

# Feet of Clay? How to Review Political Science Papers that Make Use of the Work of Historians

Jørgen Møller, Aarhus University, Denmark

**ABSTRACT** Political scientists increasingly enlist the work of historians but they often treat this work in a nonchalant or superficial way, which makes their evidentiary record questionable. It follows that we need to check the validity of the interpretation of historians' work in review processes. This article argues that enlisting historians as reviewers is not the answer. Instead, it proposes four simple criteria that can be used to flag situations in which the use of historians' work as empirical evidence is unconvincing. The general purpose of the article is to increase awareness about what is at stake when political scientists base empirical analysis on evidence gathered by historians.

Political science is in the midst of an ambitious “historical turn” (Capoccia and Ziblatt 2010; Mahoney and Thelen 2015). A key aspect of this embrace of historical analysis is that political scientists increasingly enlist the work of historians. First, political scientists sometimes use the prior work of historians to develop theoretical insights. For instance, Stasavage (2010) revisits work by Wim Blockmans to develop a hypothesis that in Medieval and Early Modern Europe, there were significant geographical barriers to political representation. Second, political scientists regularly use prior work of historians to apply their theoretical insights. This can be done by either (1) coding historical datasets based on historians' research, which then can be leveraged in quantitative analysis (e.g., Abramson and Boix 2019; Blaydes and Chaney 2013; Stasavage 2010); or (2) using historical works to create historical narratives, to make qualitative historical comparisons, or to carry out within-case process tracing (Lange 2012).

Using insights from historians to theorize is relatively unproblematic, especially if these theoretical insights are then assessed against new empirical evidence (Coppedge 2012, 148–49). However, if the empirical-evidence claims of political scientists are based on information gleaned from the work of historians, then these claims *ipso facto* hinge on the validity of the interpretation of this prior historical work, irrespective of which method of data analysis is used to process the information. Validity should be understood in a relatively undemanding way. The question is not whether representations of the past are true or false but rather would they pass muster—at least in a minimum sense in which the overall argument is not invalidated by correcting misunderstandings—for those well acquainted with the historical literature on the subject?

If this point is conceded, then it follows that editors and/or reviewers need to check this. By way of analogy, if the evidence claims in a political science article rest on advanced quantitative analysis, editors normally will seek out a methodologist or a statistically versatile specialist who can at least judge whether something looks suspicious. The problem is that editors of political science journals do not have a similar cohort of reviewers who can easily tell whether political scientists' reading of the prior work of historians looks suspect. What do we as a profession do in this situation?

This article briefly describes what is at stake when political scientists enlist the work of historians. On this basis, I discuss what type of check we need to avoid the claims about historical evidence made by political scientists resting on “feet of clay.” I argue that what might seem to be the obvious answer—that is, using historians as reviewers—creates more problems than it solves. Instead, I propose four simple criteria that reviewers and editors can use to assess historical-evidence claims—even if they do not know the historiography details on a particular subject. These criteria include basing historical-evidence claims mainly on (1) work by historians rather than social scientists; (2) historical work that does not seem blatantly outdated; (3) an awareness of differences within historiography on a particular subject; and (4) explicit page numbers in references. Two recent articles are used to illustrate these criteria.

## WHAT DO WE (REALLY) KNOW ABOUT THE PAST?

Political scientists who use historical analysis first need to recognize how uncertain our knowledge about the past often is. For many historical periods and contexts that political scientists have begun to work with as part of the ongoing historical turn (e.g., the Early Middle Ages in Europe or Ancient China), the only firsthand written sources are a small number of miscellaneous documents such as law-codes, annals, and charters—which, at best, present partial information about the societies in question.

Jørgen Møller  is professor of political science at Aarhus University, Denmark. He can be reached at [jm@ps.au.dk](mailto:jm@ps.au.dk).

Most information often comes from either archeology, which is difficult to interpret, or later secondhand narrative sources, which clearly—and often intentionally—misrepresent the past (Feng 2008; Loewe and Shaughnessy 1999; Wickham 2005).

