Rotten at the Top: A Different Form of Inaccessibility Among Top Economics Journals*

Dustin R. White†

Jadrian Wooten‡

Abstract

The credibility of academic research increasingly relies on the accessibility of digital material hosted online. However, our investigation into the top five economics journals over a ten-year period reveals a concerning trend: 51.5% of external hyperlinks are no longer accessible, resulting in link rot—a phenomenon where hyperlinks no longer function or no longer lead to the originally cited material. This decay undermines the reproducibility of research and poses challenges to maintaining the integrity of the academic record. We find that the likelihood of encountering link rot escalates as the age of the publication increases, with links in articles older than eight years being about 30% more likely to be inaccessible.

JEL classification: A11, A12, Z00

Keywords: hyperlinks, links, posterity, link rot, accessibility, reproducibility

*We would like to thank seminar participants at the University of North Texas and session participants at the 2022 Association of Private Enterprise Education Conference. In particular, we would like to thank Yang Zhou and Ben Smith for helpful feedback on the early drafts of the paper.

†Department of Economics, University of Nebraska Omaha, email: drwhite@unomaha.com

‡Corresponding Author: Department of Economics, Virginia Polytechnic Institute and State University, email: jadrian@vt.edu

1 Introduction

Scientific misbehavior manifests in a variety of forms, ranging from the overt, such as intentional misrepresentation of data, methods, or results, to the subtler, such as the questionable division of research into smaller units for separate publication. High-profile cases involving researchers like Dan Ariely and Francesca Gino have heightened awareness of these issues, underscoring the importance of integrity in research practices (Thorp, 2023; O'Grady, 2024). While data fabrication and plagiarism are almost unanimously rejected, other behaviors remain divisive among economists and can sometimes only be recognized as inappropriate in hindsight (Necker, 2014). As the reproducibility crisis has drawn increased scrutiny in recent years, so too has the need to bolster the credibility of the discipline (Christensen and Miguel, 2018).

An underlying issue of transparency and reproducibility extends to how readers access referenced materials in research papers, especially as time passes. Traditionally, verifying references meant visiting physical repositories to consult the original materials. The internet has transformed this practice: hyperlinked references promise a direct connection to an author's sources, theoretically enabling readers to verify claims and access supporting materials instantaneously. This shift towards digital referencing should simplify the process of accessing information, but it also introduces new challenges, such as link rot, which threatens the long-term viability and reliability of these digital references.

Hyperlinks in scholarly articles are frequently hosted by third parties and their accessibility relies on that third party maintaining the relevant links. As organizations evolve, they may intentionally or unintentionally sever these links, disconnecting readers from the source material. Several scenarios can lead to such link decay: a third party might remove inaccurate material, the entity responsible for the source could cease operations, or the material might be updated, altering the information that was originally cited. Even if the material itself remains online, changes to the URL referencing that material can unintentionally disrupt the connection to the original reference. Each of these possibilities significantly diminishes

the reproducibility of research findings, as they compromise the accessibility of material underpinning scholarly work.

Unlike print references stored in library archives, digital references face unique challenges. As government agencies, universities, and organizations increasingly adopt primarily digital publications, these challenges have the potential to become more widespread. While few repositories currently maintain copies of digitally cited materials, one solution that has already gained traction is the use of digital object identifiers (DOIs). A DOI provides a persistent link to content, ensuring that each cited paper is perpetually accessible via its unique string of numbers and letters.

As academic reliance on digital dissemination grows, the challenges of link rot and reference rot become increasingly important. Our investigation into five leading economics journals[1] over ten years reveals a concerning trend: over half of all links are already non-functional, underscoring the urgency of addressing these issues in academic publishing. We explore this phenomenon, beginning with a brief review of previous studies on link rot in other disciplines. We then detail the methodology and data sources used to assess the prevalence of link rot and reference rot in prominent economic research. Finally, we discuss our findings, highlighting the extent of decay in digital references and proposing solutions to mitigate these problems in future scholarly work.

