 3,936 | 3,936 | 3,936 | 3,936 | 3,936 |

Reproduced:

| | Q(0.10) | Q(0.25) | Q(0.50) | Q(0.75) | Q(0.90) |
|---|---|---|---|---|---|
| Dummy Stage 2 - Dummy Stage 1 | 2.44* | 0.07 | -0.00 | -0.06 | -0.73 |
| | (1.20) | (0.14) | (0.01) | (0.06) | (0.64) |
| Dummy Stage 3 - Dummy Stage 2 | 0.94* | 0.15 | 0.00 | -0.09 | -0.73 |
| | (0.41) | (0.09) | (0.01) | (0.05) | (0.40) |
| Dummy Stage 4 - Dummy Stage 3 | 0.21* | 0.06* | 0.00 | -0.04 | -0.25* |
| | (0.09) | (0.03) | (0.01) | (0.03) | (0.11) |
| Dummy Stage 4 - Dummy Stage 1 | 3.59** | 0.28* | -0.00 | -0.19** | -1.71** |
| | (1.24) | (0.14) | (0.01) | (0.05) | (0.51) |
| RT-hypotheses dummies | Yes | Yes | Yes | Yes | Yes |
| #Observations | 3,936 | 3,936 | 3,936 | 3,936 | 3,936 |

## 4.5. TABLE V. ANALYSIS PATHS

Original:

| RT-hypothesis | Fork | Fork description | Alternatives | Frequency |
|---|---|---|---|---|
| All | 1 | Remove open/close | No | 79% |
| | | | Yes, 30 minutes | 21% |
| All | 2 | Days excluded | None | 81% |
| | | | Settlement weeks | 19% |
| All | 3 | Outlier treatment | None | 65% |
| | | | Winsorize measure at 2.5 and 97.5 percentile[a] | 20% |
| | | | Trim measure at 2.5 and 97.5 percentile[a] | 14% |
| All | 4 | Frequency analysis | Daily | 37% |
| | | | Weekly | 1% |
| | | | Monthly | 21% |
| | | | Annual | 41% |
| All | 5 | Model | Trend stationary (regresion with linear trend) | 35% |
| | | | Log difference (trivial regression, i.e., intercept only) | 5% |
| | | | Relative difference (trivial regression) | 60% |
| 1 | 6 | Measure | Variance ratio (low-frequency in numerator) | 63% |
| | | | Autocorrelation ($R^2$ of AR model for returns) | 37% |
| 1 | 7 | Measure frequencies | Second to minute | 18% |
| | | | One to five minutes | 26% |
| | | | Five to thirty minutes | 34% |
| | | | Day to week | 13% |
| | | | Day to month | 10% |
| 2,4,5 | 6 | Tick test or aggressor flag | Aggressor flag (available only for part of the sample) | 84% |
| | | | Tick test | 16% |
| 2,4 | 7 | Post-trade value | Price 5 minutes after trade | 81% |
| | | | Price 10 minutes after trade | 6% |
| | | | Price 30 minutes after trade | 13% |
| 2,4 | 8 | Aggregation | Equal-weighted average | 47% |
| | | | Trade-size-weighted average | 53% |
| 3 | 6 | Units... | Volume expressed in #contracts | 70% |
| | | | Volume expressed in euro | 30% |
| 6 | 6 | Reference price | Last trade price in the day | 62% |
| | | | Last trade price one day later | 1% |
| | | | Volume-weighted-average-price (VWAP) full-day | 24% |
| | | | VWAP based on last five trades in the day | 0% |
| 6 | 7 | Mean or median | Mean | 96% |
| | | | Median | 4% |
| 6 | 8 | Handle non-negatives | Translate and transform ($\varepsilon = 0.001$) | 14% |
| | | | Translate and transform ($\varepsilon = 1$) | 7% |
| | | | Set to missing | 79% |
| 6 | 9 | Retain negative-trend sign | Yes | 79% |
| | | | No | 21% |

[a]: Winsorization is applied at the frequency of analysis (fork 4).

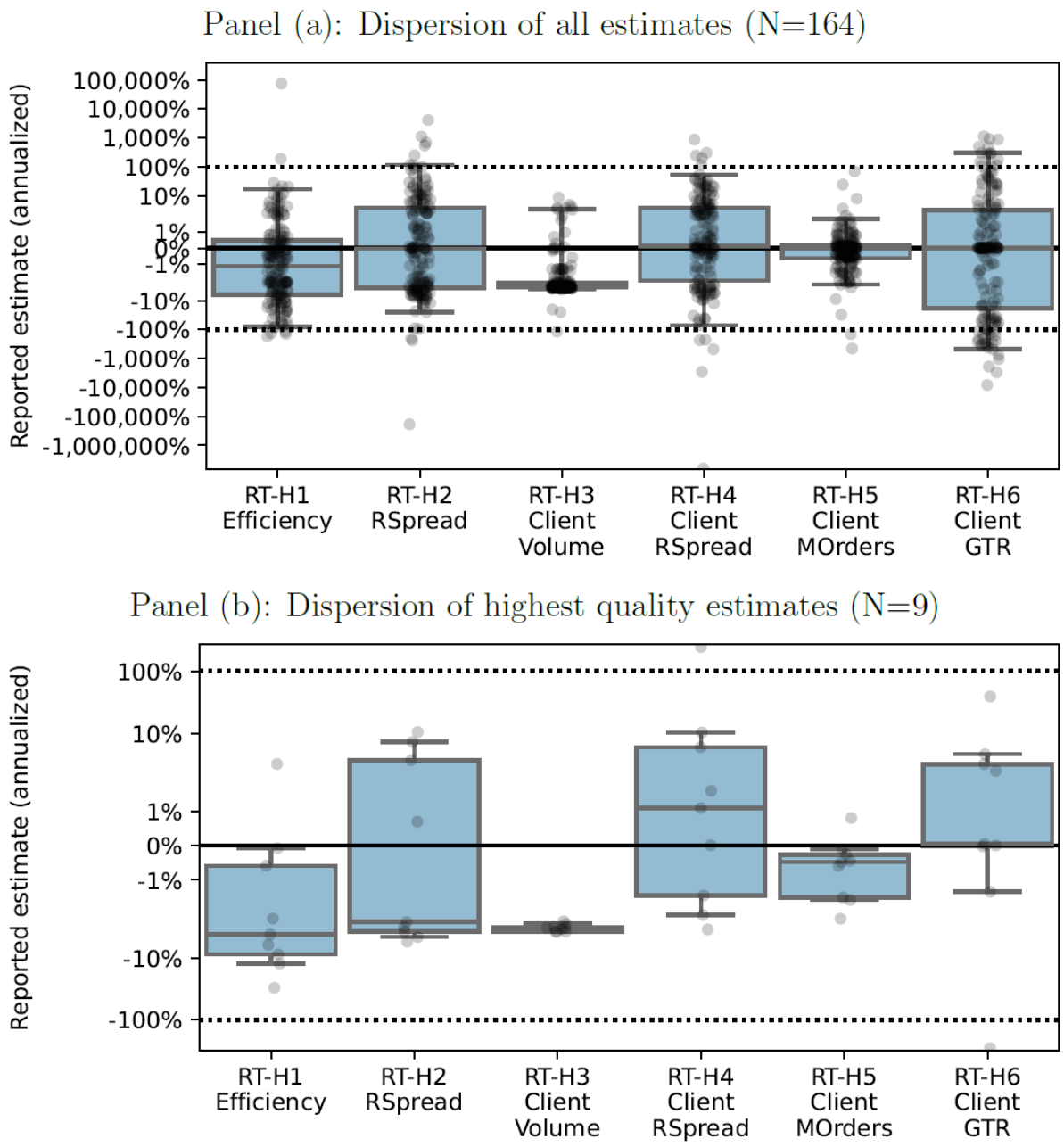
Reproduced:

| RT-hypothesis | Fork | Fork description | Alternatives | Frequency |
|---|---|---|---|---|
| All | 1 | Remove open/close | No | 79% |
| | | | Yes, 30 minutes | 21% |
| All | 2 | Days excluded | None | 81% |
| | | | Settlement weeks | 19% |
| All | 3 | Outlier treatment | None | 65% |
| | | | Winsorize measure at 2.5 and 97.5 percentile[a] | 20% |
| | | | Trim measure at 2.5 and 97.5 percentile[a] | 14% |
| All | 4 | Frequency analysis | Daily | 37% |
| | | | Weekly | 1% |
| | | | Monthly | 21% |
| | | | Annual | 41% |
| All | 5 | Model | Trend stationary (regresion with linear trend) | 35% |
| | | | Log difference (trivial regression, i.e., intercept only) | 5% |
| | | | Relative difference (trivial regression) | 60% |
| 1 | 6 | Measure | Variance ratio (low-frequency in numerator) | 63% |
| | | | Autocorrelation ($R^2$ of AR model for returns) | 37% |
| 1 | 7 | Measure frequencies | Second to minute | 18% |
| | | | One to five minutes | 26% |
| | | | Five to thirty minutes | 34% |
| | | | Day to week | 13% |
| | | | Day to month | 10% |
| 2,4,5 | 6 | Tick test or aggressor flag | Aggressor flag (available only for part of the sample) | 84% |
| | | | Tick test | 16% |
| 2,4 | 7 | Post-trade value | Price 5 minutes after trade | 81% |
| | | | Price 10 minutes after trade | 6% |
| | | | Price 30 minutes after trade | 13% |
| 2,4 | 8 | Aggregation | Equal-weighted average | 47% |
| | | | Trade-size-weighted average | 53% |
| 3 | 6 | Units... | Volume expressed in #contracts | 70% |
| | | | Volume expressed in euro | 30% |
| 6 | 6 | Reference price | Last trade price in the day | 62% |
| | | | Last trade price one day later | 1% |
| | | | Volume-weighted-average-price (VWAP) full-day | 24% |
| | | | VWAP based on last five trades in the day | 0% |
| 6 | 7 | Mean or median | Mean | 96% |
| | | | Median | 4% |
| 6 | 8 | Handle non-negatives | Translate and transform ($\varepsilon = 0.001$) | 14% |
| | | | Translate and transform ($\varepsilon = 1$) | 7% |
| | | | Set to missing | 79% |
| 6 | 9 | Retain negative-trend sign | Yes | 79% |
| | | | No | 21% |

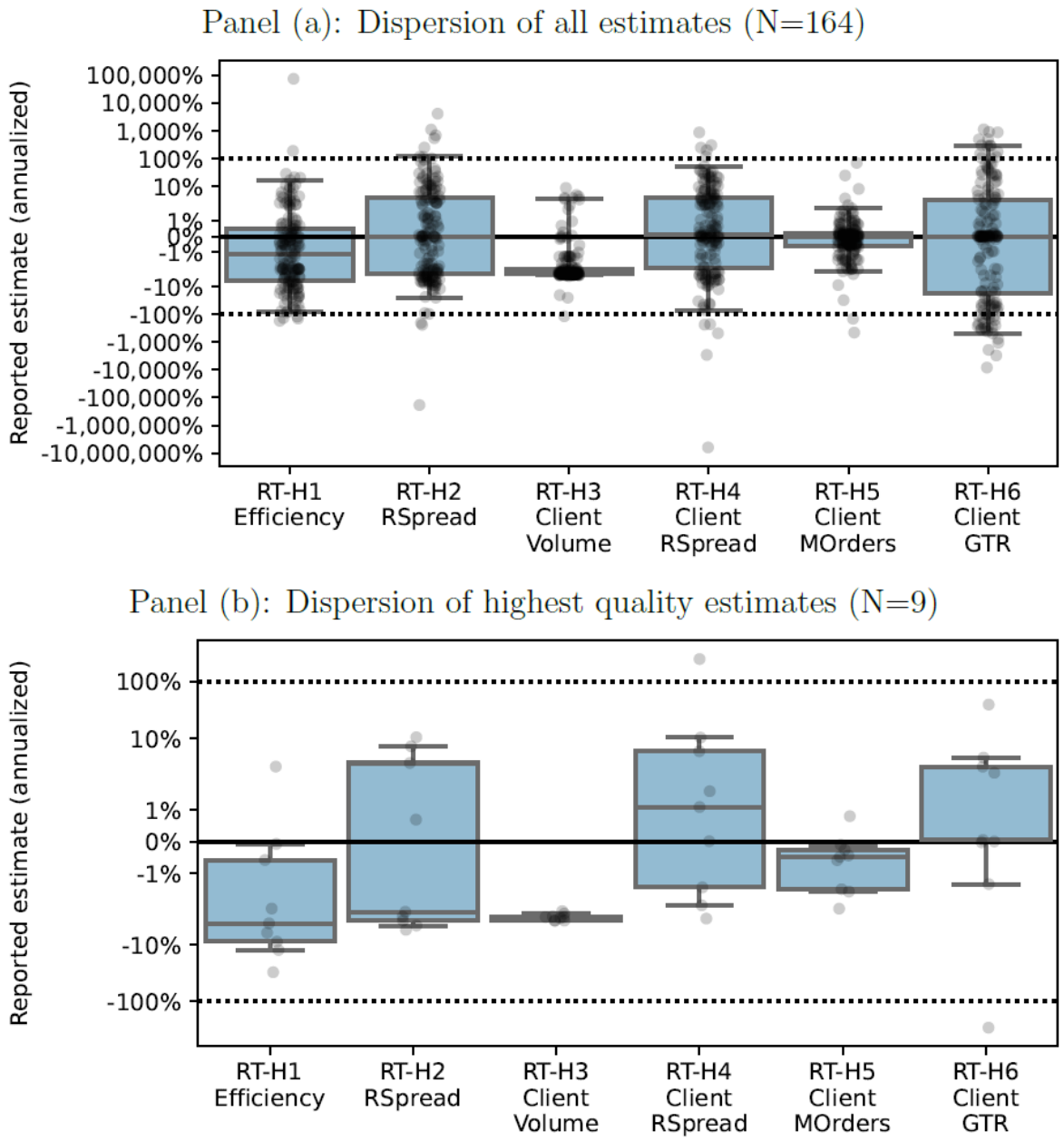
[a]: Winsorization is applied at the frequency of analysis (fork 4).

