

Data Editor's Report 2024

Koren, Miklós

2024-09-19

This report covers the period between September 1, 2023 and August 31, 2024. In this period, 79 replication packages have been submitted and 90 packages have been accepted.

We now have four complete years of data under the new Data Availability Policy. This makes it possible to make comparisons and identify changes.

Turnaround times

A key performance indicator is the number of revisions a package goes through before acceptance.

The majority of packages are accepted only after revisions, but the number and share of packages accepted without revisions have increased. This year, 13 packages were accepted without revisions. Most packages are accepted on first revision. The Figure plots the number of packages by revisions over time. Around 78 percent of packages are accepted after at most one revision (approximately same as last year).

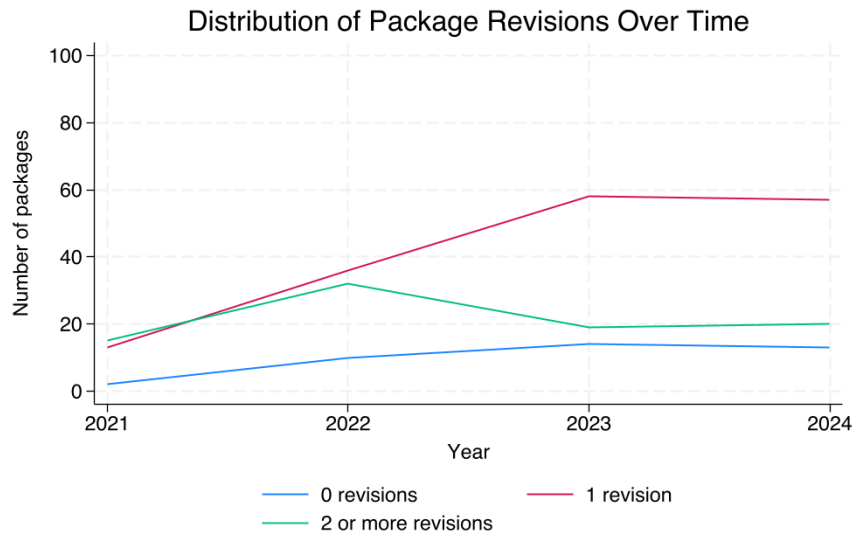


Figure 1: Three quarters of packages are accepted after at most one revision

The average package accepted this year received their first response from the Data Editor in 15 days (up from 12 days last year). This statistic excludes papers where the Data Availability Policy does not apply. For revisions, the mean response time is 6 days, with the majority of packages decided in much shorter time (see Figure).

The average time authors spend with a revision is 25 days.

