Journal of Management
Vol. 49 No. 8, November 2023 2541–2548
DOI: 10.1177/01492063231177785


# Using Old Data: When Is It Appropriate?

David J. Ketchen Jr.
*Auburn University*
Ashley Y. Roccapriore
*Auburn University*
Brian L. Connelly
*Auburn University*
*Deakin University*

*Researchers and gatekeepers lack clarity about the circumstances under which using old data to test hypotheses is appropriate or inappropriate. In response to this complex issue, we first define what makes data "old." We then suggest that using old data is justifiable (a) when examining a past event, (b) when recent data are not available, or (c) when the data were collected painstakingly. Scholars should avoid using old data if none of these conditions exist. Further, authors should be forthcoming about the age of their data and, in the case of a rejected journal submission, update the data whenever possible.*

***Keywords:*** *archival data; research methods; research design*

Publications are the currency of the academic world. Indeed, at many universities, management scholars are incentivized to publish as much as they can in high-quality journals to achieve promotions and salary increases. This creates a perverse situation whereby some scholars adopt a mercenary viewpoint—collecting acceptance letters becomes their goal instead of generating knowledge. One implication is that scholars ask, "What can I get away with?" too often when designing their research.

However, recent trends suggest that the mercenary research approach may not be rewarded with top-tier publications for much longer. The bar for publishing one's work within top journals is steadily moving in one direction, and it is not down. This places a premium on excellent research design. The answer to "What can I get away with?" is increasingly, "Not much."

*Corresponding author: David J. Ketchen Jr., Harbert College of Business, Auburn University, 415 W. Magnolia Ave., Auburn, AL 36849, USA.*

*E-mail: djk@auburn.edu*

Additionally, citation counts are growing in importance, especially when moving from being an associate to a full professor. Candidates for promotion to full professorship who accumulate a lengthy list of publications are increasingly likely to be disappointed by promotion votes if other scholars are not citing their articles. Mediocre ideas and flawed research designs occasionally get past review teams, but the resulting articles are unlikely to accumulate citations if readers recognize their weaknesses.

Significant pressure to create research with practical impact is coming, as signaled by the Association to Advance Collegiate Schools of Business (AACSB) recently adding the assessment of impact to their accreditation standards. Perhaps due to the field's "religious fervor" about theory (Hambrick, 2007: 1346), management scholars are late to this party. The need to demonstrate practical relevance is already entrenched in neighboring fields, such as finance and supply chain management, putting management at a deficit. Devising a research study that can deliver impact requires not just asking questions that managers care about but also testing our ideas with data that can provide actionable answers for managers.

Taken together, these three trends—a rising bar for publication, increased attention on citations, and pressure for practical impact—will require management scholars to shift away from a mentality of "What can I get away with?" toward one of "What is the best I can do?" As a first step toward defining the parameters of such a mindset, we consider a key design element: the age of a study's data. We call for researchers to use the most recent data they possibly can and outline when old data are—and are not—appropriate.

We focus our commentary on archival data because, in our experience, the most egregious attempts to use old data arise from studies that use archival data. By archival data, we mean measures gathered in ways that do not involve direct engagement between the research team and the observed entities (e.g., individuals, teams, firms). Examples include databases with ready-made measures (e.g., Panel Study of Entrepreneurial Dynamics, CompuStat), data that are collected by assessing publicly available resources (e.g., content analysis of media articles through websites such as Factiva), and data that are drawn from public websites (e.g., data pulled from social media platforms or organizational websites). As there is no direct researcher-observation engagement, using these data sets typically does not require the review approval of an institutional review board, unlike gathering primary data (e.g., surveys, experiments). Archival data therefore present researchers with the opportunity to explore phenomena and relationships without introducing new participants or researcher bias (Spector, 2006) but suffer from the problem that the sampling window might not be suitable for generating knowledge that informs today's organizations.