To illustrate the problems that this creates, I begin with historian Chris Wickham's (2009, 12–18) discussion about the use of these secondhand narrative sources—in this case,

say about its archeology has multiplied tenfold—in some countries, a hundredfold.” This means that historical interpretations that factor in archeological findings quickly become outdated. Wickham (2005, 10) makes clear that his own attempt to synthesize our knowledge about the Early Middle Ages—with the bulk of the evidence collected in 1997–2000—will first become dated in this respect. Therefore, political scientists who enlist older work

*The question is not whether representations of the past are true or false but rather would they pass muster—at least in a minimum sense in which the overall argument is not invalidated by correcting misunderstandings—for those well acquainted with the historical literature on the subject?*

from Early Medieval Europe. Wickham grants that many of these sources may be completely fictitious: examples include Gregory of Tours's *History of the Franks* and Paul the Deacon's *History of the Langobards* (see also Wickham 2005, 116–17). Moreover, even if they are only partly fictitious, we have little against which to check the information. However, Wickham assures us, this does not mean that these are not good sources for understanding the society for which they claim to write the history. Their moralizing observations thus are anchored in “recognizable experience.” At a minimum, therefore, these sources tell us about “the sort of things that could happen” (Wickham 2009, 14) in the world of the early Franks or the early Lombards. This insight guides how Wickham (2005; 2009) uses this narrative evidence: not as a source of actual historical facts but rather as a source of what type of potential actions and reasoning characterized these societies.

The few and scattered firsthand sources are associated with other problems. What we often have are copies of copies, which makes it difficult to tell which part of the documents are original and which parts are later (i.e., secondhand) additions (Halsall 2005, 61). Moreover, legal texts—one of the few relatively abundant sources of firsthand written material in many societies—often are better guides to the ambitions or values of elites than to the actual state of society (Wickham 2005, 95–6).

Three points follow from this. First, political scientists will not be able to eliminate trained historians as gatekeepers to the past—at least if they are dealing with premodern historical contexts. Not only does accessing the sources normally require specialist skills (e.g., a command of medieval Latin or training in locating relevant text in archives), the written material also cannot be taken at face value but rather must be interpreted by someone who understands the biases of different types of sources and who can place the interpretation within a particular historical context.

Second, political scientists must factor in the uncertainty of any historical interpretation. Here, the third historical source mentioned previously—archeology—comes into play. The interpretation of archeological findings obviously also requires expert skills and often involves creative leaps in the form of qualified guesses. However, another issue concerns the dynamic nature of our archeological knowledge. Wickham (2005, 1) justified his magisterial attempt to revisit the Early Middle Ages in Europe with the fact that during a couple of decades, “[w]hat we can

of historians—including work that may be only a few decades old—must probe whether these interpretations still pass muster among historians or whether new archeological knowledge has made them obsolete.

Third, two additional issues further complicate matters. First, political scientists face the problem that there often are competing historical interpretations based on different perspectives on history. This means that they must become acquainted with the wider historiography to understand the different positions on a particular subject (Lustick 1996, 613). Second, among these competing historical interpretations, political scientists will be prone to emphasize those that confirm their theoretical arguments at the expense of those that do not accord well with them. This problem concerns confirmation bias in historical research (Lustick 1996, 608–10; Møller and Skaaning 2018, 3–4), which is simply a more particular aspect of the general tendency that people find what they are looking for. This further underscores why we need to check the validity of the evidentiary record in historical analysis.

#### HOW TO CHECK THE READING OF THE WORK OF HISTORIANS?

That political scientists often treat history in a nonchalant or even superficial way is hardly a novel claim (e.g., Goldthorpe 1991; Kreuzer 2010; Lustick 1996; Møller and Skaaning 2018). Yet, based on my experience as both an author submitting papers with historical analysis and a reviewer of such papers (often with access to other reviewers' comments), the problem seems to be widely ignored by the profession.

It might appear that the solution is simple. Returning to my initial analogy of the methodologist, we surely need to enlist at least one trained historian as reviewer to determine whether the reading of the work of historians passes muster. However, this solution has unfortunate side effects. First, as any political scientist who has submitted a historical analysis to a trained historian for comments knows, historians use different criteria for good research. Whereas political scientists normally try to generalize, compare, and present parsimonious explanations, historians focus on the specific and “gleefully proliferate variables all the time” (Gaddis 2002, 58). Gaddis (2002, 56–58; see also 64, 88–91) presents this as the difference between the *reductionist* models of social scientists and the historians' *ecological* approach. Even if historians were instructed that a paper should be reviewed

as a product of political science, they almost certainly would be too critical of attempts to use historical material to generalize or make huge comparisons across space and/or time as well as the concomitant lack of attention to detail. Second, as previously mentioned, there are different schools within history, and representatives from one school (e.g., historians working in the Marxist tradition) would be overly critical regarding interpretations based on the work of other schools (e.g., historians working in the *Annales* tradition or political historians) (Hexter 1979, 61–145; see also Lustick 1996, 615–6; Møller and Skaaning 2018, 2–3). Editors of political science journals would be unlikely to factor in these differences between historical perspectives when reading reviews.