## 4.6. FIGURE 1. DISPERSION OF STAGE-1 ESTIMATES ACROSS RESEARCH TEAMS

Original:

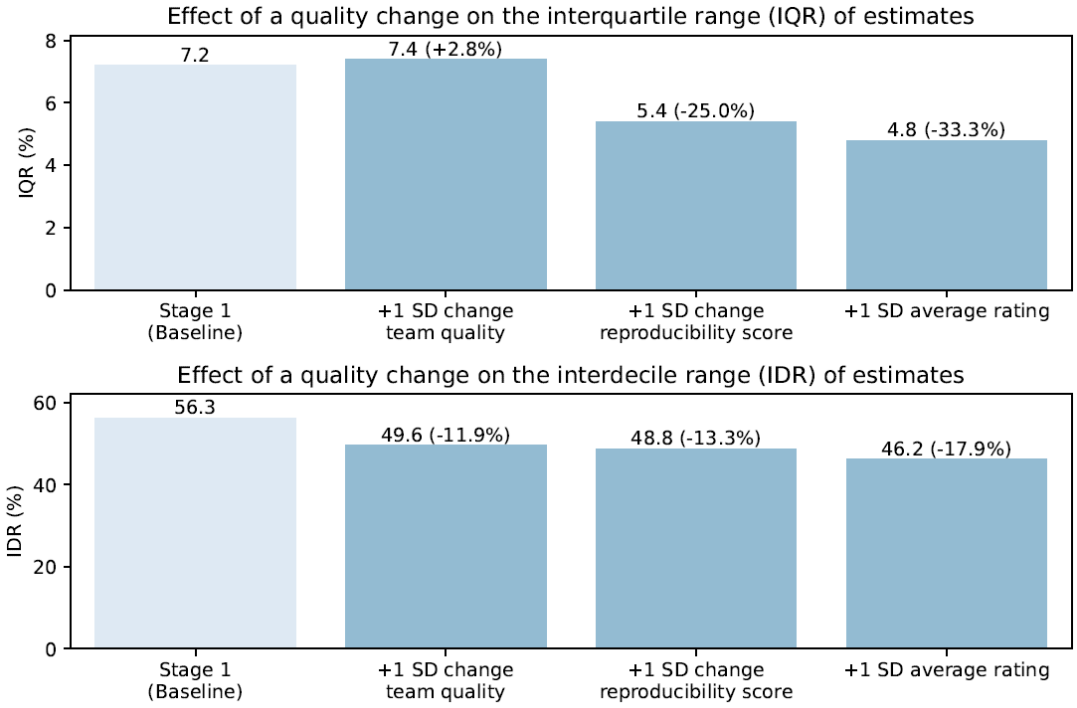


Reproduced:

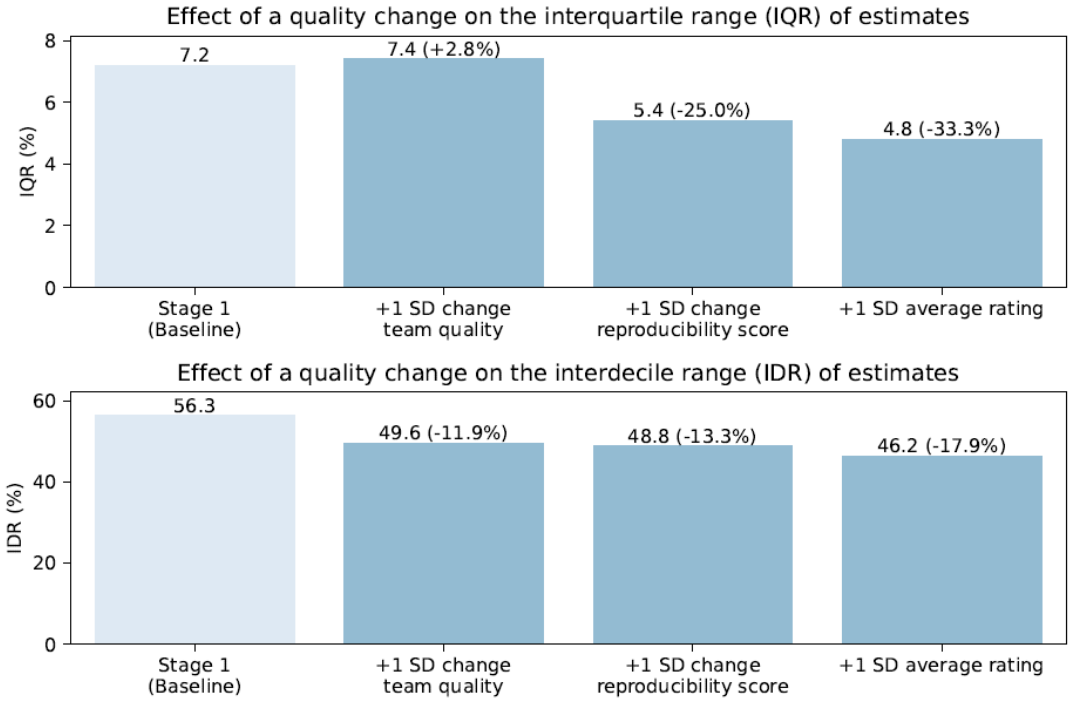

Panel (a): Dispersion of all estimates (N=164)
Reported estimate (annualized)
100,000%
10,000%
1,000%
100%
10%
1%
0%
-1%
-10%
-100%
-1,000%
-10,000%
-100,000%
-1,000,000%
-10,000,000%
RT-H1 Efficiency
RT-H2 RSpread
RT-H3 Client Volume
RT-H4 Client RSpread
RT-H5 Client MOrders
RT-H6 Client GTR
Panel (b): Dispersion of highest quality estimates (N=9)
Reported estimate (annualized)
100%
10%
1%
0%
-1%
-10%
-100%
RT-H1 Efficiency
RT-H2 RSpread
RT-H3 Client Volume
RT-H4 Client RSpread
RT-H5 Client MOrders
RT-H6 Client GTR

## 4.7. FIGURE 2. DISPERSION IN ESTIMATES RELATED TO QUALITY MEASURES

Original:

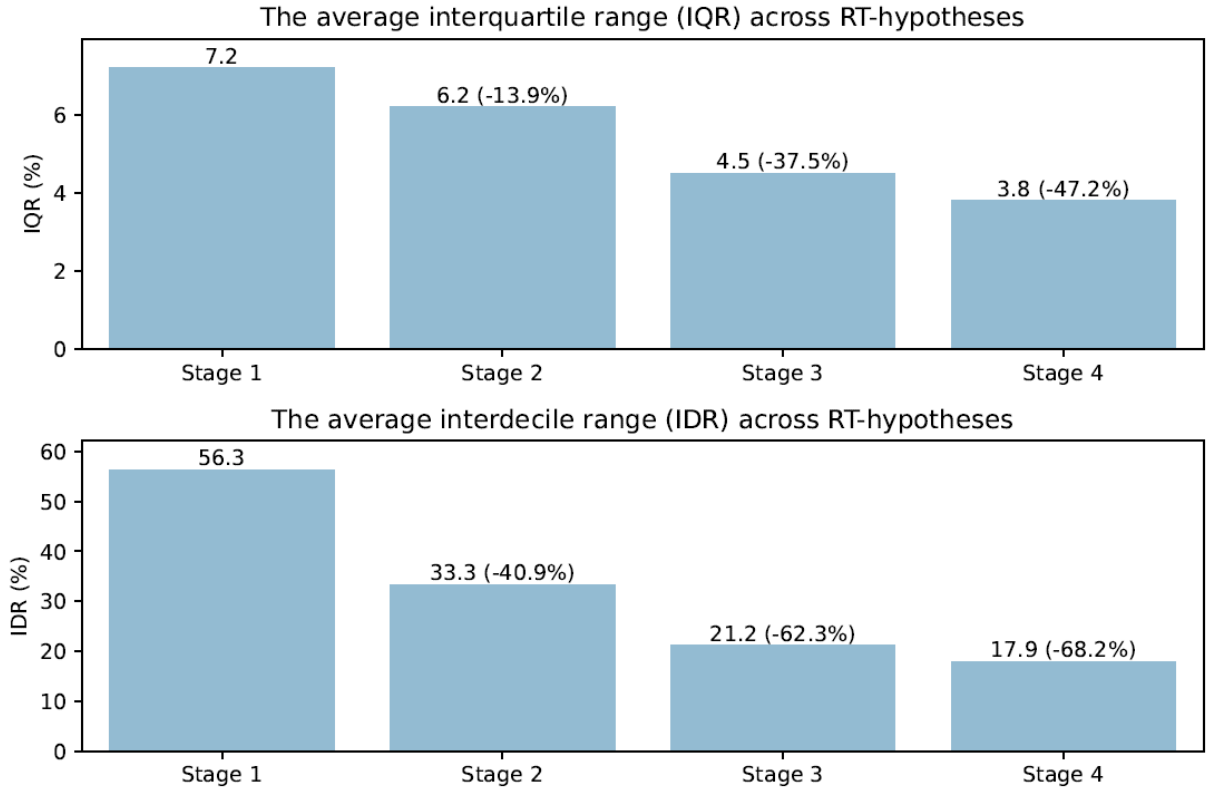



Reproduced:

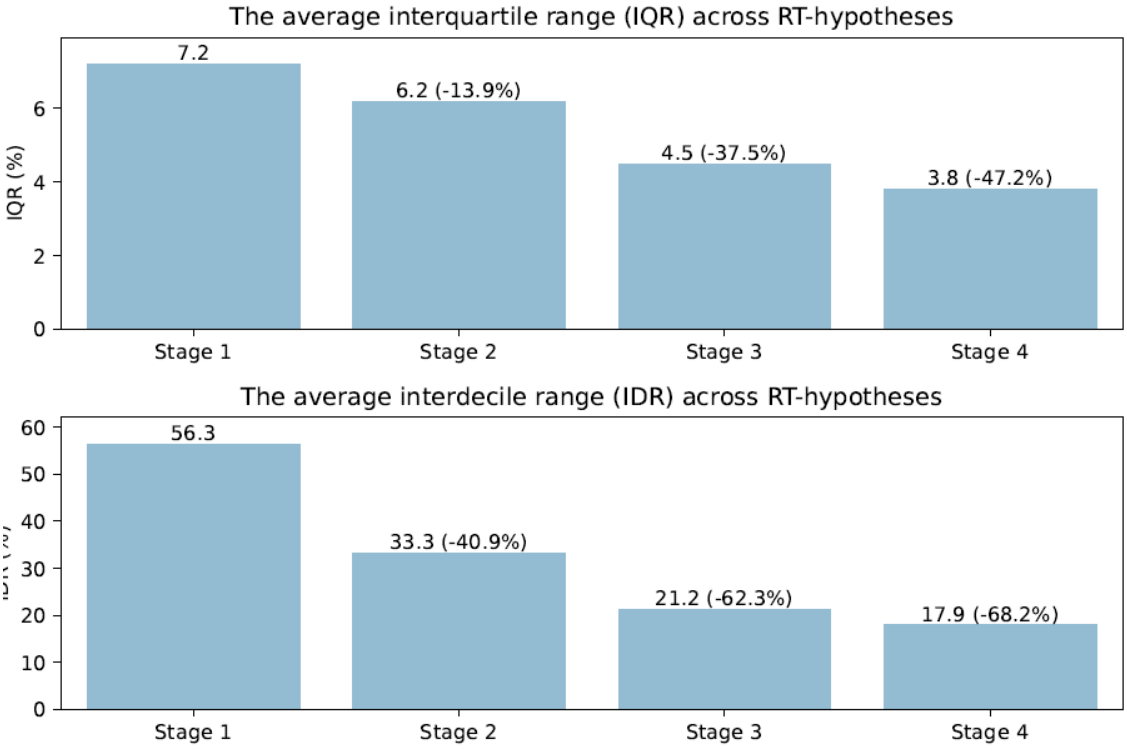


## 4.8. FIGURE 3. DISPERSION IN ESTIMATES RELATED TO FEEDBACK STAGES

Original:

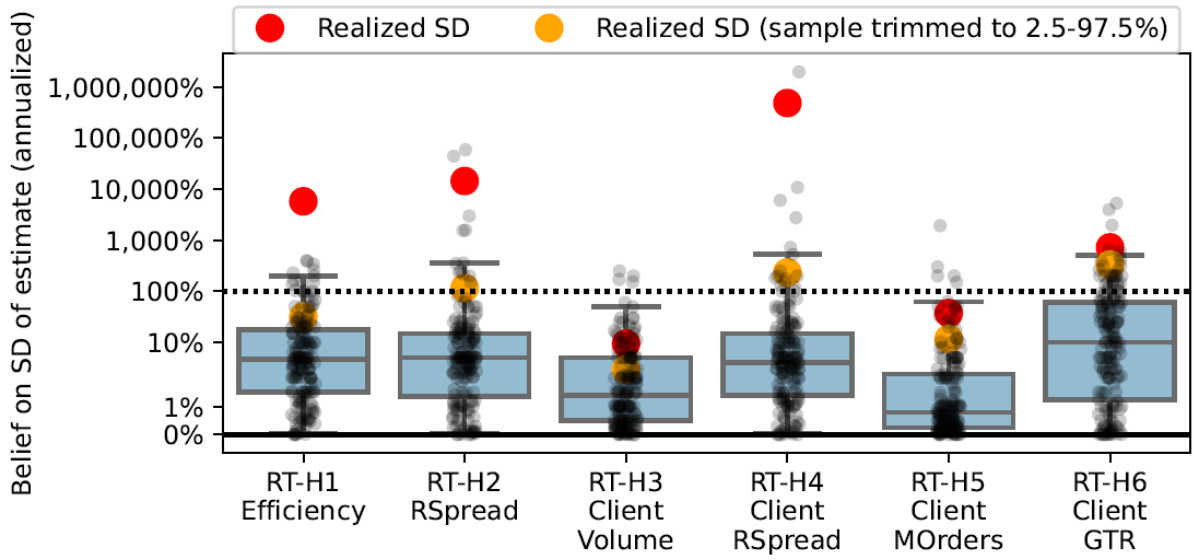



Reproduced:

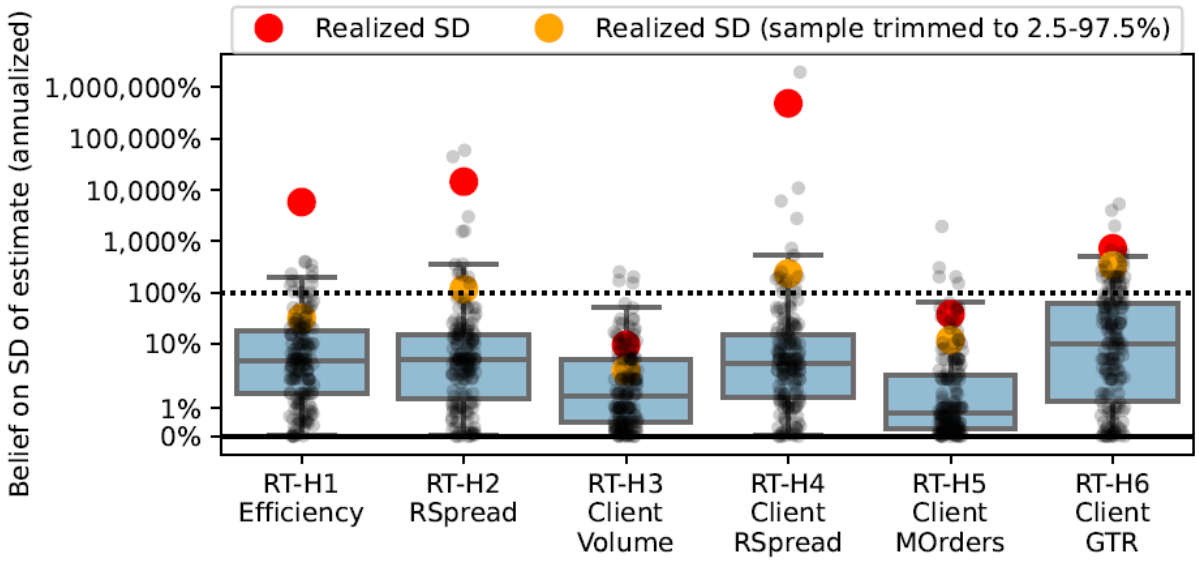


## 4.9. FIGURE 4. RESEARCH TEAM BELIEFS ON DISPERSION STAGE-1 ESTIMATES

Original:

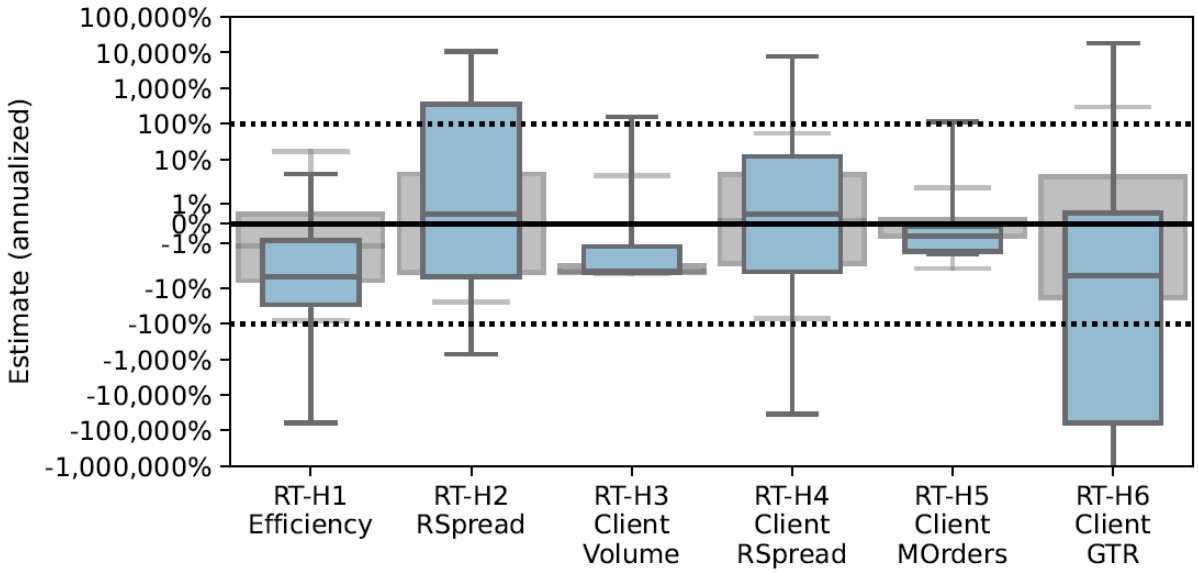



Reproduced:

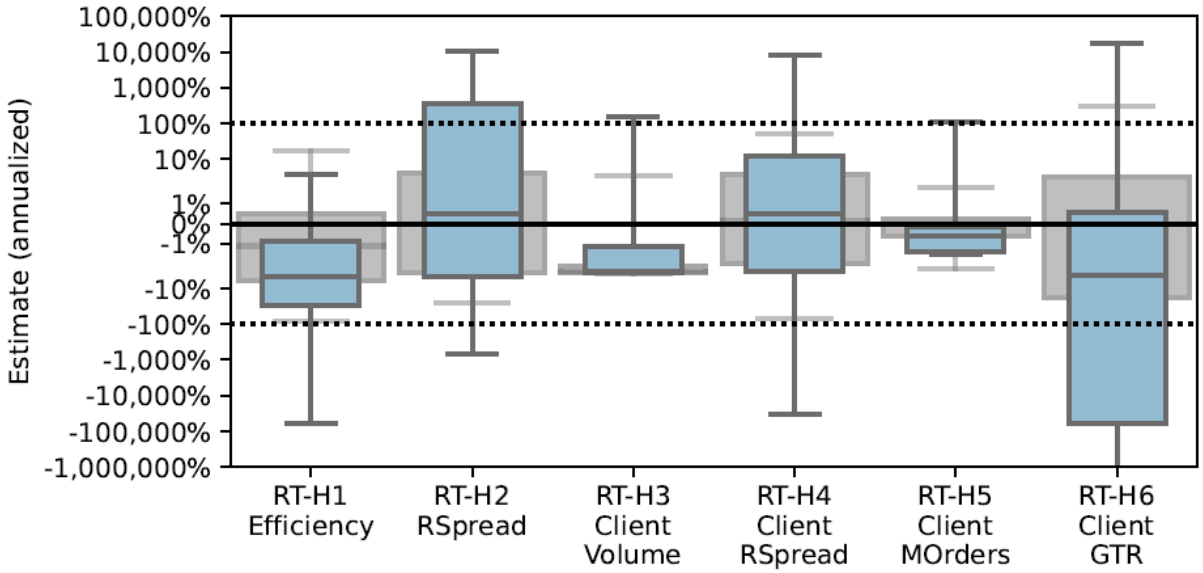


## 4.10. FIGURE 5. DISPERSION IN STAGE-1 ESTIMATES OF MULTIVERSE ANALYSIS

Original:

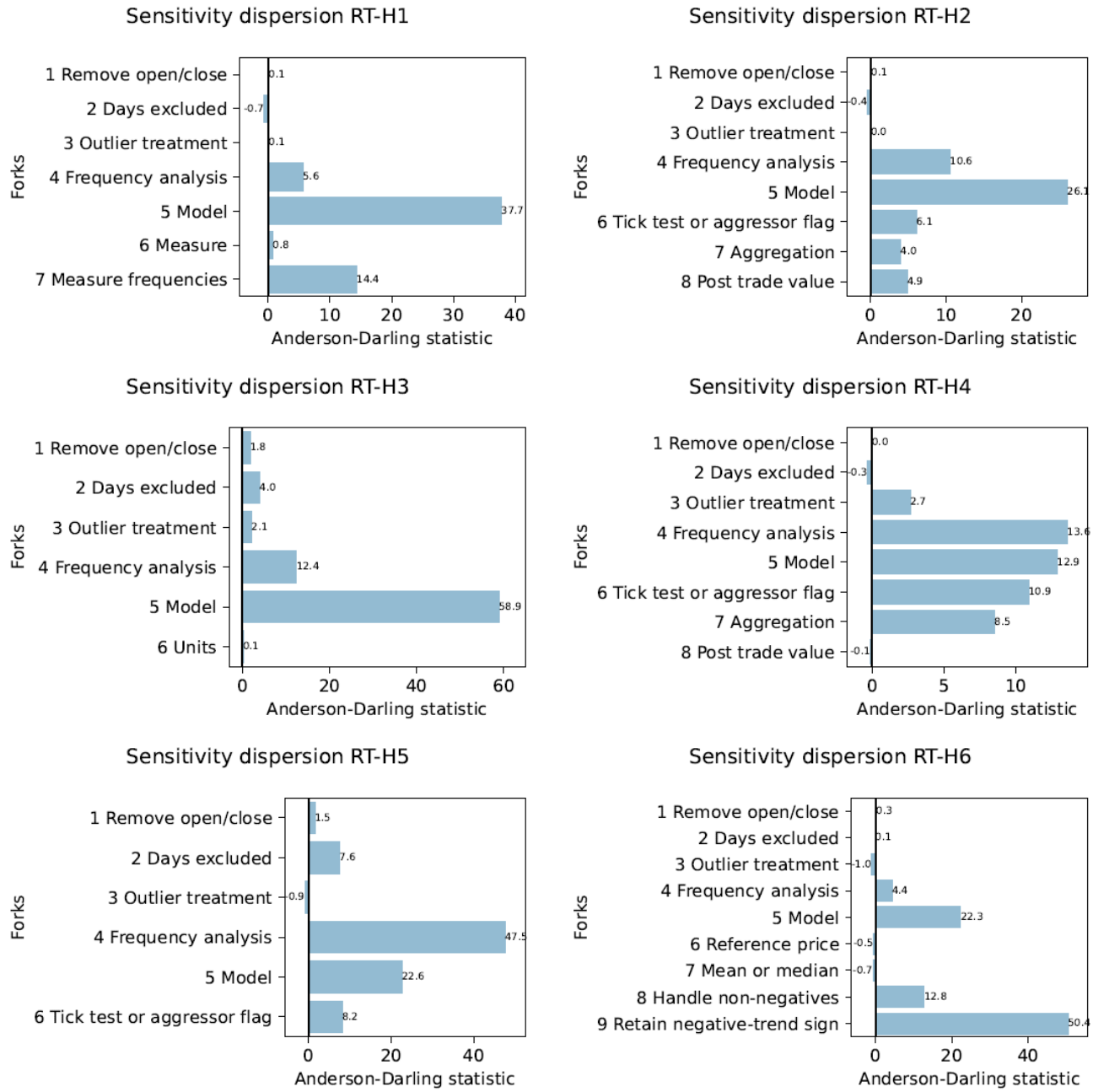



Reproduced:

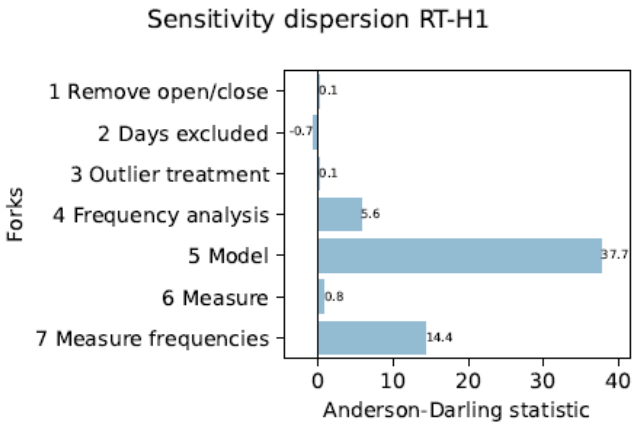


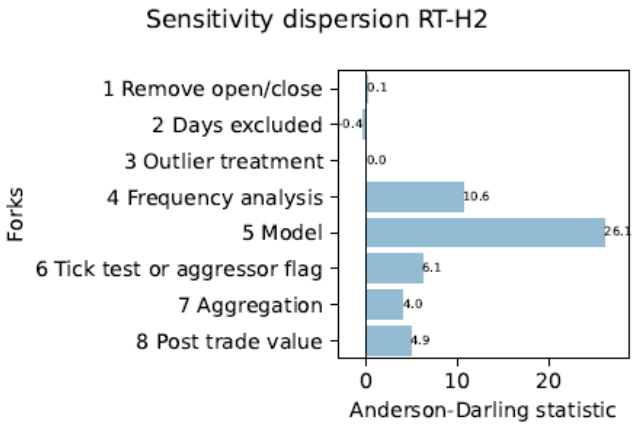
## 4.11. FIGURE 6. FORK SENSITIVITY OF ESTIMATES IN MULTIVERSE ANALYSIS

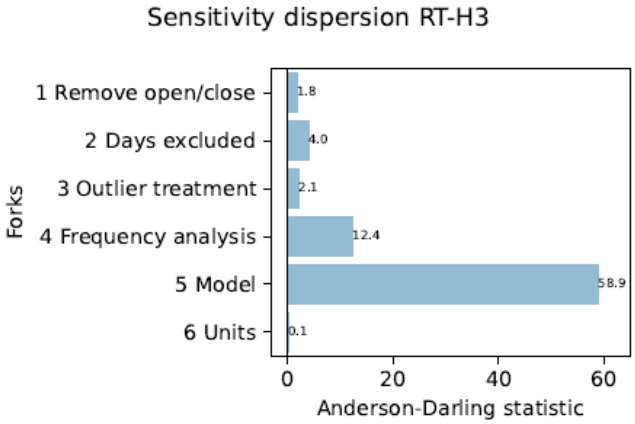
Original:

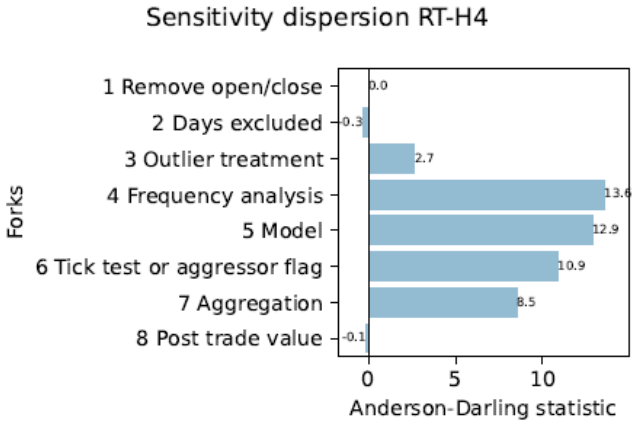


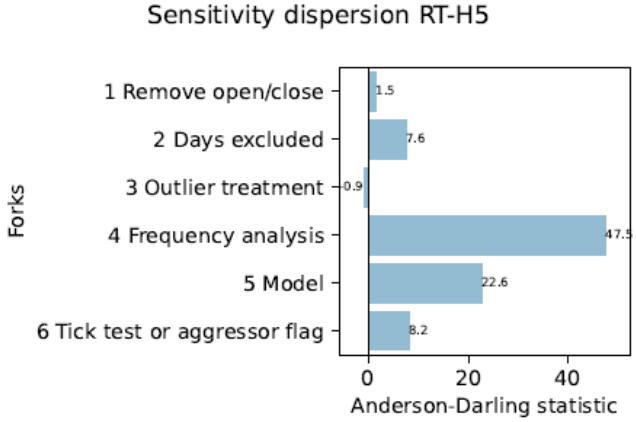
Reproduced:

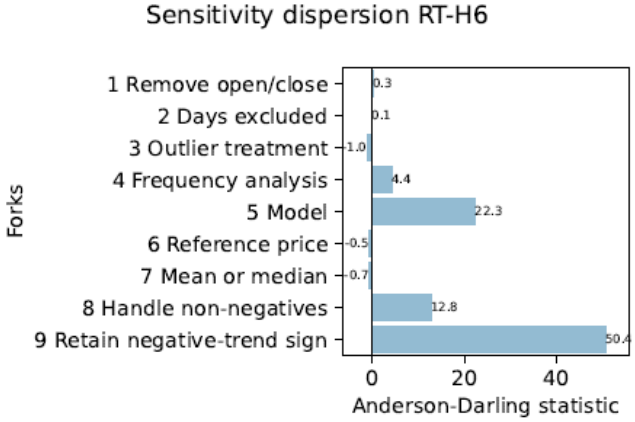

Sensitivity dispersion RT-H1
Forks
1 Remove open/close
2 Days excluded
3 Outlier treatment
4 Frequency analysis
5 Model
6 Measure
7 Measure frequencies
0.1
-0.7
0.1
5.6
37.7
0.8
14.4
0
10
20
30
40
Anderson-Darling statistic


Sensitivity dispersion RT-H2
Forks
1 Remove open/close
2 Days excluded
3 Outlier treatment
4 Frequency analysis
5 Model
6 Tick test or aggressor flag
7 Aggregation
8 Post trade value
0.1
-0.4
0.0
10.6
26.1
6.1
4.0
4.9
0
10
20
Anderson-Darling statistic


Sensitivity dispersion RT-H3
Forks
1 Remove open/close
2 Days excluded
3 Outlier treatment
4 Frequency analysis
5 Model
6 Units
1.8
4.0
2.1
12.4
58.9
0.1
0
20
40
60
Anderson-Darling statistic