2 Previous Work

Previous researchers recognized the potential for link rot in the early 2000s as citation of online material began to increase. The first researchers focused on the prevalence of dead links in legal studies (Rumsey, 2002), but soon expanded to biochemistry (Markwell and

[1]The "top five" journals were based on Card and DellaVigna (2013): *American Economic Review*, *Econometrica*, *Journal of Political Economy*, *Quarterly Journal of Economics*, and *Review of Economic Studies*. While these journals may be among the most respected journals currently, Caviglia-Harris (2023) found that a decreasing number of highly cited papers are from these journals. Heckman and Moktan (2020) support this finding but note that the next generation of economists is obsessed with publishing in these journals.

Brooks, 2003) and general webpages (Koehler, 2004).[2] Much of the work related to link rot has focused on the prevalence of link rot in case law and Supreme Court rulings (Davis, 2006; Liebler and Liebert, 2012; Zittrain et al., 2014).

Although there has been some qualitative review of link rot in other disciplines, substantive policy recommendations are rare, and no studies have definitively measured the impact of these recommendations. However, Miguel (2021) highlights that economics, along with related quantitative social science fields, are rapidly adopting new norms to enhance transparency. For example, in 2020, the American Economic Association updated its "Data and Code Availability Policy." This policy now mandates that published articles include all data and code necessary for replicating the empirical results, provided these materials are not proprietary or otherwise restricted. If restricted, authors must ensure the preservation of their data and code for a minimum of five years. Although this policy marks a significant advance in promoting transparency, it does not directly address all the concerns associated with link and reference rot, but it does set a precedent for improving access to research materials.

Zittrain et al. (2014) recommend the use of Perma.cc to preserve the integrity of sources over the long run by storing a cached version of the cited material and issuing a permanent link. Once cached, the new link can be included in the published article along with (or instead of) the original hyperlinks, and ensure that the source material remains accessible in the event of the source becoming deactivated. Another possible solution is for journals themselves to host data files, cached files, or code.

3 Methodology and Data

Every article published between 2010 to 2020 in *American Economic Review*, *Econometrica*, *Journal of Political Economy*, *Quarterly Journal of Economics*, and *Review of Economic*

[2] A 2024 report by the Pew Research Center, titled "When Online Content Disappears," indicates that as of October 2023, a quarter of all webpages that were online at any point from 2013 to 2023 are no longer accessible.

Studies was included as part of our sample of 1,926 papers. Hyperlink information was extracted from each article and associated with a particular journal. Hyperlinked email addresses, internal URL references (links within the same document), JSTOR permalinks, and DOI addresses were excluded from the analysis. The sample only includes the first instance of a link in each article such that any article that may have referenced the same hyperlink multiple times is only counted once for that article.

We conducted HTTP status checks on each hyperlink in our sample to determine its functionality. An HTTP status code, sent from a webpage's server to the browser, indicates the browser's interpreted result of attempting to access a website. Appendix A summarizes relevant codes that could be generated. A '200' status code means the page loaded as expected, confirming the link is functional. Conversely, if the domain no longer existed, the status check would return a specific error such as 'OPEN', indicating a non-functional link. The most common error we encountered was the '404' code, signaling that the page was not found. This same code, however, could also be reported if the website is temporarily down at the time of our request.[3] Due to the extensive number of hyperlinks—6,892 in total—it was not feasible to manually verify if the content on each hyperlink remained unchanged from its initial citation. Therefore, our initial analysis primarily focuses on determining whether the links were operational at the time of the request.

HTTP status codes can reflect a variety of outcomes, indicating whether hyperlinks are functioning or not. While a '200' code confirms that a link is active, it does not guarantee the accuracy or relevance of the current content. It's possible for an active link to lead to a custom error page or to redirect to a different page, thus diverging from the content originally cited. Initial results from our status checks are documented in Table 1, highlighting the diversity of status outcomes in our study.