I believe the problem calls for another solution, which comes in two parts. First, and perhaps most important, editors and reviewers must be conscious of the fact that the use of historians' work—whether for qualitative or quantitative purposes—cannot be taken simply at face value but instead needs to be questioned like any other type of data enlisted by political scientists. Second, we need a list of simple criteria against which historical work by political scientists can be assessed at minimal cost.

*... editors and reviewers must be conscious of the fact that the use of historians' work—whether for qualitative or quantitative purposes—cannot be taken simply at face value but instead needs to be questioned like any other type of data enlisted by political scientists.*

#### FOUR CRITERIA

I am not proposing a checklist that papers will either pass or fail. Rather, it should be perceived as a quickly applied set of criteria that can help reviewers—who are providing a community service—and editors identify whether historical claims seem tenuous. The metaphor is a warning signal, not a red light. If there is a warning, it often would make sense to raise this in reviews and, if the decision is to revise and resubmit, to ask authors to better substantiate and/or defend their reading of the work of historians, if necessary, in online appendices. I propose the following four criteria:

- Are evidence claims about historical facts mainly based on work by historians rather than by social scientists?
- Are evidence claims about historical facts mainly based on relatively new rather than relatively old historical sources?
- Is there discussion or at least awareness of different historical interpretations and an attempt to adjudicate between these?
- Does the author use page references for specific claims about historical facts?

These four criteria are not exhaustive; however, they are easy to apply and should flag blatant cases in which authors need to buttress their historical claims by invoking more relevant and/or unbiased historical data. The first three criteria are simpler versions of the more demanding criteria presented by Lustick (1996, 615–6) and Møller and Skaaning (2018, 6–13). Their criteria cannot be used in the present form because they presuppose precisely what editors and reviewers normally

lack—namely, a detailed knowledge of the historiography of a particular subject.

I previously clarified why we need the second and third criteria. The fourth—specific page numbers when referencing historical material—is particularly important when the claims that are invoked are not the central arguments or findings of the source but rather peripheral to it—or perhaps even made implicitly. In this case, the lack of page numbers makes it extremely cumbersome for readers to check these claims—or to read about them for their own use. Moreover, it is an affront to the historians who are invoked for evidence, considering that historians normally cite the particular pages for their evidence claims (Trachtenberg 2015, 13).

The first criterion—basing evidence claims mainly on work by historians rather than by social scientists—is a way to diminish the risk of confirmation bias by steering clear of a theoretically polluted selection or even reading of the evidence that historians initially produced (Møller and Skaaning 2018, 8–11). This criterion is not meant to imply that the work of historians is unbiased; it also is affected by implicit theories, and historians often use social science concepts or theories that

have long since been abandoned elsewhere. However, all else equal, work by historians has not to the same extent preselected evidence to provide support for *one* central theoretical claim (Gaddis 2002, 56–58), which means that it is likely to be less polluted by the theoretical vantage point. Moreover, this criterion can be seen as a prudent way to check whether historical claims first encountered by reading other social scientists can be corroborated based on work by historians, including more recent work.

To illustrate how the four criteria can be applied, I use examples from two recent public-choice articles on the “Rise of Europe”: Salter (2015) and Salter and Young (2018).<sup>1</sup> Salter (2015, 725) offers “a novel interpretation of the development of Western political institutions.” His claim is that when political authority is bundled with economic property rights, the result is a “shareholder state,” which he sees as responsible for the development of the European “nonpredatory governance” (Salter 2015, 726). The main aim of his article is to present this theoretical argument. But Salter (2015, 728–9) makes clear that it is based on empirical underpinnings—that is, something resembling the “shareholder state” existed in Medieval Western Europe.

On what type of historical evidence does he base this claim? Salter (2015, 729) references the following authorities: “Anderson (1991); Baechler (1975); Benson (1990); Berman (1983); Raico (1994); and Stark (2011, chaps. 14–6).” Only one of these scholars is a general historian (i.e., Baechler 1975); another is a legal historian (i.e., Berman 1983). Both works are clearly dated and, throughout Salter's article, there is no attempt to provide page references for the historical claims being made—this despite the fact that the main point of these sources is surely not that medieval monarchies resembled shareholder states. Neither is there

any discussion about differences within historiography or even among the social scientists who are referred to for evidence about the medieval context. In a sense, Salter seems to take for granted that the medieval political order in fact did resemble his ideal type in which claims to political power and revenue derived from property rights.