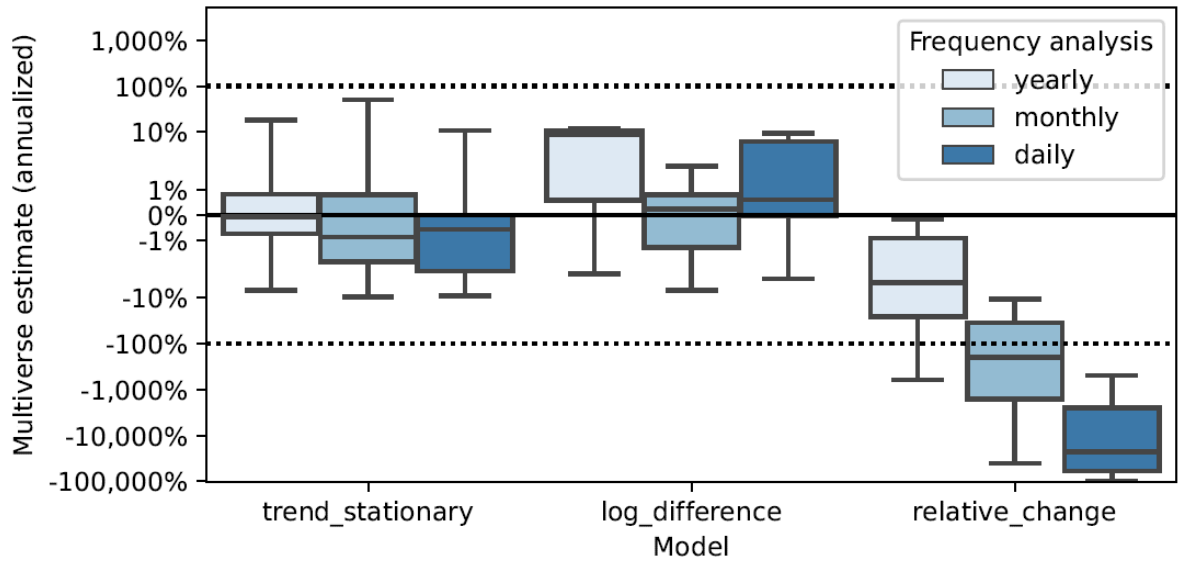

Sensitivity dispersion RT-H4
Forks
1 Remove open/close
2 Days excluded
3 Outlier treatment
4 Frequency analysis
5 Model
6 Tick test or aggressor flag
7 Aggregation
8 Post trade value
0.0
-0.3
2.7
13.6
12.9
10.9
8.5
-0.1
0
5
10
Anderson-Darling statistic


Sensitivity dispersion RT-H5
Forks
1 Remove open/close
2 Days excluded
3 Outlier treatment
4 Frequency analysis
5 Model
6 Tick test or aggressor flag
1.5
7.6
-0.9
47.5
22.6
8.2
0
20
40
Anderson-Darling statistic

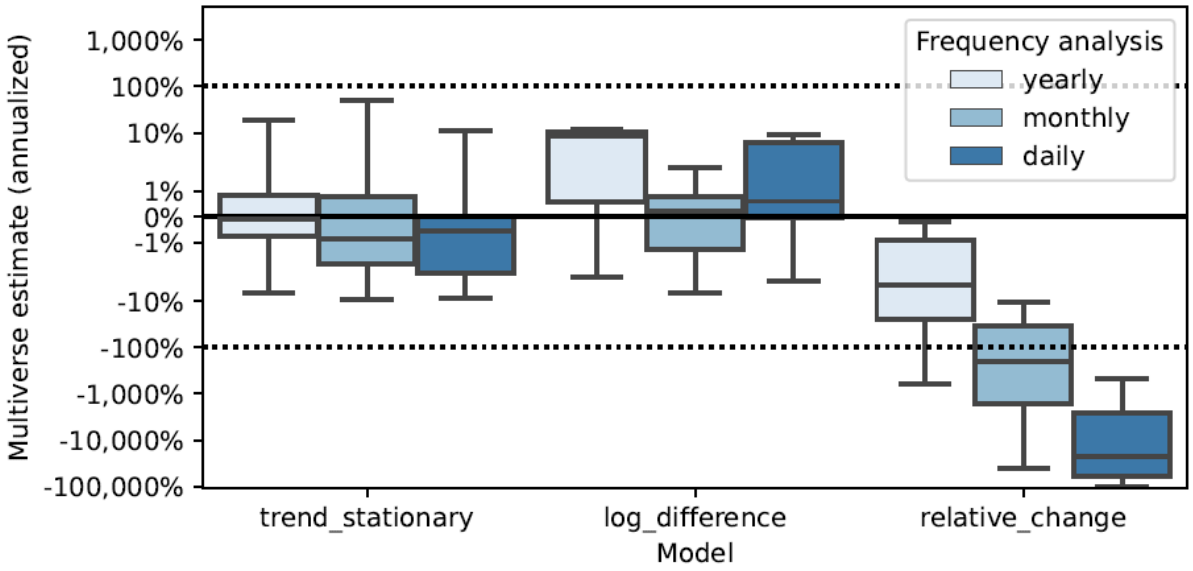

Sensitivity dispersion RT-H6
Forks
1 Remove open/close
2 Days excluded
3 Outlier treatment
4 Frequency analysis
5 Model
6 Reference price
7 Mean or median
8 Handle non-negatives
9 Retain negative-trend sign
0.3
0.1
-1.0
4.4
22.3
-0.5
-0.7
12.8
50.4
0
20
40
Anderson-Darling statistic

## 4.12. FIGURE 7. SENSITIVITY OF ESTIMATES IN MULTIVERSE ANALYSIS OF RT-H1

Original:




Reproduced:



## 4.13. TABLE IA.1. SUMMARY STATISTICS STAGE-2 DISPERSION IN RT-RESULTS

Original:

| | RT-H1 Efficiency | RT-H2 RSpread | RT-H3 Client Volume | RT-H4 Client RSpread | RT-H5 Client MOrders | RT-H6 Client GTR |
|---|---|---|---|---|---|---|
| *Estimate (yearly change, %)* | | | | | | |
| Mean | 451.2 | -1,122.2 | -3.6 | -38,254.7 | 0.4 | -37.1 |
| SD | 5,817.1 | 14,531.4 | 9.2 | 490,025.9 | 7.5 | 264.7 |
| Min | -291.3 | -186,074.5 | -117.5 | -6,275,383.0 | -30.8 | -3,024.9 |
| Q(0.10) | -13.1 | -7.7 | -3.8 | -5.8 | -1.8 | -83.4 |
| Q(0.25) | -4.4 | -4.7 | -3.7 | -2.7 | -0.6 | -9.4 |
| Median | -1.2 | -0.9 | -3.3 | 0.0 | -0.0 | -0.0 |
| Q(0.75) | 0.3 | 2.5 | -2.1 | 3.5 | 0.2 | 2.1 |
| Q(0.90) | 3.4 | 12.6 | -0.3 | 14.2 | 1.1 | 33.7 |
| black!10 IQR (i.e., NSE) | 4.7 | 7.1 | 1.7 | 6.2 | 0.8 | 11.6 |
| IDR | 16.5 | 20.2 | 3.5 | 20.0 | 2.8 | 117.1 |
| Max | 74,491.1 | 1,098.0 | 4.8 | 870.2 | 86.1 | 486.5 |
| *Standard error* | | | | | | |
| Mean | 462.2 | 1,166.4 | 3.5 | 38,279.4 | 2.0 | 86.3 |
| SD | 5,811.0 | 14,710.6 | 29.6 | 489,931.2 | 8.0 | 308.4 |
| Min | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| Q(0.10) | 0.1 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| Q(0.25) | 0.4 | 0.5 | 0.1 | 0.7 | 0.1 | 0.3 |
| Median | 1.4 | 3.3 | 0.4 | 3.0 | 0.5 | 4.9 |
| Q(0.75) | 5.8 | 8.4 | 1.9 | 9.0 | 1.4 | 45.5 |
| Q(0.90) | 29.0 | 28.5 | 2.0 | 24.7 | 2.8 | 160.7 |
| IQR | 5.4 | 7.8 | 1.8 | 8.3 | 1.2 | 45.2 |
| IDR | 28.9 | 28.4 | 2.0 | 24.7 | 2.8 | 160.7 |
| Max | 74,425.5 | 188,404.1 | 378.8 | 6,274,203.0 | 86.1 | 2,740.2 |
| *t-value* | | | | | | |
| Mean | -4.8 | 24.9 | -99.4 | 28.8 | -3.5 | -0.1 |
| SD | 28.4 | 434.5 | 627.1 | 400.5 | 73.8 | 4.4 |
| Min | -322.3 | -907.7 | -7,208.7 | -160.0 | -876.2 | -38.3 |
| Q(0.10) | -7.3 | -8.4 | -45.4 | -5.0 | -3.7 | -2.0 |
| Q(0.25) | -2.3 | -2.5 | -17.4 | -1.2 | -0.8 | -0.8 |
| Median | -0.8 | -0.4 | -3.5 | 0.0 | -0.1 | -0.0 |
| Q(0.75) | 0.5 | 0.7 | -1.7 | 1.0 | 1.0 | 1.0 |
| Q(0.90) | 2.3 | 1.6 | -0.4 | 2.5 | 3.1 | 1.4 |
| IQR | 2.8 | 3.3 | 15.7 | 2.3 | 1.7 | 1.8 |
| IDR | 9.6 | 10.0 | 45.0 | 7.5 | 6.8 | 3.4 |
| Max | 30.8 | 5,479.5 | 56.1 | 5,120.8 | 318.8 | 25.2 |

Reproduced:

| | RT-H1 Efficiency | RT-H2 RSpread | RT-H3 Client Volume | RT-H4 Client RSpread | RT-H5 Client MOrders | RT-H6 Client GTR |
|---|---|---|---|---|---|---|
| *Estimate (yearly change, %)* | | | | | | |
| Mean | 451.2 | -1,122.2 | -3.6 | -38,254.7 | 0.4 | -37.1 |
| SD | 5,817.1 | 14,531.4 | 9.2 | 490,025.9 | 7.5 | 264.7 |
| Min | -291.3 | -186,074.5 | -117.5 | -6,275,383.0 | -30.8 | -3,024.9 |
| Q(0.10) | -13.1 | -7.7 | -3.8 | -5.8 | -1.8 | -83.4 |
| Q(0.25) | -4.4 | -4.7 | -3.7 | -2.7 | -0.6 | -9.4 |
| Median | -1.2 | -0.9 | -3.3 | 0.0 | -0.0 | -0.0 |
| Q(0.75) | 0.3 | 2.5 | -2.1 | 3.5 | 0.2 | 2.1 |
| Q(0.90) | 3.4 | 12.6 | -0.3 | 14.2 | 1.1 | 33.7 |
| black!10 IQR (i.e., NSE) | 4.7 | 7.1 | 1.7 | 6.2 | 0.8 | 11.6 |
| IDR | 16.5 | 20.2 | 3.5 | 20.0 | 2.8 | 117.1 |
| Max | 74,491.1 | 1,098.0 | 4.8 | 870.2 | 86.1 | 486.5 |
| *Standard error* | | | | | | |
| Mean | 462.2 | 1,166.4 | 3.5 | 38,279.4 | 2.0 | 86.3 |
| SD | 5,811.0 | 14,710.6 | 29.6 | 489,931.2 | 8.0 | 308.4 |
| Min | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| Q(0.10) | 0.1 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| Q(0.25) | 0.4 | 0.5 | 0.1 | 0.7 | 0.1 | 0.3 |
| Median | 1.4 | 3.3 | 0.4 | 3.0 | 0.5 | 4.9 |
| Q(0.75) | 5.8 | 8.4 | 1.9 | 9.0 | 1.4 | 45.5 |
| Q(0.90) | 29.0 | 28.5 | 2.0 | 24.7 | 2.8 | 160.7 |
| IQR | 5.4 | 7.8 | 1.8 | 8.3 | 1.2 | 45.2 |
| IDR | 28.9 | 28.4 | 2.0 | 24.7 | 2.8 | 160.7 |
| Max | 74,425.5 | 188,404.1 | 378.8 | 6,274,203.0 | 86.1 | 2,740.2 |
| *t-value* | | | | | | |
| Mean | -4.8 | 24.9 | -99.4 | 28.8 | -3.5 | -0.1 |
| SD | 28.4 | 434.5 | 627.1 | 400.5 | 73.8 | 4.4 |
| Min | -322.3 | -907.7 | -7,208.7 | -160.0 | -876.2 | -38.3 |
| Q(0.10) | -7.3 | -8.4 | -45.4 | -5.0 | -3.7 | -2.0 |
| Q(0.25) | -2.3 | -2.5 | -17.4 | -1.2 | -0.8 | -0.8 |
| Median | -0.8 | -0.4 | -3.5 | 0.0 | -0.1 | -0.0 |
| Q(0.75) | 0.5 | 0.7 | -1.7 | 1.0 | 1.0 | 1.0 |
| Q(0.90) | 2.3 | 1.6 | -0.4 | 2.5 | 3.1 | 1.4 |
| IQR | 2.8 | 3.3 | 15.7 | 2.3 | 1.7 | 1.8 |
| IDR | 9.6 | 10.0 | 45.0 | 7.5 | 6.8 | 3.4 |
| Max | 30.8 | 5,479.5 | 56.1 | 5,120.8 | 318.8 | 25.2 |