To further investigate the prevalence of reference rot among hyperlinks that appeared functional, we manually examined a random sample of 20% of the links that returned an

[3]To confirm that error codes indicating potential link rot was not due to temporary outages, each hyperlink that returned such an error was retested approximately one month later.

active status. This process involved checking 649 hyperlinks to verify whether the content accessible via these links still matched the content originally cited in the respective articles. The results of this detailed examination are presented in Table 2, which outlines the proportion of links within each journal that accurately maintained the intended content at the time of our evaluation.

While the findings discussed below provide insight into the prevalence of reference rot, they likely represent a conservative estimate of the true extent of link and reference rot in the top economics journals. Our analysis suggests that the observed rates of decay may serve as a lower bound, indicating that the actual prevalence of outdated or erroneous links could be even higher in recently published scholarly work.

4 Discussion

Throughout our study period, approximately 600 hyperlinks were published annually in articles across the top economics journals, as detailed in Figure 1. Collectively, 51.6% of the hyperlinks included in published papers over the ten-year sample were inaccessible. The prevalence of link rot, however, shows significant variation among these journals and over time. Our results are heavily influenced by hyperlinks that appear in articles from the *American Economic Review*, which make up nearly half of all the hyperlinks in our sample. This is not to insist that link rot is contained to only papers published there. Even the journal with the lowest incidence of link rot, *Econometrica*, still presented substantial accessibility issues, with approximately one-third of its links failing within the study period. These findings underscore the widespread and significant challenge of maintaining digital accessibility in academic research.

4.1 Inaccessibility by Journal

Each journal in our study publishes articles that include at least one hyperlink, yet the frequency of linked sources and the reliance on such digital references varies significantly across journals. This variation may reflect the editorial policies of the journals aimed at minimizing link and reference rot, or it could be influenced by the nature of the research typically published within each journal.

Table 1 presents the raw count of status codes for each hyperlink in our entire sample, broken down by journal. It highlights that the percentage of hyperlinks considered 'rotten'—those returning an error code other than '200'—varies considerably across journals. Notably, the *American Economic Review*, which constitutes nearly half of all hyperlinks in our sample, also features the highest incidence of link rot. This predominance significantly influences the overall results, suggesting a greater vulnerability among the hyperlinks in this journal.

It is important to remember that an active code does not necessarily confirm that the content accessed remains unchanged from what was originally cited. We believe our findings might underrepresent the true extent of inaccessibility. Further illustrating this issue, Table 2 examines a subsample of active hyperlinks to identify cases of reference rot, revealing trends consistent with those in Table 1. We note an exception with the *Journal of Political Economy*, where the fewest hyperlinks appear during our sample and it is hard to make meaningful conclusions from the small number of observations.

4.2 The Impact of Time

The frequency of hyperlinks in academic articles has significantly increased over the past few years, as illustrated in Figure 1. Despite this growing reliance on digital references, many of these links become inaccessible by the time the article reaches its readers, illustrating just how rapid the decay process can be. Given that the review process in economics is notably slower compared to other disciplines (Hadavand et al., 2024), lengthy delays mean that by

the time research is finally published, a substantial proportion of hyperlinks are already non-functional. This can be most clearly seen in Figure 2 as nearly 45% of hyperlinks that appear at the end of our sample (2020) are already rotten. It's important to remember that once a hyperlink becomes non-functional, it is rare for it to be restored; thus, broken links tend to remain so permanently.

Our results show a growing prevalence of link rot each year removed from an article's initial year of publication. After about 5-7 years, that prevalence plateaus, as seen in Figure 2. We see similar patterns when looking at the results for each journal. Figure 3 provides a visual representation of this decay, displaying a smoothed trendline for each journal and the overall trajectory of link rot across the top five economics journals. For journals that publish a sufficient number of links, a consistent decay pattern emerges. However, the Quarterly Journal of Economics stands out as an exception, exhibiting a stable rate of link rot from 2013 onwards, with no instances of link rot in articles published within the first three years of our sample. The persistence of link rot aligns with findings from other disciplines, suggesting that link decay is a common challenge across scholarly publishing.