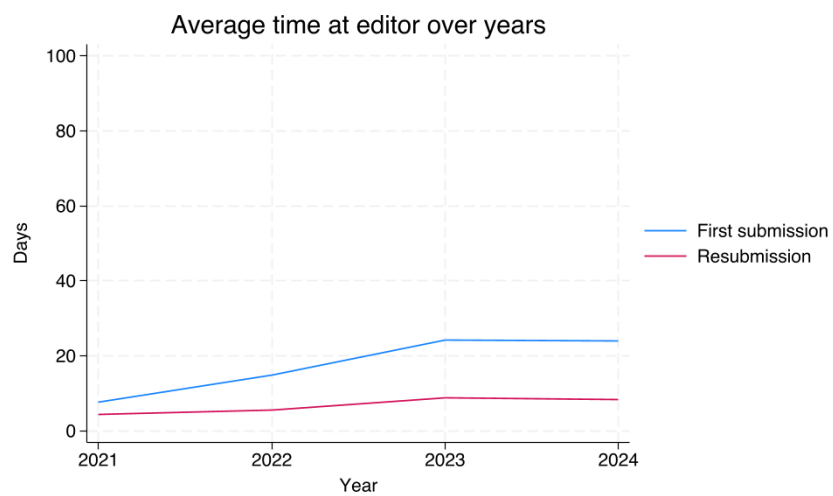


Figure 2: Average time at editorial team

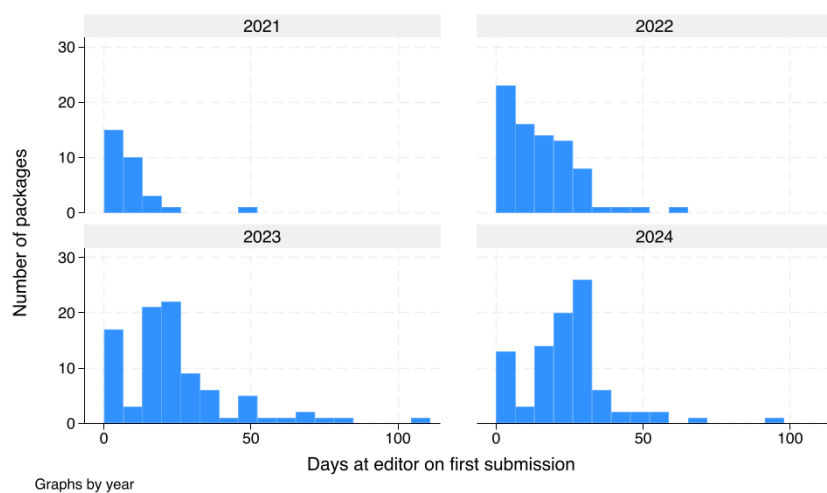


Figure 3: Distribution of time at editorial team

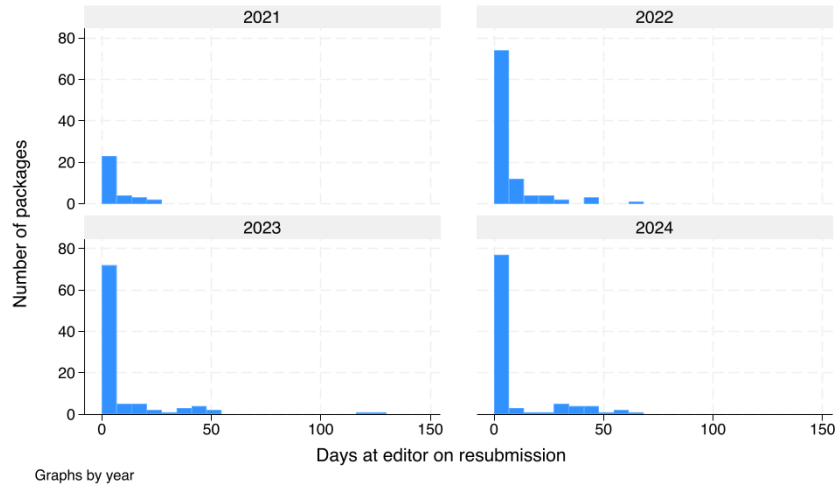


Figure 4: Distribution of time at editorial team

Reasons for revision

The Figure shows the main reasons for sending back the first version of the package for revision. About half of the packages miss data citations, a Data Availability Statement, or both. About a third of the packages do not save or include the reproducible exhibits to be verified. Specific Stata requirements are also often inadequately explained.

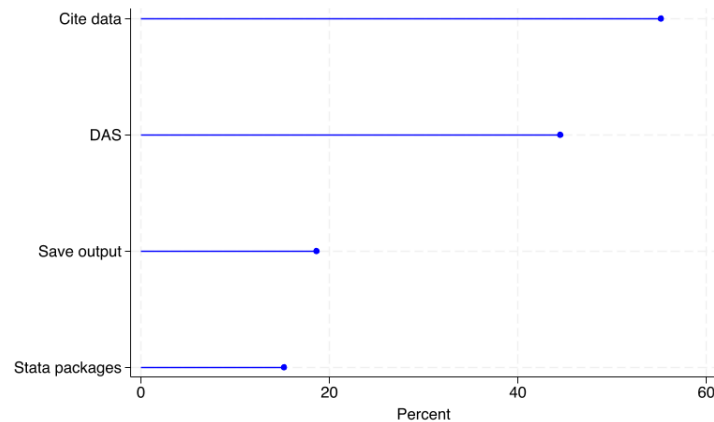


Figure 5: Main reasons for revision

Impact

The median replication package is downloaded from Zenodo 2 times a month. This includes views and downloads by author and editorial team. As the figure shows, there is substantial heterogeneity across packages in their download statistics, but all packages have some visibility.

Download statistics are highly correlated over time. Packages that were popular last year are often downloaded also this year (next Figure). This suggests that the download statistics capture genuine interest and not only driven by early downloads by authors and the editorial team.

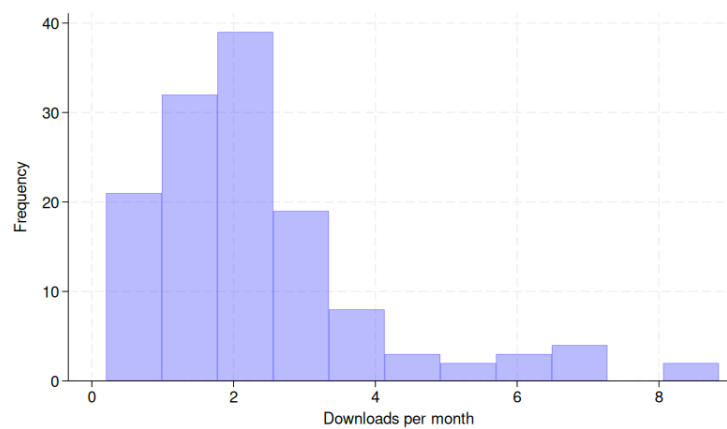


Figure 6: Unique downloads of packages from Zenodo

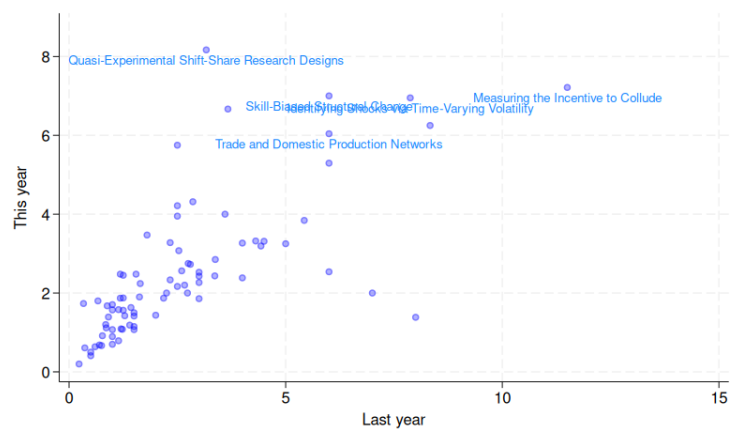


Figure 7: Downloads are highly correlated over time