## 4.14. TABLE IA.2. SUMMARY STATISTICS STAGE-3 DISPERSION IN RT-RESULTS

Original:

| | RT-H1 Efficiency | RT-H2 RSpread | RT-H3 Client Volume | RT-H4 Client RSpread | RT-H5 Client MOrders | RT-H6 Client GTR |
|---|---|---|---|---|---|---|
| *Estimate (yearly change, %)* | | | | | | |
| Mean | 453.4 | -1,130.6 | -3.1 | -38,263.4 | -2.4 | -2.4 |
| SD | 5,816.9 | 14,530.5 | 10.7 | 490,025.2 | 36.2 | 105.9 |
| Min | -70.5 | -186,074.0 | -117.5 | -6,275,383.0 | -452.9 | -898.7 |
| Q(0.10) | -6.7 | -8.0 | -3.9 | -7.1 | -1.7 | -15.1 |
| Q(0.25) | -3.2 | -5.7 | -3.8 | -3.4 | -0.6 | -0.5 |
| Median | -1.0 | -1.8 | -3.3 | -0.3 | -0.0 | 0.0 |
| Q(0.75) | -0.0 | 0.0 | -1.3 | 0.8 | 0.2 | 1.4 |
| Q(0.90) | 2.2 | 5.5 | -0.4 | 5.3 | 0.9 | 12.8 |
| black!10 IQR (i.e., NSE) | 3.2 | 5.8 | 2.4 | 4.1 | 0.7 | 1.8 |
| IDR | 8.9 | 13.6 | 3.5 | 12.4 | 2.6 | 28.0 |
| Max | 74,491.1 | 1,098.0 | 66.7 | 302.4 | 86.1 | 486.5 |
| *Standard error* | | | | | | |
| Mean | 458.6 | 1,156.2 | 3.0 | 38,264.0 | 4.2 | 40.5 |
| SD | 5,811.3 | 14,711.3 | 29.6 | 489,932.4 | 36.8 | 149.5 |
| Min | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| Q(0.10) | 0.1 | 0.0 | 0.0 | 0.0 | 0.1 | 0.0 |
| Q(0.25) | 0.3 | 0.4 | 0.1 | 0.4 | 0.1 | 0.0 |
| Median | 0.6 | 1.1 | 0.2 | 1.2 | 0.3 | 1.6 |
| Q(0.75) | 2.1 | 3.8 | 0.7 | 3.4 | 0.7 | 8.0 |
| Q(0.90) | 7.9 | 10.0 | 2.0 | 10.1 | 1.9 | 58.5 |
| IQR | 1.8 | 3.5 | 0.6 | 3.0 | 0.6 | 8.0 |
| IDR | 7.8 | 9.9 | 1.9 | 10.1 | 1.9 | 58.5 |
| Max | 74,425.5 | 188,404.0 | 378.8 | 6,274,203.0 | 463.7 | 1,149.3 |
| *t-value* | | | | | | |
| Mean | -3.7 | 25.2 | -56.5 | 29.8 | -3.3 | 0.3 |
| SD | 12.5 | 434.4 | 363.3 | 400.3 | 73.2 | 3.3 |
| Min | -131.7 | -908.3 | -3,800.0 | -160.0 | -876.2 | -10.2 |
| Q(0.10) | -7.9 | -8.8 | -36.5 | -5.5 | -4.2 | -1.6 |
| Q(0.25) | -3.8 | -5.1 | -28.5 | -2.9 | -1.5 | -0.6 |
| Median | -1.8 | -1.2 | -11.1 | -0.3 | -0.0 | 0.1 |
| Q(0.75) | 0.1 | 0.3 | -3.1 | 0.9 | 1.0 | 1.0 |
| Q(0.90) | 2.2 | 1.6 | -1.4 | 2.5 | 3.5 | 1.6 |
| IQR | 3.8 | 5.4 | 25.4 | 3.9 | 2.5 | 1.6 |
| IDR | 10.0 | 10.5 | 35.0 | 8.0 | 7.7 | 3.2 |
| Max | 11.9 | 5,479.5 | 56.1 | 5,120.8 | 318.8 | 25.2 |

Reproduced:

| | RT-H1 Efficiency | RT-H2 RSpread | RT-H3 Client Volume | RT-H4 Client RSpread | RT-H5 Client MOrders | RT-H6 Client GTR |
|---|---|---|---|---|---|---|
| *Estimate (yearly change, %)* | | | | | | |
| Mean | 453.4 | -1,130.6 | -3.1 | -38,263.4 | -2.4 | -2.4 |
| SD | 5,816.9 | 14,530.5 | 10.7 | 490,025.2 | 36.2 | 105.9 |
| Min | -70.5 | -186,074.0 | -117.5 | -6,275,383.0 | -452.9 | -898.7 |
| Q(0.10) | -6.7 | -8.0 | -3.9 | -7.1 | -1.7 | -15.1 |
| Q(0.25) | -3.2 | -5.7 | -3.8 | -3.4 | -0.6 | -0.5 |
| Median | -1.0 | -1.8 | -3.3 | -0.3 | -0.0 | 0.0 |
| Q(0.75) | -0.0 | 0.0 | -1.3 | 0.8 | 0.2 | 1.4 |
| Q(0.90) | 2.2 | 5.5 | -0.4 | 5.3 | 0.9 | 12.8 |
| black!10 IQR (i.e., NSE) | 3.2 | 5.8 | 2.4 | 4.1 | 0.7 | 1.8 |
| IDR | 8.9 | 13.6 | 3.5 | 12.4 | 2.6 | 28.0 |
| Max | 74,491.1 | 1,098.0 | 66.7 | 302.4 | 86.1 | 486.5 |
| *Standard error* | | | | | | |
| Mean | 458.6 | 1,156.2 | 3.0 | 38,264.0 | 4.2 | 40.5 |
| SD | 5,811.3 | 14,711.3 | 29.6 | 489,932.4 | 36.8 | 149.5 |
| Min | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| Q(0.10) | 0.1 | 0.0 | 0.0 | 0.0 | 0.1 | 0.0 |
| Q(0.25) | 0.3 | 0.4 | 0.1 | 0.4 | 0.1 | 0.0 |
| Median | 0.6 | 1.1 | 0.2 | 1.2 | 0.3 | 1.6 |
| Q(0.75) | 2.1 | 3.8 | 0.7 | 3.4 | 0.7 | 8.0 |
| Q(0.90) | 7.9 | 10.0 | 2.0 | 10.1 | 1.9 | 58.5 |
| IQR | 1.8 | 3.5 | 0.6 | 3.0 | 0.6 | 8.0 |
| IDR | 7.8 | 9.9 | 1.9 | 10.1 | 1.9 | 58.5 |
| Max | 74,425.5 | 188,404.0 | 378.8 | 6,274,203.0 | 463.7 | 1,149.3 |
| *t-value* | | | | | | |
| Mean | -3.7 | 25.2 | -56.5 | 29.8 | -3.3 | 0.3 |
| SD | 12.5 | 434.4 | 363.3 | 400.3 | 73.2 | 3.3 |
| Min | -131.7 | -908.3 | -3,800.0 | -160.0 | -876.2 | -10.2 |
| Q(0.10) | -7.9 | -8.8 | -36.5 | -5.5 | -4.2 | -1.6 |
| Q(0.25) | -3.8 | -5.1 | -28.5 | -2.9 | -1.5 | -0.6 |
| Median | -1.8 | -1.2 | -11.1 | -0.3 | -0.0 | 0.1 |
| Q(0.75) | 0.1 | 0.3 | -3.1 | 0.9 | 1.0 | 1.0 |
| Q(0.90) | 2.2 | 1.6 | -1.4 | 2.5 | 3.5 | 1.6 |
| IQR | 3.8 | 5.4 | 25.4 | 3.9 | 2.5 | 1.6 |
| IDR | 10.0 | 10.5 | 35.0 | 8.0 | 7.7 | 3.2 |
| Max | 11.9 | 5,479.5 | 56.1 | 5,120.8 | 318.8 | 25.2 |

## 4.15. TABLE IA.3. SUMMARY STATISTICS STAGE-4 DISPERSION IN RT-RESULTS

Original:

| | RT-H1 Efficiency | RT-H2 RSpread | RT-H3 Client Volume | RT-H4 Client RSpread | RT-H5 Client MOrders | RT-H6 Client GTR |
|---|---|---|---|---|---|---|
| *Estimate (yearly change, %)* | | | | | | |
| Mean | 453.5 | 1,138.6 | -1.8 | -38,263.2 | -2.9 | 4.9 |
| SD | 5,816.9 | 14,529.9 | 9.6 | 490,025.2 | 35.4 | 71.2 |
| Min | -70.5 | -90.1 | -6.9 | -6,275,383.0 | -452.9 | -360.7 |
| Q(0.10) | -6.2 | -6.9 | -3.8 | -5.8 | -1.3 | -5.0 |
| Q(0.25) | -2.8 | -4.4 | -3.8 | -2.0 | -0.5 | -0.2 |
| Median | -1.1 | -2.3 | -2.9 | -0.2 | 0.0 | 0.0 |
| Q(0.75) | -0.2 | -0.1 | -2.0 | 0.4 | 0.1 | 0.8 |
| Q(0.90) | 1.2 | 2.2 | -1.1 | 3.6 | 0.8 | 5.7 |
| black!10 IQR (i.e., NSE) | 2.6 | 4.3 | 1.7 | 2.4 | 0.6 | 1.1 |
| IDR | 7.4 | 9.1 | 2.7 | 9.5 | 2.1 | 10.8 |
| Max | 74,491.1 | 186,074.5 | 117.5 | 302.4 | 7.1 | 486.5 |
| *Standard error* | | | | | | |
| Mean | 457.9 | 1,155.3 | 3.0 | 38,261.5 | 3.5 | 24.7 |
| SD | 5,811.4 | 14,711.4 | 29.6 | 489,932.6 | 36.2 | 88.8 |
| Min | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| Q(0.10) | 0.1 | 0.1 | 0.1 | 0.0 | 0.1 | 0.0 |
| Q(0.25) | 0.3 | 0.6 | 0.1 | 0.6 | 0.1 | 0.1 |
| Median | 0.5 | 1.2 | 0.3 | 1.3 | 0.3 | 2.0 |
| Q(0.75) | 1.5 | 3.0 | 0.5 | 2.7 | 0.6 | 5.2 |
| Q(0.90) | 5.2 | 7.0 | 1.8 | 5.5 | 1.2 | 46.6 |
| IQR | 1.2 | 2.4 | 0.4 | 2.1 | 0.4 | 5.1 |
| IDR | 5.1 | 7.0 | 1.8 | 5.5 | 1.1 | 46.6 |
| Max | 74,425.5 | 188,404.1 | 378.8 | 6,274,203.0 | 463.7 | 786.1 |
| *t-value* | | | | | | |
| Mean | -3.7 | 25.1 | -54.7 | 29.7 | -3.6 | 0.6 |
| SD | 12.1 | 434.4 | 363.4 | 400.2 | 73.0 | 6.4 |
| Min | -131.7 | -911.2 | -3,801.4 | -159.8 | -876.2 | -9.0 |
| Q(0.10) | -7.5 | -8.0 | -33.5 | -3.6 | -3.6 | -1.3 |
| Q(0.25) | -3.1 | -4.0 | -18.5 | -1.8 | -1.1 | -0.3 |
| Median | -2.0 | -1.9 | -11.5 | -0.2 | 0.0 | 0.1 |
| Q(0.75) | -0.4 | -0.4 | -4.1 | 0.3 | 1.0 | 0.8 |
| Q(0.90) | 1.8 | 0.8 | -1.7 | 1.8 | 2.9 | 1.4 |
| IQR | 2.7 | 3.6 | 14.3 | 2.1 | 2.1 | 1.1 |
| IDR | 9.4 | 8.8 | 31.8 | 5.4 | 6.5 | 2.7 |
| Max | 8.0 | 5,479.5 | 19.5 | 5,120.8 | 318.8 | 80.2 |

Reproduced:

| | RT-H1 Efficiency | RT-H2 RSpread | RT-H3 Client Volume | RT-H4 Client RSpread | RT-H5 Client MOrders | RT-H6 Client GTR |
|---|---|---|---|---|---|---|
| *Estimate (yearly change, %)* | | | | | | |
| Mean | 453.5 | 1,138.6 | -1.8 | -38,263.2 | -2.9 | 4.9 |
| SD | 5,816.9 | 14,529.9 | 9.6 | 490,025.2 | 35.4 | 71.2 |
| Min | -70.5 | -90.1 | -6.9 | -6,275,383.0 | -452.9 | -360.7 |
| Q(0.10) | -6.2 | -6.9 | -3.8 | -5.8 | -1.3 | -5.0 |
| Q(0.25) | -2.8 | -4.4 | -3.8 | -2.0 | -0.5 | -0.2 |
| Median | -1.1 | -2.3 | -2.9 | -0.2 | 0.0 | 0.0 |
| Q(0.75) | -0.2 | -0.1 | -2.0 | 0.4 | 0.1 | 0.8 |
| Q(0.90) | 1.2 | 2.2 | -1.1 | 3.6 | 0.8 | 5.7 |
| black!10 IQR (i.e., NSE) | 2.6 | 4.3 | 1.7 | 2.4 | 0.6 | 1.1 |
| IDR | 7.4 | 9.1 | 2.7 | 9.5 | 2.1 | 10.8 |
| Max | 74,491.1 | 186,074.5 | 117.5 | 302.4 | 7.1 | 486.5 |
| *Standard error* | | | | | | |
| Mean | 457.9 | 1,155.3 | 3.0 | 38,261.5 | 3.5 | 24.7 |
| SD | 5,811.4 | 14,711.4 | 29.6 | 489,932.6 | 36.2 | 88.8 |
| Min | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| Q(0.10) | 0.1 | 0.1 | 0.1 | 0.0 | 0.1 | 0.0 |
| Q(0.25) | 0.3 | 0.6 | 0.1 | 0.6 | 0.1 | 0.1 |
| Median | 0.5 | 1.2 | 0.3 | 1.3 | 0.3 | 2.0 |
| Q(0.75) | 1.5 | 3.0 | 0.5 | 2.7 | 0.6 | 5.2 |
| Q(0.90) | 5.2 | 7.0 | 1.8 | 5.5 | 1.2 | 46.6 |
| IQR | 1.2 | 2.4 | 0.4 | 2.1 | 0.4 | 5.1 |
| IDR | 5.1 | 7.0 | 1.8 | 5.5 | 1.1 | 46.6 |
| Max | 74,425.5 | 188,404.1 | 378.8 | 6,274,203.0 | 463.7 | 786.1 |
| *t-value* | | | | | | |
| Mean | -3.7 | 25.1 | -54.7 | 29.7 | -3.6 | 0.6 |
| SD | 12.1 | 434.4 | 363.4 | 400.2 | 73.0 | 6.4 |
| Min | -131.7 | -911.2 | -3,801.4 | -159.8 | -876.2 | -9.0 |
| Q(0.10) | -7.5 | -8.0 | -33.5 | -3.6 | -3.6 | -1.3 |
| Q(0.25) | -3.1 | -4.0 | -18.5 | -1.8 | -1.1 | -0.3 |
| Median | -2.0 | -1.9 | -11.5 | -0.2 | 0.0 | 0.1 |
| Q(0.75) | -0.4 | -0.4 | -4.1 | 0.3 | 1.0 | 0.8 |
| Q(0.90) | 1.8 | 0.8 | -1.7 | 1.8 | 2.9 | 1.4 |
| IQR | 2.7 | 3.6 | 14.3 | 2.1 | 2.1 | 1.1 |
| IDR | 9.4 | 8.8 | 31.8 | 5.4 | 6.5 | 2.7 |
| Max | 8.0 | 5,479.5 | 19.5 | 5,120.8 | 318.8 | 80.2 |

## 4.16. TABLE IA.4. PRINCIPAL COMPONENT ANALYSIS TEAM QUALITY

Original:

### Panel (a): Correlation team quality measures

| | Publications | Experience | Big Data | Position | #Members |
|---|---|---|---|---|---|
| Publications | | 0.34 | 0.10 | 0.54 | 0.30 |
| Experience | | | -0.18 | 0.25 | 0.12 |
| Big Data | | | | 0.14 | 0.14 |
| Position | | | | | 0.16 |

### Panel (b): Fraction of variance explained

| | PC1 | PC2 | PC3 | PC4 | PC5 |
|---|---|---|---|---|---|
| Variance explained | 38.3% | 23.6% | 17.1% | 12.4% | 8.6% |

### Panel (c): Loading of principal components on variables

| | Publications | Experience | Big Data | Position | #Members |
|---|---|---|---|---|---|
| PC1 | 0.61 | 0.40 | 0.13 | 0.55 | 0.37 |
| PC2 | -0.01 | -0.55 | 0.79 | 0.05 | 0.26 |
| PC3 | -0.10 | 0.06 | -0.21 | -0.46 | 0.86 |
| PC4 | -0.20 | 0.71 | 0.56 | -0.35 | -0.12 |
| PC5 | -0.76 | 0.14 | -0.02 | 0.60 | 0.22 |