4.3 Source of Inaccessibility

Our analysis deliberately excluded certain types of hyperlinks: permalinks (like DOIs), email addresses, and internal links that connect in-text citations to reference pages. Permalinks, primarily used for citation-tracking and identifying peer-reviewed works, along with internal links, generally offer stable connections and are not prone to decay. Email addresses simply do not offer the same kind of access to information. These exclusions allow us to focus on external hyperlinks, which are more susceptible to link rot.

We also examine the top-level domains of all hyperlinks in our sample to assess which types of sources are most likely to experience link decay. Table 3 shows the predominance of the various domains broken out across journals, with commercial (.com), government (.gov), and education (.edu) websites making up the largest categories. Our findings, illustrated in

Figure 4, reveal a significant disparity in the durability of links based on their domain type. Alarmingly, nearly 75% of hyperlinks directing readers to academic webpages were found to be rotten, contrasting sharply with a 30% to 40% rot rate observed in commercial domains.

This trend is most likely attributed to the transient nature of academic appointments: researchers frequently post important materials such as online appendices, data, or code on their institutional webpages, but these resources become inaccessible if the researcher leaves the institution. In contrast, commercial websites, while seemingly more reliable, often employ tactics such as redirect pages or custom 404 error pages that can misleadingly appear as functional during automated checks.

5 Conclusion

The prevalence of link and reference rot in leading economics journals over a ten-year period reveals a significant challenge to the reproducibility of academic research. We find that more than half of all hyperlinks referenced in these journals are no longer accessible, and nearly 25% of active links do not lead to the originally cited material. These discrepancies, which vary across journals, may reflect differing editorial policies or the nature of the researchers publishing in those outlets, but show consistent cause for alarm across all journals.

The implications of these findings are profound: the decay of references compromises the foundation of academic scholarship and makes it difficult for future researchers who wish to replicate or advance those findings to gather the requisite information. Adaptive policy solutions may be able to combat this issue going forward. Requiring archives of data and code can facilitate the replication of results, so long as the archives exist within the journal's servers. Making these resources available only on an institutional website, as we've seen, may be futile. There may be additional challenges for the material cited from external websites, where copyright restrictions often limit the ability of journals to host original third-party content.

To ensure the accessibility of referenced materials for future researchers, the adoption of permalinks and the promotion of durable digital repositories are critical. Although potentially costly, these strategies are investments in the reliability of academic work. Journals and academic institutions should collaborate to develop standardized practices for digital preservation and encourage authors to use more stable and verifiable sources (Zittrain et al., 2014).

Beyond institutional and editorial policy changes, the scholarly community must first recognize the urgency of addressing link and reference rot. Without buy-in from the academic community, the accessibility of academic content will wane with each passing year. As digital scholarship has evolved, so too must our approaches to maintaining its integrity.

6 Tables & Figures

Table 1: Count of hyperlinks returning a particular status code and percentage of rotten hyperlinks by journal

HTTP Status	Total	AER	ECN	JPE	QJE	ReStud
200 (working)	3,337	1,280	1,392	56	492	117
400	5	4	0	1	0	0
403	55	50	2	0	3	0
404	898	448	305	14	99	32
Other	2,039	1,185	369	51	362	72
Timeout	558	413	61	5	57	22
Total	6,892	3,380	2,129	127	1,013	243
Percent Rotten	51.6%	62.1%	34.6%	55.9%	51.4%	51.9%

Note: Each row shows raw counts of the number of links that return a particular status code during our queries. The results are divided by journal for articles published in *American Economic Review* (AER), *Econometrica* (ECN), *Journal of Political Economy* (JPE), *Qaurterly Journal of Economics* (QJE), *Review of Economic Studies* (ReStud). The percentage of hyperlinks determined to be rotten is based on the total number of links returned a code other than '200" as a share of the total number of links.