## Foundational Principles

Defining what is "old" depends on the context. A 15-year-old dog is old, but a 15-year-old human is young. A good starting point for deciding if data are old is considering whether leading popular press publications, such as the *Wall Street Journal*, would report the results of a study using the data. The answer, no doubt, varies depending on the context. A study centered on Twitter usage that relies on 3-year-old data might not appeal to popular press outlets, especially given the recent dramatic changes at Twitter. In contrast, a study of stock market reactions to spin-offs using data that end 3 years ago might appeal to press outlets, given that spin-offs are a relatively stable phenomenon over time. This latter

example suggests a second criterion that resonates with us: whether a study uses the most recent data available. In the case of a study into spin-offs, it would be reasonable to critique a research design for not being recent if 2 years of data were available but not used.

The age of one's data depends on both when the sampling window ended and when it began. Academics should build their empirical knowledge base with all the available information because, if it is handled appropriately, more information is better. However, there is often an extent to which going further back in time could yield less accurate or less relevant results if old data are simply bundled together with recent data. Consider, for example, a study on virtual teams with data that extend backward in time before the rapid ascent of videoconferencing due to the COVID-19 pandemic. While including older data offers a larger sample size, the age of the data from the early part of the sampling window could bias the results from the latter part of the sample. While archival data can sometimes reach back into the 1990s or earlier, researchers should consider whether adding samples from far in the past will make their data and analysis worse.

To determine the time boundaries of their data, researchers should factor into their decision-making the variables and relationships that are most central to their ideas. Researchers will sometimes terminate their data based on the availability of a control variable. We suggest that control variables should not be determinants of either the end or the beginning of a sampling window. If the dependent variable, predictors, and the nature of their relationship can support more up-to-date data, one could drop the control, use the control in a supplementary test with a limited sample, or use techniques to impute the missing control observations.

Last, researchers need to realize that the age of their data, as determined by the start and end of their sampling window, is fundamentally an issue of research integrity. Without a proper justification for these boundaries, researchers could conceivably alter their data's start or end dates to obtain the best results. Even if the end date was recent, unscrupulous researchers could play with the start date of their data, thereby making it older (but larger) or newer (but smaller) to optimize the outcome. Therefore, we suggest that samples should begin at the earliest possible point and end at the latest possible point that can be justified by the phenomena under investigation and the availability of the study's most important variables.

## When Using Old Data Makes Sense

There are situations where old data might be necessary and appropriate. Gatekeepers (editors and reviewers) should recognize and allow for this possibility. We discuss three of these scenarios next.

### *When Examining a Past Event*

Some events have an immediate major impact on organizations but then quickly cease to be causal agents. Studying these events requires restricting the time frame of a data set. For example, the Jumpstart Our Business Startups Act of 2012 made investing via equity crowdfunding legal for all U.S. citizens. Before its passage, the U.S. Securities and Exchange Commission allowed only wealthy people to fund ventures directly. A scholar asking the

research question, "To what extent, if any, does relaxing regulatory restriction of qualified investors increase or decrease crowdfunding success?" might find that there were strong effects immediately after 2012, but these effects dissipated over time as other causes of crowdfunding success interfered with them. Therefore, answering the question using data from 2012 to 2017 would be acceptable.

There is often value in historical analysis, such as learning how an industry emerged, grew, shrunk, and dissolved within a given period. Navis and Glynn (2010) explored how new markets emerged from 1990 to 2005 using satellite radio's introduction and evolution. The two competitors in this industry—Sirius and XM—merged in 2008. Current authors interested in studying the competitive dynamics between Sirius and XM could reasonably end their data in 2008. There can be value in studying an old phenomenon using old data if the phenomenon is generalizable as an event that could be repeated in the future, such as a competitive war in a duopoly.