Reproduced:

## Panel (a): Correlation team quality measures

| | Publications | Experience | Big Data | Position | #Members |
|---|---|---|---|---|---|
| Publications | | 0.34 | 0.10 | 0.54 | 0.30 |
| Experience | | | -0.18 | 0.25 | 0.12 |
| Big Data | | | | 0.14 | 0.14 |
| Position | | | | | 0.16 |

## Panel (b): Fraction of variance explained

| | PC1 | PC2 | PC3 | PC4 | PC5 |
|---|---|---|---|---|---|
| Variance explained | 38.3% | 23.6% | 17.1% | 12.4% | 8.6% |

## Panel (c): Loading of principal components on variables

| | Publications | Experience | Big Data | Position | #Members |
|---|---|---|---|---|---|
| PC1 | 0.61 | 0.40 | 0.13 | 0.55 | 0.37 |
| PC2 | -0.01 | -0.55 | 0.79 | 0.05 | 0.26 |
| PC3 | -0.10 | 0.06 | -0.21 | -0.46 | 0.86 |
| PC4 | -0.20 | 0.71 | 0.56 | -0.35 | -0.12 |
| PC5 | -0.76 | 0.14 | -0.02 | 0.60 | 0.22 |

## 4.17. TABLE IA.5. STAGE-1 QUANTILE REGRESSIONS WITH ALL TEAM QUALITY VARIABLES

Original:

| | Q(0.10) | Q(0.25) | Q(0.50) | Q(0.75) | Q(0.90) |
|---|---|---|---|---|---|
| Top publications (standardized/scaled) | -0.200** (0.063) | -0.213** (0.020) | 0.000 (0.009) | 0.056** (0.013) | 0.648** (0.037) |
| Experience in field (standardized/scaled) | -0.370** (0.040) | 0.031* (0.015) | 0.000 (0.007) | 0.002 (0.012) | -0.106** (0.028) |
| Experience with big data (standardized/scaled) | 0.074 (0.048) | -0.025 (0.015) | 0.000 (0.007) | 0.030* (0.011) | 0.138** (0.027) |
| Academic seniority (standardized/scaled) | 1.666** (0.058) | 0.171** (0.019) | 0.001 (0.008) | -0.026* (0.013) | -1.022** (0.035) |
| Team size (1 or 2 members) (standardized/scaled) | 0.819** (0.044) | 0.068** (0.015) | 0.004 (0.007) | -0.061** (0.011) | -0.150** (0.028) |
| Reproducibility score (standardized/scaled) | 0.494** (0.043) | 0.186** (0.015) | -0.001 (0.007) | -0.116** (0.011) | -0.521** (0.028) |
| Average rating (standardized/scaled) | 0.521** (0.041) | 0.158** (0.014) | -0.001 (0.007) | -0.067** (0.011) | -0.486** (0.029) |
| Dummy RT-H1 Efficiency | -31.553** (1.032) | -6.779** (0.354) | -1.107** (0.164) | 0.844** (0.271) | 9.262** (0.664) |
| Dummy RT-H2 RSpread | -22.463** (0.990) | -4.543** (0.356) | -0.032 (0.165) | 3.846** (0.270) | 20.961** (0.647) |
| Dummy RT-H3 Client Volume | -6.548** (1.007) | -3.776** (0.354) | -3.315** (0.164) | -2.361** (0.270) | 0.023 (0.661) |
| Dummy RT-H4 Client RSpread | -18.079** (0.952) | -2.875** (0.351) | 0.158 (0.164) | 4.365** (0.270) | 18.189** (0.645) |
| Dummy RT-H5 Client MOrders | -3.169** (1.007) | -0.853* (0.354) | -0.003 (0.164) | 0.286 (0.270) | 1.817** (0.625) |
| Dummy RT-H6 GTR | -178.868** (0.993) | -20.853** (0.351) | -0.020 (0.166) | 5.672** (0.265) | 60.704** (0.676) |
| #Observations | 984 | 984 | 984 | 984 | 984 |

Reproduced:

| | Q(0.10) | Q(0.25) | Q(0.50) | Q(0.75) | Q(0.90) |
|---|---|---|---|---|---|
| Top publications (standardized/scaled) | -0.200** (0.064) | -0.213** (0.020) | 0.000 (0.009) | 0.056** (0.014) | 0.648** (0.034) |
| Experience in field (standardized/scaled) | -0.370** (0.040) | 0.031* (0.015) | 0.000 (0.007) | 0.002 (0.012) | -0.106** (0.028) |
| Experience with big data (standardized/scaled) | 0.074 (0.048) | -0.025 (0.015) | 0.000 (0.007) | 0.030* (0.012) | 0.138** (0.028) |
| Academic seniority (standardized/scaled) | 1.666** (0.058) | 0.171** (0.019) | 0.001 (0.008) | -0.026* (0.013) | -1.022** (0.034) |
| Team size (1 or 2 members) (standardized/scaled) | 0.819** (0.044) | 0.068** (0.015) | 0.004 (0.007) | -0.061** (0.011) | -0.150** (0.028) |
| Reproducibility score (standardized/scaled) | 0.494** (0.043) | 0.186** (0.015) | -0.001 (0.007) | -0.116** (0.012) | -0.521** (0.026) |
| Average rating (standardized/scaled) | 0.521** (0.042) | 0.158** (0.014) | -0.001 (0.007) | -0.067** (0.011) | -0.486** (0.029) |
| Dummy RT-H1 Efficiency | -31.553** (1.032) | -6.779** (0.354) | -1.107** (0.164) | 0.844** (0.270) | 9.262** (0.664) |
| Dummy RT-H2 RSpread | -22.463** (1.018) | -4.543** (0.356) | -0.032 (0.165) | 3.846** (0.270) | 20.961** (0.647) |
| Dummy RT-H3 Client Volume | -6.548** (1.007) | -3.776** (0.354) | -3.315** (0.164) | -2.361** (0.270) | 0.023 (0.661) |
| Dummy RT-H4 Client RSpread | -18.079** (0.952) | -2.875** (0.351) | 0.158 (0.164) | 4.365** (0.270) | 18.189** (0.663) |
| Dummy RT-H5 Client MOrders | -3.169** (0.980) | -0.853* (0.354) | -0.003 (0.164) | 0.286 (0.270) | 1.817** (0.643) |
| Dummy RT-H6 GTR | -178.868** (1.020) | -20.853** (0.351) | -0.020 (0.166) | 5.672** (0.274) | 60.704** (0.656) |
| #Observations | 984 | 984 | 984 | 984 | 984 |

## 4.18. TABLE IA.6. DISPERSION IN RESEARCH TEAM BELIEFS

Original:

| | RT-H1 Efficiency | RT-H2 RSpread | RT-H3 Client Volume | RT-H4 Client RSpread | RT-H5 Client MOrders | RT-H6 Client GTR | All |
|---|---|---|---|---|---|---|---|
| Estimate | -99.5%** (0.00) | -95.4%** (0.00) | -9.0% (0.74) | -97.5%** (0.00) | -45.3% (0.63) | -83.3%** (0.00) | -71.7%** (0.00) |

Reproduced:

| | RT-H1 Efficiency | RT-H2 RSpread | RT-H3 Client Volume | RT-H4 Client RSpread | RT-H5 Client MOrders | RT-H6 Client GTR | All |
|---|---|---|---|---|---|---|---|
| Estimate | -99.5%** (0.00) | -95.4%** (0.00) | -9.0% (0.74) | -97.5%** (0.00) | -45.3% (0.63) | -83.3%** (0.00) | -71.7%** (0.00) |

## 4.19. TABLE IA.7. STATISTICAL SIGNIFICANCE UNDER MULTIPLE TESTING (PANEL A ONLY)

Original:

Panel (a): Multiple tests (Bonferroni)

| | Reject at 0.5%? | $p$-value of family test | Mean (SD) correlation test statistics | Effective number of tests |
|---|---|---|---|---|
| RT-H1 | Yes (31, 4) | $<0.0001$ | 0.00 (0.00) | 164 |
| RT-H2 | Yes (38, 3) | $<0.0001$ | 0.00 (0.00) | 164 |
| RT-H3 | Yes (123, 2) | $<0.0001$ | 0.00 (0.00) | 164 |
| RT-H4 | Yes (15, 8) | $<0.0001$ | 0.00 (0.00) | 164 |
| RT-H5 | Yes (13, 9) | $<0.0001$ | 0.00 (0.00) | 164 |
| RT-H6 | Yes (3, 3) | $<0.0001$ | 0.00 (0.00) | 164 |

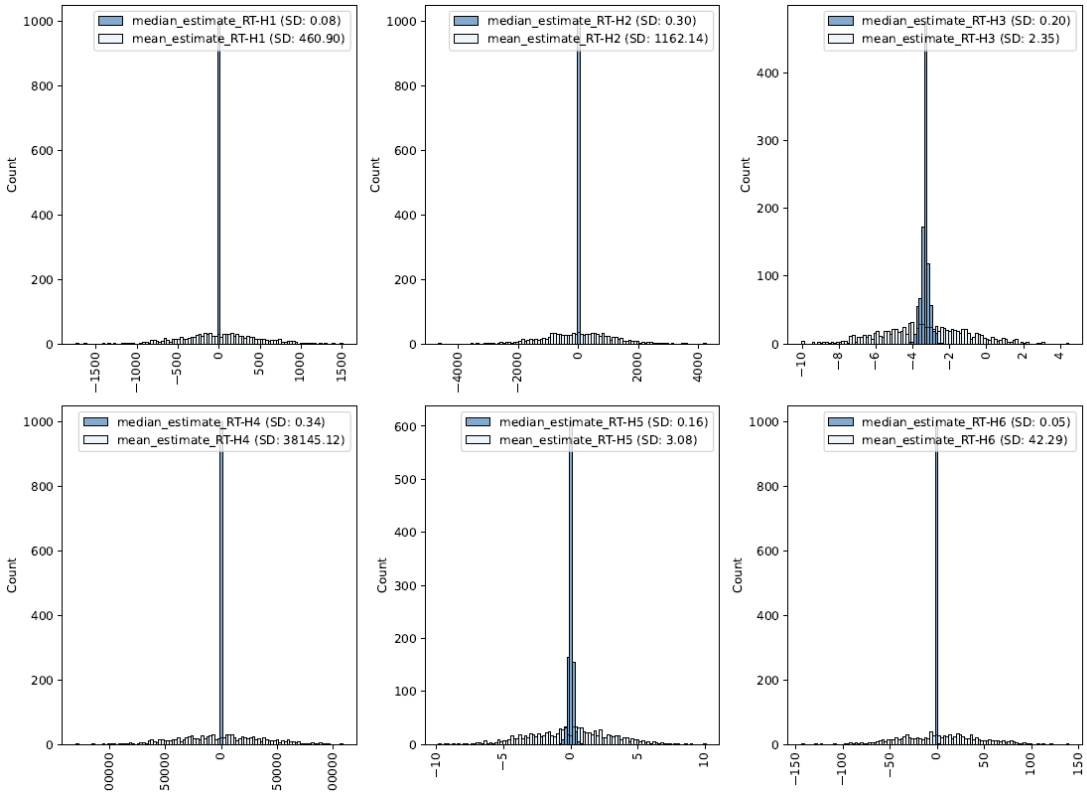
Reproduced:

Panel (a): Multiple tests (Bonferroni)

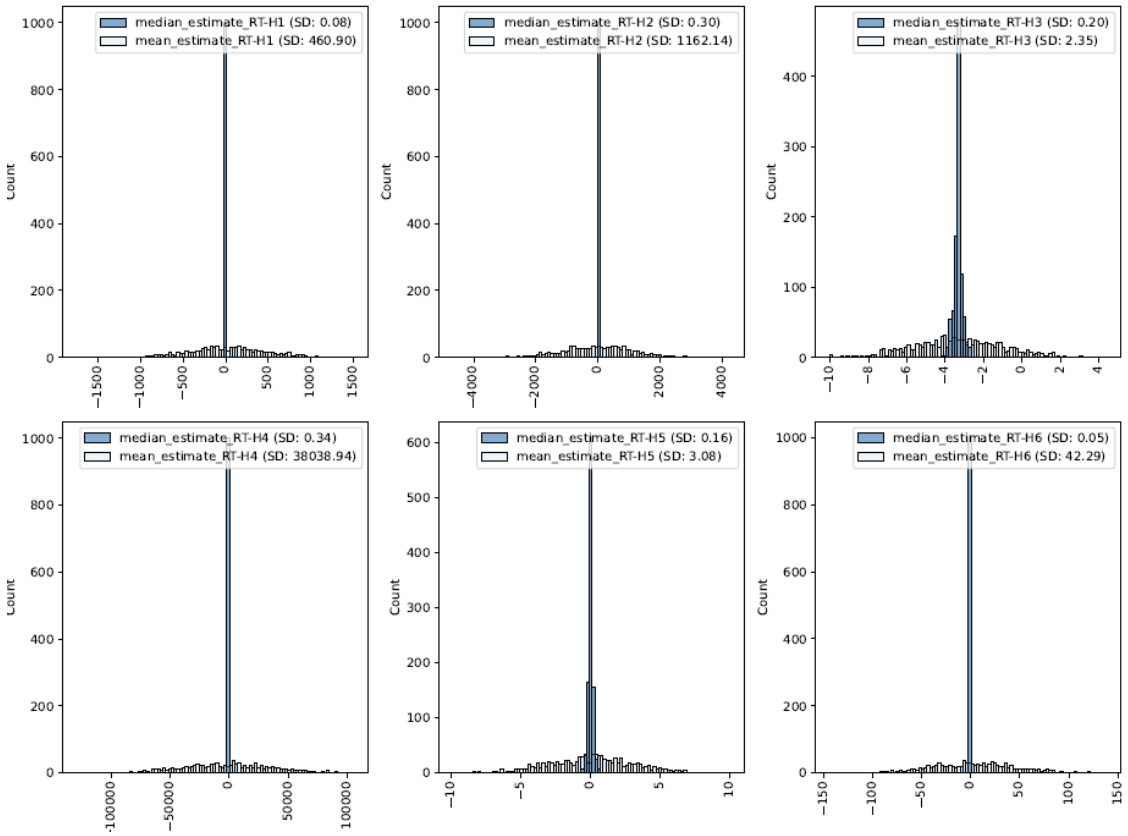
| | Reject at 0.5%? | $p$-value of family test | Mean (SD) correlation test statistics | Effective number of tests |
|---|---|---|---|---|
| RT-H1 | Yes (31, 4) | $<0.0001$ | 0.00 (0.00) | 164 |
| RT-H2 | Yes (38, 3) | $<0.0001$ | 0.00 (0.00) | 164 |
| RT-H3 | Yes (123, 2) | $<0.0001$ | 0.00 (0.00) | 164 |
| RT-H4 | Yes (15, 8) | $<0.0001$ | 0.00 (0.00) | 164 |
| RT-H5 | Yes (13, 9) | $<0.0001$ | 0.00 (0.00) | 164 |
| RT-H6 | Yes (3, 3) | $<0.0001$ | 0.00 (0.00) | 164 |