Table 2: Prevalence of reference rot among a subsample of active hyperlinks

Link Status	Total	AER	ECN	JPE	QJE	ReStud
Accessible Content	490	223	176	4	72	15
No Longer Accessible	159	56	74	5	19	5
Total	649	279	250	9	91	20
Percent Rotten	24.5%	20.1%	29.6%	55.6%	20.9%	25.0%

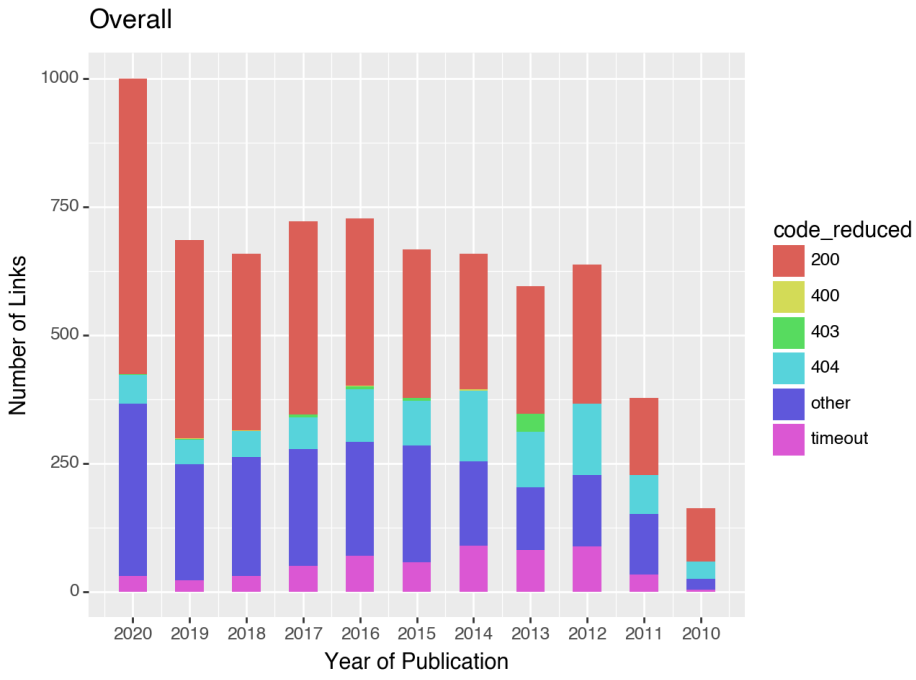
Note: We sampled 20% of all active hyperlinks that were found in papers published between 2010 and 2020 in the *American Economic Review* (AER), *Econometrica* (ECN), *Journal of Political Economy* (JPE), *Qaurterly Journal of Economics* (QJE), *Review of Economic Studies* (ReStud). We then manually inspected the content on these webpages to verify if the material cited in the paper was still present at the link. To calculate the proportion of reference rot, we divided the number of links with outdated or incorrect content by the total number of active links in our subsample.

Table 3: Count of hyperlinks by domain type by journal

Domain Category	Total	AER	ECN	JPE	QJE	ReStud
.com	1494	925	200	33	258	78
.gov	939	672	67	13	153	34
.edu	756	493	121	24	76	42
Other	3703	1290	1741	57	526	89
Total	6,892	3,380	2,129	127	1,013	243

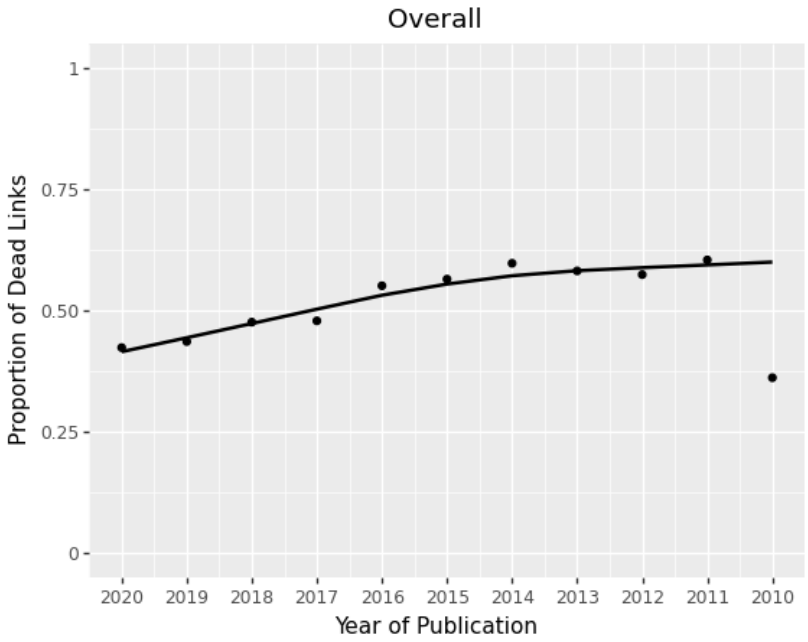
Note: Each row shows raw counts of the number of links that belong to a particular domain category. The results are divided by journal for articles published in *American Economic Review* (AER), *Econometrica* (ECN), *Journal of Political Economy* (JPE), *Qaurterly Journal of Economics* (QJE), *Review of Economic Studies* (ReStud).