### *When Recent Data Are Not Available*

Some archival data sets offer important insights but are not updated to the recent past. For example, the U.S. census is updated every 10 years. A scholar asking a research question requiring census data for an answer might be forced to use data up to 10 years old. A second example is the Pew religious data, which are updated every 7 years. Depending on when a study is initiated, scholars incorporating aspects of religion could be forced to use noncurrent data. A smaller yet still meaningful data collection interval surrounds the Global Entrepreneurship Monitor data set, which is updated every 3 years (Reynolds et al., 2005). In each of these three cases, researchers are constrained by the entity that gathers the data. Here, we suggest that using old data is reasonable, and we encourage gatekeepers to be flexible in their expectations about the age of the data. At the same time, authors should not assume that reviewers are aware of the occasional nature of data collection and should describe the relevant interval in their paper's Method section.

A second scenario occurs when a unique archival data set is no longer updated but it offers insight that more recent alternative data sources cannot. The Kauffman Firm Survey (KFS) is a good example. The KFS is a panel study that began data collection in 2004 and ended in 2011. KFS provides information on business formation for over 4,000 entrepreneurs, from their launch through the early years of operationalization, essentially following the firm's life cycle. While other data sets explore venture formation, the KFS particularly focuses on the launch process (Robb et al., 2009). KFS provides information on entrepreneurs who are engaged in starting and growing their ventures. It is reasonable, in our view, for researchers interested in variables that KFS includes and that are not available in other archival data sets to use KFS data despite their age.

### *When Data Were Painstakingly Collected*

A piece of advice we often share with doctoral students is to create a database that nobody else possesses. Doctoral students, after all, typically have considerable time, energy, and resources that they can devote to one specific project. Having detailed hand-collected data can allow them to develop unique insights that are unavailable to other scholars. In the

parlance of the resource-based view of the firm, a scholar accomplishing this task creates a strategic resource (i.e., one that is valuable, rare, and difficult to imitate or substitute) that can, in turn, create a competitive advantage in the marketplace of ideas (Barney, 1991). We believe that authors should be able to leverage such data without mandating that they update it, which would engender a laborious process. Imposing such a mandate could prevent authors from stepping away from data collection to focus on the equally important tasks of theory development and analysis.

Trickier situations arise when new archival data are available on an almost daily basis, such as social media data that must be drawn from websites. Take, for example, Roccapriore and Pollock (in press), who used data from 2019, 4 years before their article's publication date. Their data were painstakingly scraped from the social media platform Instagram. Given how social media platforms and their corresponding algorithms change, reviewers could have considered data from 2019 to be too old and demanded an update at the revise-and-resubmit stage. No such requests were made, which we believe was a just and fair approach. Given the length of the publication process and the difficulty of collecting this sort of data, gatekeepers should be flexible in making demands of authors about the age of data that were especially difficult to collect.

## What to Avoid

### *Spaghetti Syndrome*

Cooks can tell if spaghetti is ready to eat by throwing it against the wall and seeing if it sticks. A similar phenomenon occurs when authors take readily available data, fish around looking for significant relationships between variables, and write a paper about what they find. Sometimes these data sets were gathered years ago and used in past publications, but the author hopes to squeeze one more article out of the data despite the passage of time.

While spaghetti-flinging authors risk harm to their reputations among editors, this behavior is encouraged because journal submissions involve little cost to authors. Changing the incentives might help (Kerr, 1975). For example, submission fees could go a long way toward curing spaghetti syndrome because few authors would pay a fee for a bad paper to get reviewed. Historically, discussing submission fees has been like a politician pondering changes to social security: sure to elicit gasps. However, journals could waive fees for authors who lack financial support or refund fees for authors who earn a revise-and-resubmit request. The nuances of such a system are beyond this editorial's scope, and no system would be perfect, but we believe there is merit in the academic community at least discussing mechanisms that could help to reduce the amount of not-fully-cooked spaghetti being flung at journals.