## 4.20. FIGURE IA.1. DISTRIBUTION OF THE MEDIAN/MEAN ESTIMATE

Original:

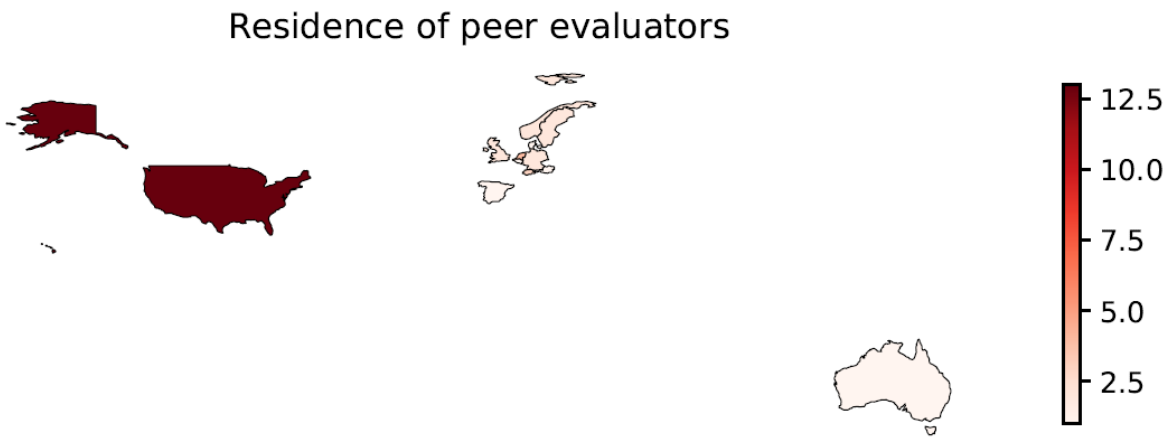



Reproduced:



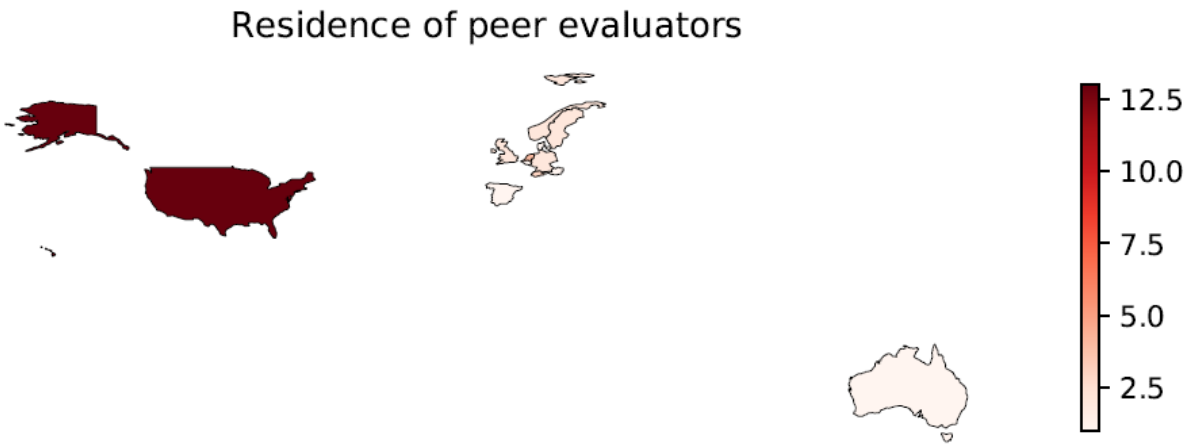
## 4.21. FIGURE IA.2. COUNTRIES OF #FINCAP COMMUNITY

Original:



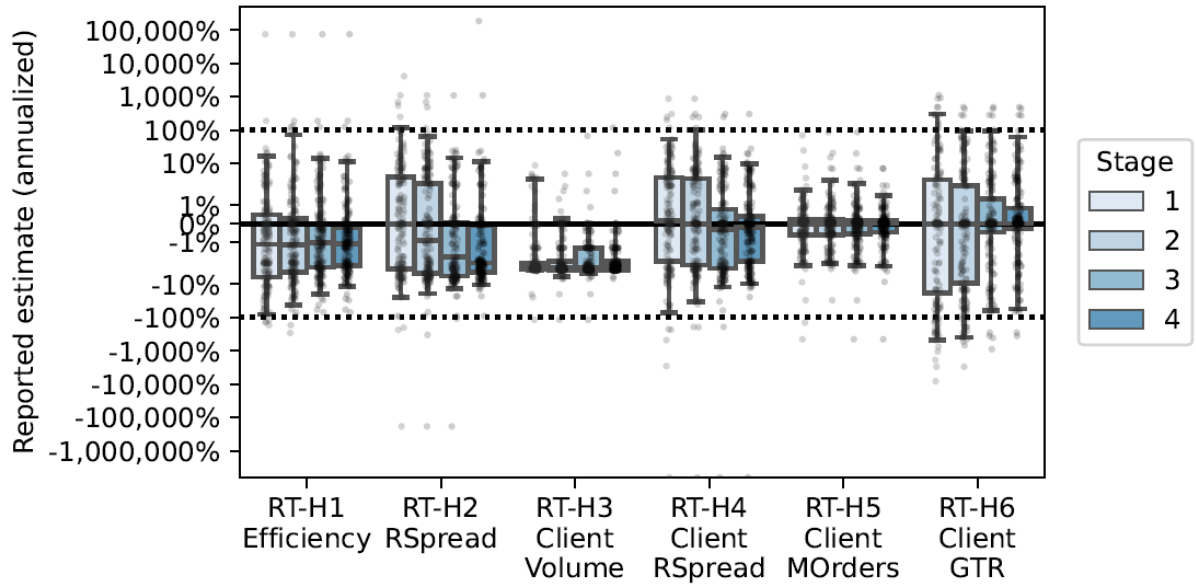
Reproduced:

Residence of research-team members
50
40
30
20
10

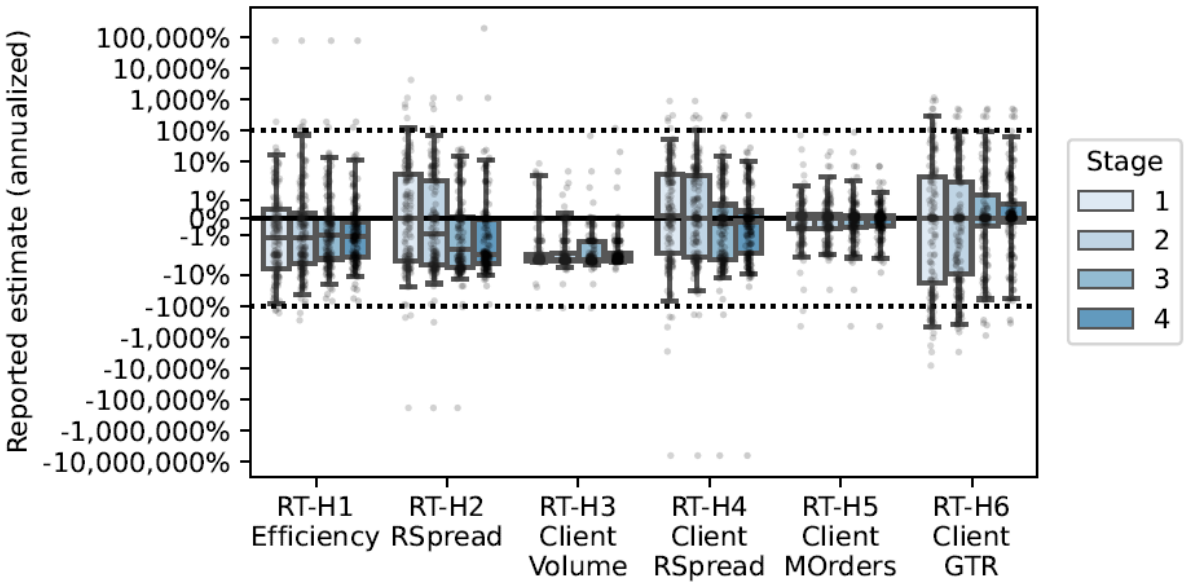

Residence of peer evaluators
12.5
10.0
7.5
5.0
2.5

## 4.22. FIGURE IA.3. DISPERSION ESTIMATES ACROSS ALL FEEDBACK STAGES

Original:

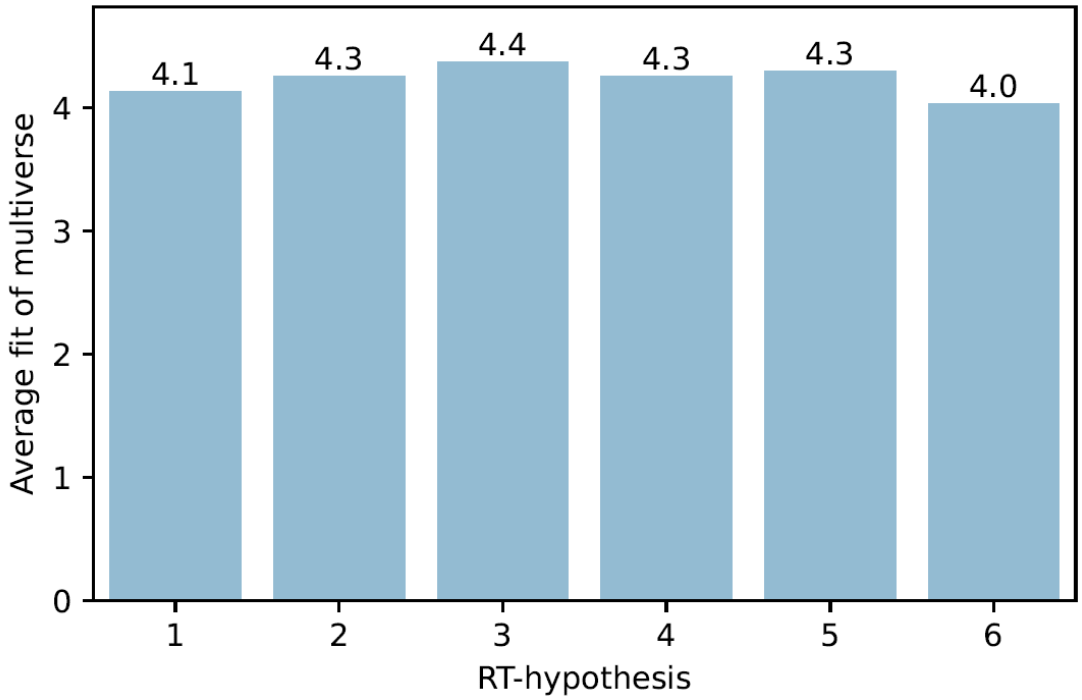



Reproduced:

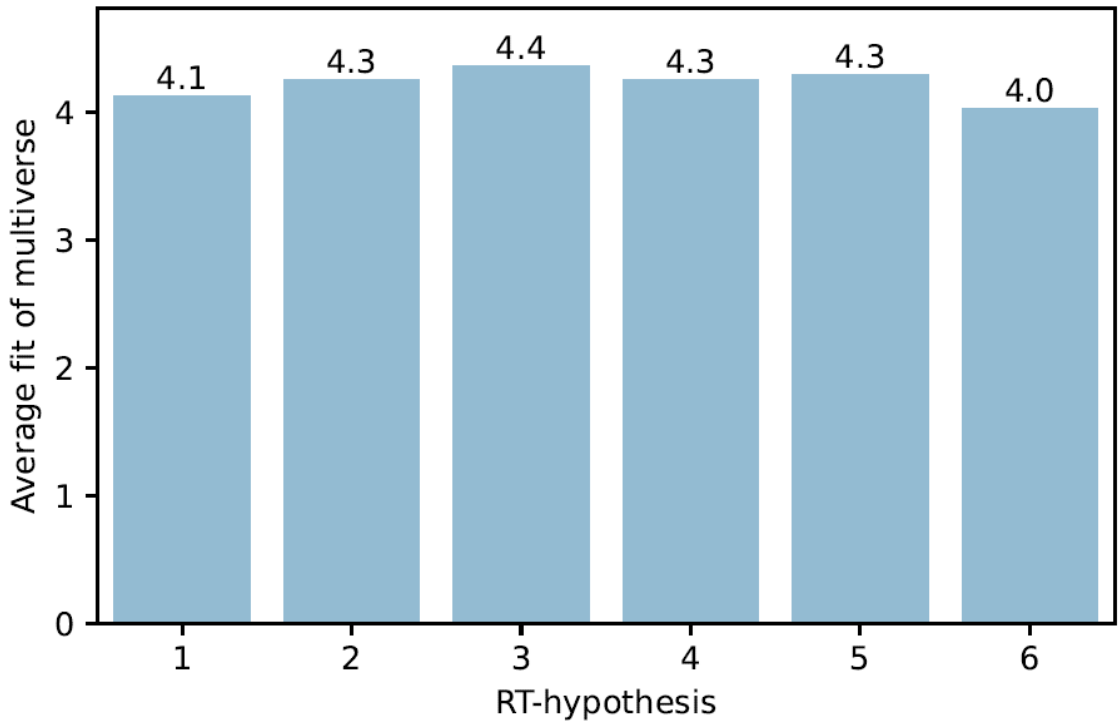


## 4.23. FIGURE IA.4. MULTIVERSE FIT

Original:

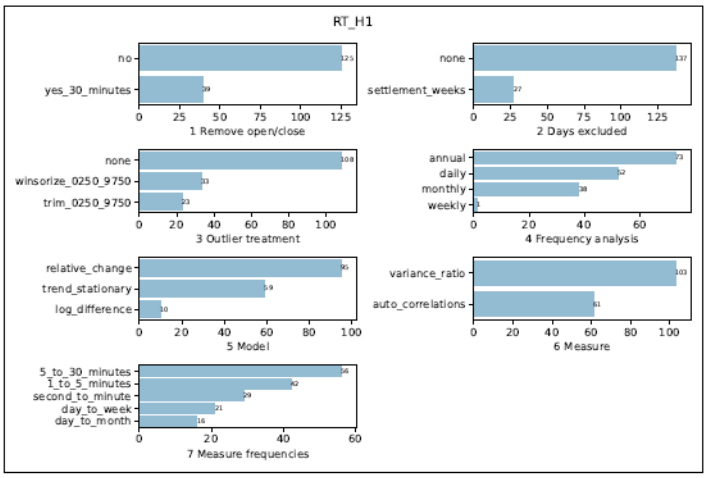


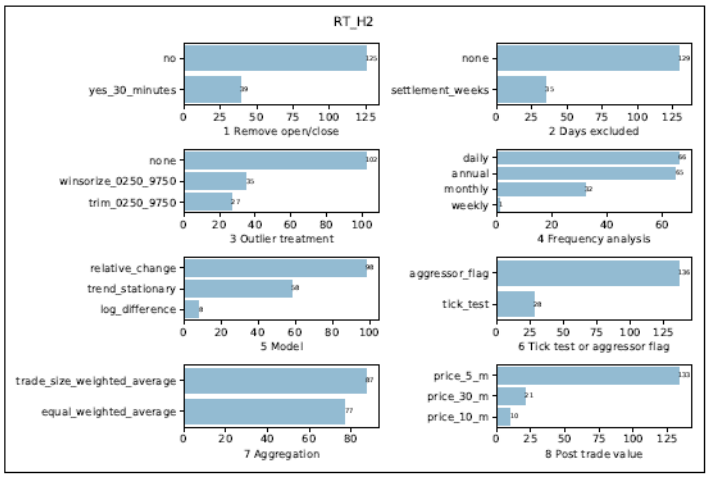

Reproduced:

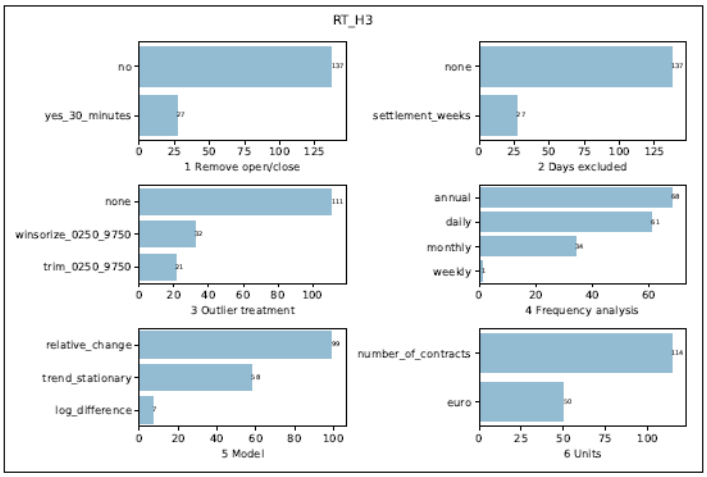


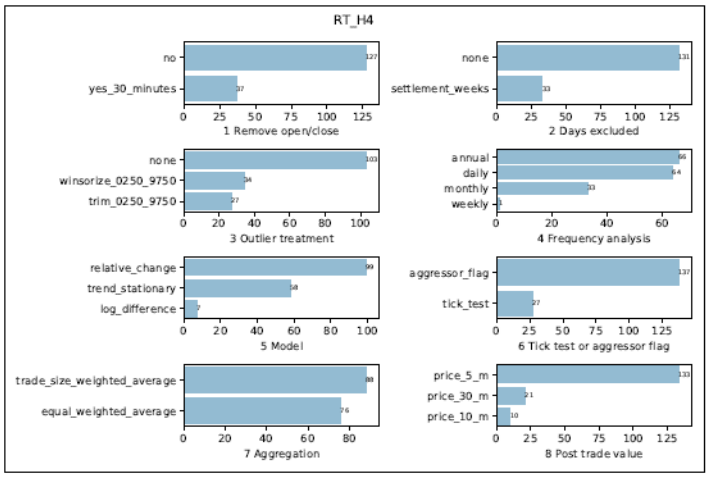
## 4.24. FIGURE IA.5. DISTRIBUTION DECISIONS AT THE FORKS

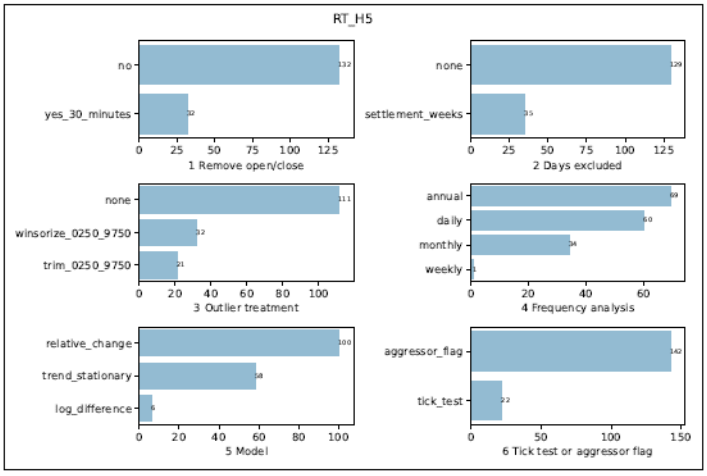
Original:

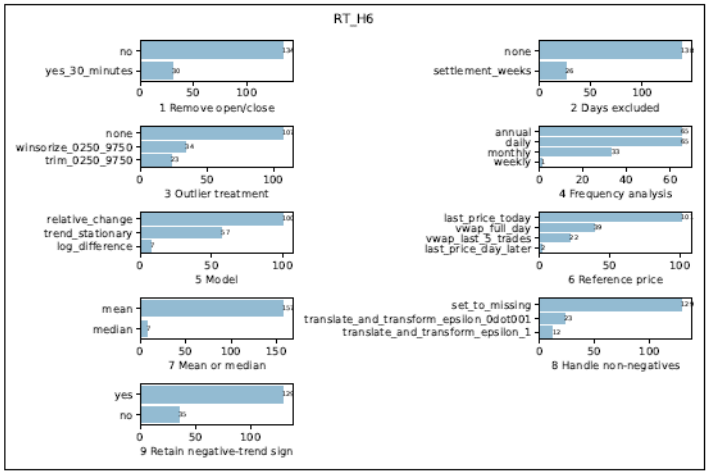


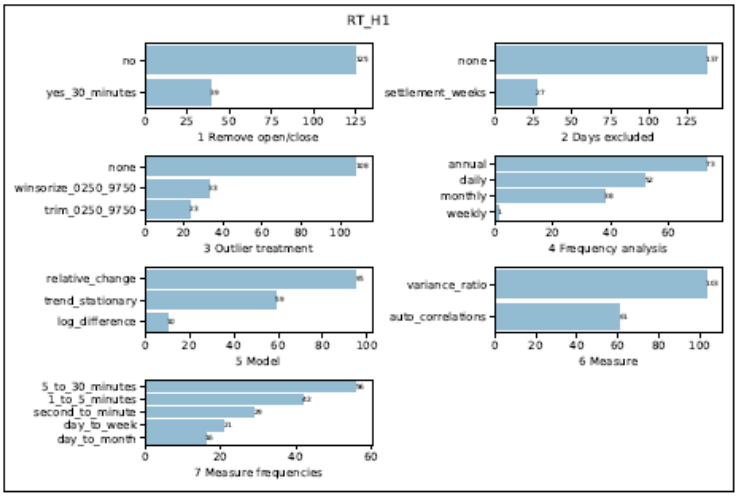



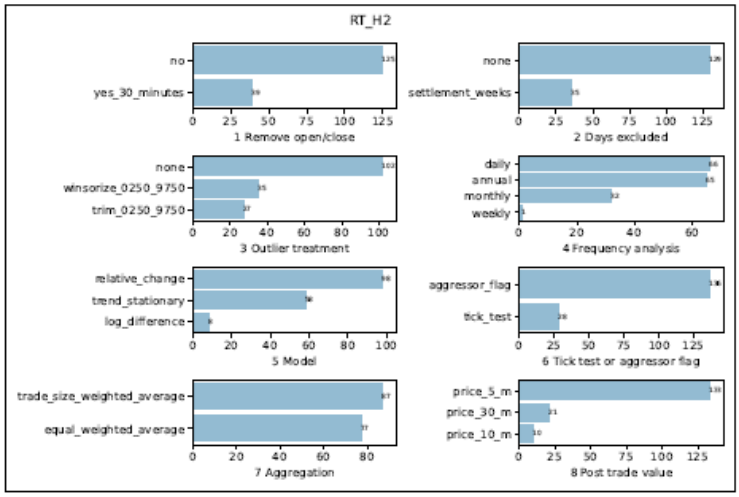



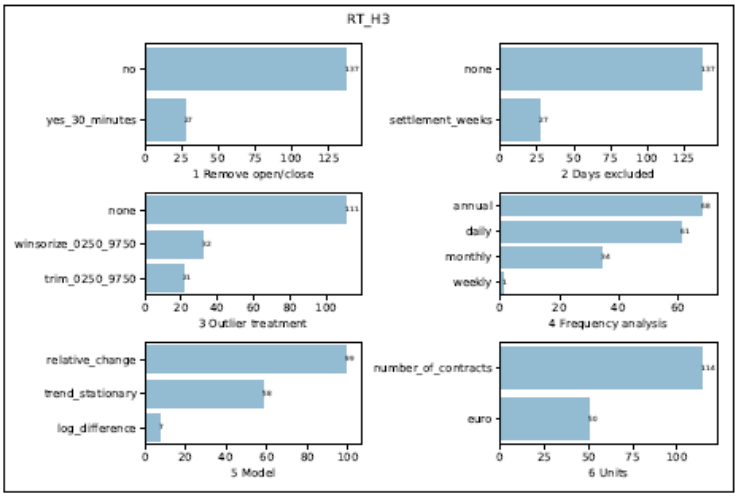



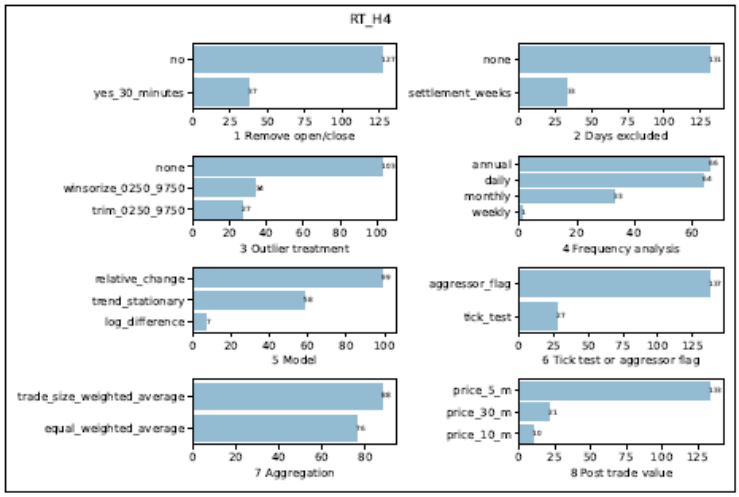



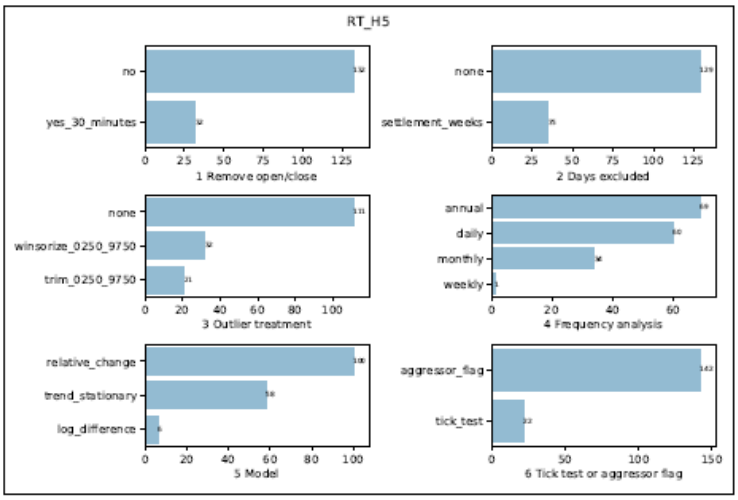
Reproduced:

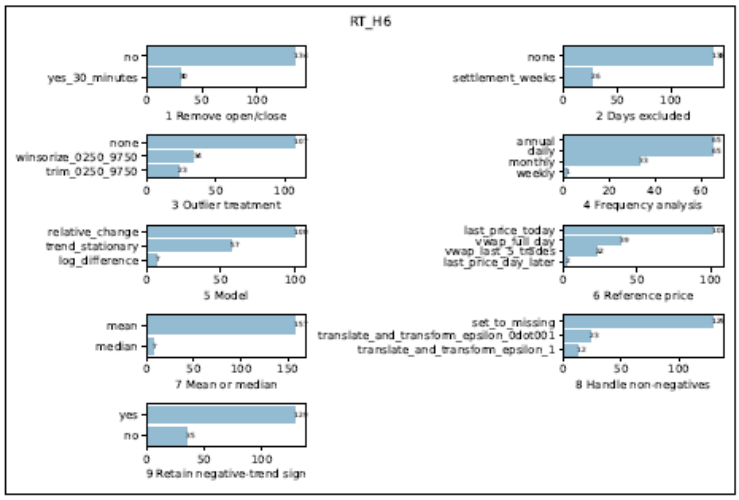

RT_H1
no
yes_30_minutes
1 Remove open/close
none
settlement_weeks
2 Days excluded
none
winsorize_0250_9750
trim_0250_9750
3 Outlier treatment
annual
daily
monthly
weekly
4 Frequency analysis
relative_change
trend_stationary
log_difference
5 Model
variance_ratio
auto_correlations
6 Measure
5_to_30_minutes
1_to_5_minutes
second_to_minute
day_to_week
day_to_month
7 Measure frequencies


RT_H2
no
yes_30_minutes
1 Remove open/close
none
settlement_weeks
2 Days excluded
none
winsorize_0250_9750
trim_0250_9750
3 Outlier treatment
daily
annual
monthly
weekly
4 Frequency analysis
relative_change
trend_stationary
log_difference
5 Model
aggressor_flag
tick_test
6 Tick test or aggressor flag
trade_size_weighted_average
equal_weighted_average
7 Aggregation
price_5_m
price_30_m
price_10_m
8 Post trade value


RT_H3
no
yes_30_minutes
1 Remove open/close
none
settlement_weeks
2 Days excluded
none
winsorize_0250_9750
trim_0250_9750
3 Outlier treatment
annual
daily
monthly
weekly
4 Frequency analysis
relative_change
trend_stationary
log_difference
5 Model
number_of_contracts
euro
6 Units


RT_H4
no
yes_30_minutes
1 Remove open/close
none
settlement_weeks
2 Days excluded
none
winsorize_0250_9750
trim_0250_9750
3 Outlier treatment
annual
daily
monthly
weekly
4 Frequency analysis
relative_change
trend_stationary
log_difference
5 Model
aggressor_flag
tick_test
6 Tick test or aggressor flag
trade_size_weighted_average
equal_weighted_average
7 Aggregation
price_5_m
price_30_m
price_10_m
8 Post trade value


RT_H5
no
yes_30_minutes
1 Remove open/close
none
settlement_weeks
2 Days excluded
none
winsorize_0250_9750
trim_0250_9750
3 Outlier treatment
annual
daily
monthly
weekly
4 Frequency analysis
relative_change
trend_stationary
log_difference
5 Model
aggressor_flag
tick_test
6 Tick test or aggressor flag


RT_H6
no
yes_30_minutes
1 Remove open/close
none
settlement_weeks
2 Days excluded
none
winsorize_0250_9750
trim_0250_9750
3 Outlier treatment
annual
daily
monthly
weekly
4 Frequency analysis
relative_change
trend_stationary
log_difference
5 Model
last_price_today
vwap_full_day
vwap_last_5_trades
last_price_day_later
6 Reference price
mean
median
7 Mean or median
set_to_missing
translate_and_transform_epsilon_0dot001
translate_and_transform_epsilon_1
8 Handle non-negatives
yes
no
9 Retain negative-trend sign