Figure 1: Number of hyperlinks published in all articles from the Top 5 economics journals during each year.



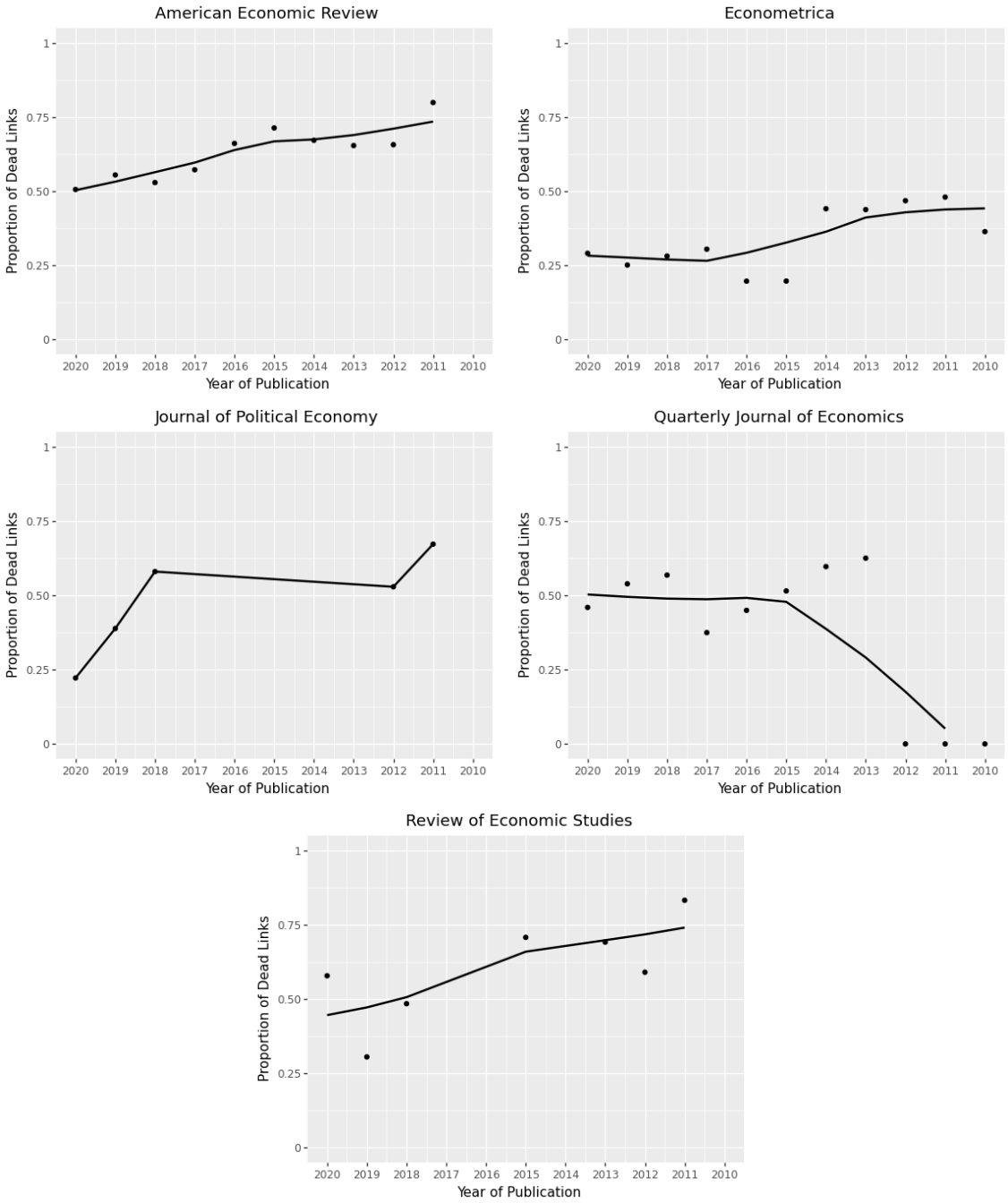
Note: A raw count of the number of hyperlinks that appear in articles published in the *American Economic Review* (AER), *Econometrica* (ECN), *Journal of Political Economy* (JPE), *Qaurterly Journal of Economics* (QJE), *Review of Economic Studies* (ReStud) based on the year in which the article was published. We further differentiate the hyperlinks in each year based on the status code returned during our query.