### *Gamesmanship*

We have grouped these tactics because they appear to be based on the same underlying logic. It goes like this: "Let's omit this information for now. Reviewers probably won't consider this as a rejectable offense, so we are not hurting the paper's chances of moving forward. If we are invited to revise the paper, we will likely be asked about the age of our data. The

paper has limited space, but we can justify our old data at length in our responses to reviewers. Part of that justification will be a friendly and subtle reminder to the review team to assess the overall design. This will help us argue that while our use of old data can be viewed as a weakness, it is not a fatal flaw, and the stronger elements of our design offset this weakness."

We must admit that the logic behind this gamesmanship is sound, even if the ethics are questionable. We are mindful, however, that such tactics will not sit well with reviewers. It is frustrating to conscientiously review a paper that you think might warrant publication only to discover at the revision stage that the authors cut corners by needlessly using old data. What's the solution? Simply, editors need to ensure that adequate information about the sample is provided before sending a paper to reviewers. If there are strategic omissions in the information provided, the action editor can ask the authors to add the relevant material before seeking reviewers for the paper.

### *Recirculation*

Urban legend says that if a frog is placed in a pot of tepid water that is slowly heated, the frog will not realize the mounting trouble and will be boiled to death. A similar phenomenon can arise when a study based on archival data is rejected by a series of journals. The data set may have initially been current, but after several years and numerous rejections, it becomes outdated. As the data become incrementally older with each rejection, a study's authors may never realize that their chances of publication are being boiled away. The solution is straightforward: Whenever a journal submission gets rejected, authors should check to see if new data are available, and if so, they should update their data set.

An anecdote from 2023 is instructive. An associate editor conveyed a "desk reject and resubmit" decision on a study that relied on publicly available—and easily updatable—archival data on initial public offerings (IPOs). The data ended in 2015, with little justification for this endpoint. A quick internet search revealed that a paper with the same title was presented at a conference in 2016. It did not take Sherlock Holmes to deduce what had happened. The authors were asked to bring their data set forward as many years as possible before the paper would be reviewed.

A good question for authors and gatekeepers to ask in this situation is, "Would I feel confident about providing guidance to students in an MBA class about IPOs based on data that end in 2015?" To the associate editor, the answer was a firm no. As an applied field, we want to confidently give those students actionable insights (cf. Craighead, Ketchen, & Darby, 2019). Adding years of data has definite costs in terms of labor and logistics, and maybe even money, but we should test hypotheses and provide conclusions in ways that include the most current archival data available.

Conversely, a deceptively inadequate question for authors and gatekeepers to ask in this situation is, "Has the phenomenon changed over time?" It is a fair question, but answering it in the affirmative is a necessary but insufficient condition for determining if old data are too old. If the phenomenon has presumably not changed, but new data are readily available, then authors should endeavor to use the most complete and recent information to show definitively that the phenomenon has not changed.

Relatedly, based on submissions we are seeing at the *Journal of Management*, some authors are beginning to stop their data collection in 2019 so that their results will not be

influenced by the COVID-19 pandemic. Doing so, however, can be counterproductive as authors miss an opportunity to establish boundary conditions on their theorizing (Busse, Kach, & Wagner, 2017). Instead, they should gather data from the years of the pandemic and the following years. If including the COVID-affected years creates empirical problems, authors can exclude these years from their hypothesis-testing analysis along with an explanation of why they are doing this. But, in exploring the COVID years, authors might find something interesting and worthy of discussion.

## Old Primary Data

Although primary data collection methods, such as surveys and interviews, lie outside the scope of our discussion, we would be remiss if we did not acknowledge that similar issues surround old data collected via these techniques. These methods are used to gather information about informants' perceptions of the working world. A good screen for authors is to consider whether the underlying phenomenon they investigated has changed since their primary data were collected. For example, perceptual data about working from home gathered in 2019 would not necessarily be chronologically old, but the subsequent quarantines triggered by COVID-19 changed many workers' and managers' perceptions about working from home. In this instance, gathering new data and comparing it with the 2019 data might be revealing.

