

Peer Review

Review of: "A Statistical Investigation of the Synoptic Problem"

Shira Faigenbaum-Golovin¹

1. Bar-Ilan University, Israel

My main concern, which can also help to address the other issues, is:

- **Absence of Ground-Truth Validation:** The primary limitation is the lack of empirical validation on known, non-biblical manuscript trees. Without testing this specific Poisson/quasi-Poisson framework on synthetic or historical datasets with known editorial histories (e.g., medieval scribal adaptations or controlled modern rewriting experiments), it is impossible to verify whether relative vocabulary frequency alone can reliably differentiate a copyist from an independent user of a common source.

In addition,

Loss of Local Context: Aggregating the 807 word counts across entire Gospels strips away scene-by-scene context and narrative structure, treating distant occurrences as independent.

Residual Overdispersion: Global quasi-Poisson and negative binomial corrections struggle with extreme, localized topic clustering ("burstiness"), which can still artificially inflate p -values.

Function Word Bias: Predictions are heavily driven by high-frequency function words (e.g., $\acute{\upsilon}\mu\tilde{\iota}\tilde{\nu}$, $\tilde{\eta}\nu$), which often reflect genre or direct speech rather than direct manuscript copying.

Imperfect Baseline Categories: Assuming unique material (e.g., *Sondergut*) represents a pure "authorial fingerprint" ignores the impact of oral traditions or lost minor sources.

Directionality Assumptions: Inferring borrowing direction from vocabulary profiles overlooks asymmetrical editing, such as an author adopting source vocabulary into their own unique writing.

Rigid Data Categorization: Using strict word-matching toggles (present/absent) makes the input dataset overly sensitive to the concordance editors' specific decisions.

Declarations

Potential competing interests: No potential competing interests to declare.