Figure 2: Proportion of inactive hyperlinks over time for all papers published between 2010 and 2020 in the Top 5 economics journals



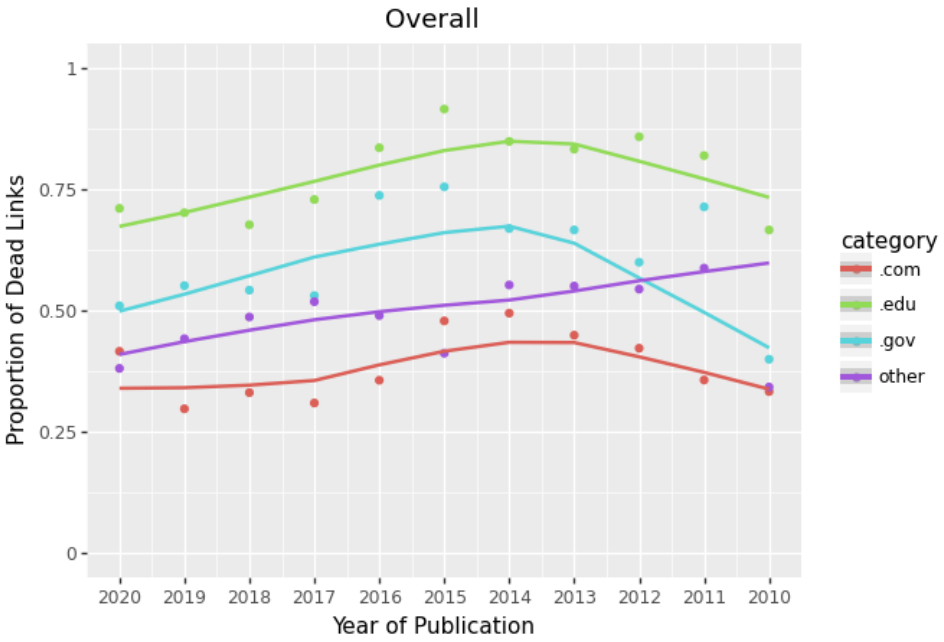
Note: All hyperlinks that appear in articles published in the *American Economic Review* (AER), *Econometrica* (ECN), *Journal of Political Economy* (JPE), *Qaurterly Journal of Economics* (QJE), *Review of Economic Studies* (ReStud) based on the year in which the article was published. We overlay a smoothed trendline based on these points.

Figure 3: Proportion of inactive hyperlinks over time for each of the top 5 journals



Note: Each graph shows the proportion of inaccessible links for a particular journal based on the year in which the hyperlink appears in an article. We overlay a smoothed trendline for each journal.