*... when participating in review processes either as author or reviewer, it often has seemed to me as if the problems identified in this article are simply “not on the radar” of those who assess work that is based on historical evidence.*

Salter and Young (2018) focus on the ways in which medieval representative institutions contributed to the Western European traditions of limited government. They ask why “the establishment of those traditions were uneven across countries” (2018, 173). To answer this, they revisit and elaborate on the German historian Otto Hintze’s (1962 [1930], 173) seminal argument “that two-chamber representative assemblies were more effective at resisting absolutist tendencies than three-chamber assemblies.” When presenting Hintze’s claim, Salter and Young (2018, 178, fn. 14) use political scientist Thomas Ertman’s (1997, 20–22) version of it.<sup>2</sup> When they describe the functioning of two- and three-chamber assemblies, respectively, they mainly refer to outdated historical evidence, including Bloch 1968a [1939], pp. 145–162; Vinogradoff 1968 [1922], p. 180; Lord 1930, p. 37; Myers 1975, p. 64 (fn. 20). Although Salter and Young should be praised for using page numbers, it seems fair to say that there is no attempt to probe the historical underpinnings of Hintze’s argument based on recent work by historians. In general, Salter and Young (2018, 174) “take as historical givens the two types of assembly structures.”

In both cases, the question is why, during the review process, were the authors not tasked to provide firmer evidence in favor of these sweeping historical claims about the medieval context? Does the shareholder-state description of Medieval Europe and Hintze’s old claim about the institutional differences between the two categories of representative institutions, respectively, hold sway among historians and, if so, is it also corroborated by the most recent evidence? After reading the papers, we have almost no clue about the validity of the evidentiary record. As the historian J. H. Hexter (1979, 243–48) once critiqued the work of one of his colleagues, what we are left with is “an attractive hypothesis treading water, as it were, waiting for the solid ground of proof to be shoved under it. And that is not good enough.”

## CONCLUSIONS

I make a claim that is hardly bold: If the evidentiary record of political scientists is based on the reading of prior work by historians, then we need to check the validity of these data during the review process. However, the reason that I try to push in what may seem to be an open door is that this is not part of best practice today. As mentioned previously, when participating in review processes either as author or reviewer, it often has seemed to me as if the problems identified in this article are simply “not on the radar” of those who assess work that is based on historical evidence. The main purpose of this article is to alert the political science profession about this problem. More specifically, I argue

that the solution is not to enlist historians as reviewers. Rather, we need simple criteria against which to flag problems in work by political scientists that invokes evidence from historians. The four criteria presented are easily applied, and two recent publications illustrate how they can be used to encourage authors to do a better job when using historical evidence.

## ACKNOWLEDGMENTS

I am indebted to Svend-Erik Skaaning and an anonymous reviewer at *PS: Political Science & Politics* for helpful comments. Any errors are mine. ■

## NOTES

1. For other examples of an unconvincing use of the prior work of historians, see Lustick (1996, 608–10), Kreuzer (2010), and Møller and Skaaning (2018, 16–21).
2. Ertman’s reading of Hintze’s conceptual work is, in fact, very convincing, and his attempt to show how Hintze’s types empirically map onto the European space is probably more valid than Hintze’s own (rather sweeping) attempt. For example, Ertman (1997, 21) corrected a mistake that Hintze (1962 [1930], 124) made when he categorized Castile as a member of the set with stronger parliaments due to the lack of a Carolingian legacy. In that sense, Salter and Young (2018) chose a good source on which to rely. However, they make no attempt to argue that this is the case—and without reading Hintze in the original, they of course are unable to show (or even know) how valid Ertman’s interpretation is.

## REFERENCES

- Abramson, Scott, and Carles Boix. 2019. “Endogenous Parliaments: The Domestic and International Roots of Long-Term Economic Growth and Executive Constraints in Europe.” *International Organization* 73 (4): 793–837.
- Anderson, J. L. 1991. *Explaining Long-Term Economic Change*. London: Macmillan.
- Baechler, Jean. 1975. *The Origins of Capitalism*. Barry Cooper, trans. Oxford: Basil Blackwell.
- Benson, Bruce L. 2011 [1990]. *The Enterprise of Law: Justice Without the State*. Oakland, CA: The Independent Institute.
- Berman, Harold J. 1983. *Law and Revolution: The Formation of the Western Legal Tradition*. Cambridge, MA: Harvard University Press.
- Blaydes, Lisa, and Eric Chaney. 2013. “The Feudal Revolution and Europe’s Rise: Political Divergence of the Christian West and the Muslim World before 1500 CE.” *American Political Science Review* 107 (1): 16–34.
- Bloch, Marc. 1968 [1939]. *Feudal Society: Volume 1, The Growth of Ties of Dependence*. Chicago: University of Chicago Press.
- Capoccia, Giovanni, and Daniel Ziblatt. 2010. “The Historical Turn in Democratization Studies: A New Research Agenda for Europe and Beyond.” *Comparative Political Studies* 43 (8/9): 931–68.
- Coppedge, Michael. 2012. *Democratization and Research Methods*. New York: Cambridge University Press.
- Ertman, Thomas. 1997. *Birth of the Leviathan: Building States and Regimes in Medieval and Early Modern Europe*. Cambridge: Cambridge University Press.
- Feng, Li. 2008. *Bureaucracy and the State in Early China: Governing the Western Zhou*. Cambridge: Cambridge University Press.
- Gaddis, John Lewis. 2002. *The Landscape of History: How Historians Map the Past*. Oxford: Oxford University Press.
- Goldthorpe, Ian. 1991. *The Uses of History in Sociology: Reflections on Some Recent Tendencies*. Oxford: Nuffield College.
- Halsall, Guy. 2005. “The Sources and Their Interpretation.” In *The New Cambridge Medieval History: Volume 1, c. 500–c. 700*, ed. Paul Fourace, 56–92. Cambridge: Cambridge University Press.