However, old primary data can be appropriate when they supplement archival data. One of the main downsides of archival data is that they do not allow scholars to investigate causal mechanisms. Methods that rely on primary data, such as interviews and experiments, can be used in conjunction with archival data to address this weakness. It is reasonable to create multimethod designs where older primary data are supplemented with subsequent archival data to investigate causality in a unique way. A pioneering example of this is Thomas, Clark, and Gioia (1993), where the authors test a model linking scanning, interpretation, action, and performance. Scanning and interpretation were measured using a 1987 survey of hospital chief executive officers, action was measured using archival data from 1987 to 1989 on what services each hospital offered, and performance was measured using archival data from 1990 and 1991. The research design would have been less rigorous if scanning and interpretation were measured using archival data or if action and performance were assessed subjectively.

## Conclusion

In his 2009 editorial missive in the *Academy of Management Journal*, Duane Ireland posed and answered the question, "When is a new paper really new?" Ireland provided an important service for the organizational sciences by laying out specific criteria for deciding whether a paper is new and thus appropriate to enter into the review process. We endeavored to offer the same value by posing and answering a parallel question: When are old data too old? Our view is that using old data is appropriate (a) when examining a past event, (b) when recent data are not available, or (c) when the data were collected painstakingly. We suggest that studies should avoid using only old data if none of those three conditions exist. In reporting a study, authors should be transparent about the age of their data, not opaque. When journal submissions are rejected, authors should update their data whenever

possible. Simply, when it comes to using old data, authors should not ask what they *can* do but what they *should* do.

## References

Barney, J. 1991. Firm resources and sustained competitive advantage. *Journal of Management,* 17: 99-120.

Busse, C., Kach, A. P., & Wagner, S. M. 2017. Boundary conditions: What they are, how to explore them, why we need them, and when to consider them. *Organizational Research Methods,* 20: 574-609.

Craighead, C. W., Ketchen, D. J., & Darby, J. L. 2019. Taking AIM at theoretical and pragmatic impact: A call for actionable, insightful, and measurable research. *Journal of Business Logistics,* 40: 289-298.

Hambrick, D. C. 2007. The field of management's devotion to theory: Too much of a good thing? *Academy of Management Journal,* 50: 1346-1352.

Ireland, R. D. 2009. From the editors: When is a "new" paper really new? *Academy of Management Journal,* 52: 9-10.

Kerr, S. 1975. On the folly of rewarding A, while hoping for B. *Academy of Management Journal,* 18: 769-783.

Navis, C., & Glynn, M. A. 2010. How new market categories emerge: Temporal dynamics of legitimacy, identity, and entrepreneurship in satellite radio, 1990-2005. *Administrative Science Quarterly,* 55: 439-471.

Reynolds, P., Bosma, N., Autio, E., Hunt, S., De Bono, N., Servais, I., Lopez-Garcia, P., & Chin, N. 2005. Global entrepreneurship monitor: Data collection design and implementation 1998-2003. *Small Business Economics,* 24: 205-231.

Robb, A., Ballou, J., DesRoches, D., Potter, F., Zhao, Z., & Reedy, E. J. 2009. *An overview of the Kauffman Firm Survey: Results from the 2004-2007 data*. https://mathematica.org/-/media/publications/pdfs/labor/kfs_overview.pdf.

Roccapriore, A. Y., & Pollock, T. in press. I don't need a degree, I've got abs: Influencer warmth and competence, communication mode, and stakeholder engagement on social media. *Academy of Management Journal.*

Spector, P. E. 2006. Method variance in organizational research: Truth or urban legend? *Organizational Research Methods,* 9: 221-232.

Thomas, J. B., Clark, S. M., & Gioia, D. A. 1993. Strategic sensemaking and organizational performance: Linkages among scanning, interpretation, action, and outcomes. *Academy of Management Journal,* 36: 239-270.