Figure 4: Proportion of each top-level domain that is rotten based on the year those links appeared in a Top 5 article



Note: Each point shows the proportion of links with a particular top-level domain (.com, .edu, .gov, and others) in a given year that are no longer accessible. We overlay a smooth trendline based on points within a particular domain class.

Appendices

A Summary of relevant error codes

The following codes were summarized by the Mozilla Developer Network.[4]

200 OK

The request succeeded. A successful response depends on the method used in the request, for example:

GET: The resource has been fetched and transmitted in the message body.

HEAD: The representation headers are included in the response without any message body.

PUT or POST: The resource describing the result of the action is transmitted in the message body.

TRACE: The message body contains the request message as received by the server.

400 Bad Request

The server did not process the request because of something that was perceived to be a client error. Examples that produce this error code include malformed syntax, an invalid request message framing, or deceptive request routing.

401 Unauthorized

This material requires that the client accessing the link be authenticated.

[4]https://developer.mozilla.org/en-US/docs/Web/HTTP/Status

403 Forbidden

The server refused the access request because the client does not have the appropriate access rights to the content at the hyperlink. Unlike a 401 Unauthorized error, the client's identity is known to the server.

404 Not Found

The server was unable to find the requested source because the URL was not recognized. It is possible that the serve sent this response instead of a 403 Forbidden error in order to hide the existence of a resource from an unauthorized client. This is the most well known error on the web.

405 Method Not Allowed

The request method was known by the server, but was not supported by the target resource.

410 Gone

This response was sent because the requested content has been permanently deleted from the server and no forwarding address exists.

500 Internal Server Error

The server encountered a situation that it did not know how to handle.

502 Bad Gateway

The server received an invalid response as it was working as a gateway to get a response needed to handle the request.

References

Card, D. and DellaVigna, S. (2013). Nine facts about top journals in economics. *Journal of Economic Literature*, 51(1):144–61.

Caviglia-Harris, J. (2023). Opening the gates: The increasing impact of papers beyond the top five and other changes in economic publishing. *Southern Economic Journal*, 90(2):474–496.

Christensen, G. and Miguel, E. (2018). Transparency, reproducibility, and the credibility of economics research. *Journal of Economic Literature*, 56(3):920–980.

Davis, H. E. (2006). Keeping validity in cite: Web resources cited in select washington law reviews, 2001-03. *Law Library Journal*, 98:639.

Hadavand, A., Hamermesh, D. S., and Wilson, W. W. (2024). Publishing economics: How slow? Why slow? Is slow productive? How to fix slow? *Journal of Economic Literature*, 62(1):269–293.

Heckman, J. J. and Moktan, S. (2020). Publishing and promotion in economics: The tyranny of the top five. *Journal of Economic Literature*, 58(2):419–470.

Koehler, W. (2004). A longitudinal study of web pages continued: a report after six years. *Information Research*, 9(2):174.

Liebler, R. and Liebert, J. (2012). Something rotten in the state of legal citation: The life span of a United States Supreme Court citation containing an internet link (1996-2010). *Yale Journal of Law & Technology*, 15:273.

Markwell, J. and Brooks, D. W. (2003). "Link rot" limits the usefulness of web-based educational materials in biochemistry and molecular biology. *Biochemistry and Molecular Biology Education*, 31(1):69–72.

Miguel, E. (2021). Evidence on research transparency in economics. *Journal of Economic Perspectives*, 35(3):193–214.

Necker, S. (2014). Scientific misbehavior in economics. *Research Policy*, 43(10):1747–1759.

O'Grady, C. (2024). Embattled Harvard honesty professor accused of plagiarism. *Science*, 384(6692):148–149.

Rumsey, M. (2002). Runaway train: Problems of permancence, accessibility, and stability in the use of web sources in law review citations. *Law Library Journal*, 94:27.

Thorp, H. H. (2023). Generative approach to research integrity. *Science*, 381(6658):587.

Zittrain, J., Albert, K., and Lessig, L. (2014). Perma: Scoping and addressing the problem of link and reference rot in legal citations. *Legal Information Management*, 14(2):88–99.