- 
- Hexter, Jack H. 1979. "The Historical Method of Christopher Hill." In *On Historians*, ed. Jack H. Hexter, 227–51. London: HarperCollins.
- Hintze, Otto. 1962 [1930]. "Typologie der ständischen Verfassungen des Abenlandes." In *Staat und Verfassung*, Vol. I, ed. Otto Hintze, 120–39. Göttingen, Germany: Vandenhoeck & Ruprecht.
- Kreuzer, Marcus. 2010. "Historical Knowledge and Quantitative Analysis: The Case of PR Adoption in Early Democracies." *American Political Science Review* 104 (2): 369–92.
- Lange, Matthew. 2012. *Comparative–Historical Methods*. London: Sage Publications.
- Loewe, Michael, and Edward L. Shaughnessy. 1999. "Introduction." In *The Cambridge History of Ancient China*, ed. Michael Loewe and Edward L. Shaughnessy, 1–18. Cambridge: Cambridge University Press.
- Lord, R. H. 1930. "The Parliaments of the Middle Ages and the Early Modern Period." *Catholic Historical Review* 16: 125–44.
- Lustick, Ian S. 1996. "History, Historiography, and Political Science: Multiple Historical Records and the Problem of Selection Bias." *American Political Science Review* 90 (3): 605–18.
- Mahoney, James, and Kathleen Thelen (eds.). 2015. *Advances in Comparative–Historical Analysis*. Cambridge: Cambridge University Press.
- Møller, Jørgen, and Svend-Erik Skaaning. 2018. "The Ulysses Principle: A Criterial Framework for Reducing Bias When Enlisting the Work of Historians." *Sociological Methods & Research*, Online First.
- Myers, A. R. 1975. *Parliaments and Estates in Europe to 1789*. London: Thames & Hudson.
- Raico, Ralph. 1994. "The Theory of Economic Development and the 'European Miracle.'" In *The Collapse of Development Planning*, ed. Peter J. Boettke. New York: New York University Press.
- Salter, Alexander William. 2015. "Rights to the Realm: Reconsidering Western Political Development." *American Political Science Review* 109 (4): 725–34.
- Salter, Alexander William, and Andrew T. Young. 2018. "Medieval Representative Assemblies: Collective Action and Antecedents of Limited Government." *Constitutional Political Economy* 29 (2): 171–92.
- Stark, Rodney. 2011. *The Triumph of Christianity: How the Jesus Movement Became the World's Largest Religion*. New York: HarperCollins.
- Stasavage, David. 2010. "When Distance Mattered: Geographic Scale and the Development of European Representative Assemblies." *American Political Science Review* 104 (4): 625–43.
- Trachtenberg, Marc. 2015. "Transparency in Practice: Using Written Sources." In *Symposium: Transparency in Qualitative and Multi-Method Research*, ed. Tim Büthe and Alan Jacobs, 13–17. Newsletter of the American Political Science Association Organized Section for Qualitative and Multi-Method Research 13 (1).
- Vinogradoff, Paul. 1968 [1922]. "Feudalism." In *The Cambridge Medieval History, Volume 3: Germany and the Western Empire*, ed. H. M. Gwatkin, J. P. Whitney, J. R. Tanner, and C. W. Previté-Orton, chap. 18. New York: Cambridge University Press.
- Wickham, Chris. 2005. *Framing the Early Middle Ages: Europe and the Mediterranean 400–800*. Oxford: Oxford University Press.
- Wickham, Chris. 2009. *The Inheritance of Rome: A History of Europe from 400 to 1000*. London: Penguin Books.