AARHUS UNIVERSITY

Coversheet

This is the accepted manuscript (post-print version) of the article.
Contentwise, the post-print version is identical to the final published version, but there may be differences in typography and layout.

How to cite this publication
Please cite the final published version:

Publication metadata

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

Jørgen Møller

Professor, PhD

Department of Political Science

Aarhus University

Bartholins Allé 7

8000 Aarhus C

Denmark

E-mail: jm@ps.au.dk

Author bio: Jørgen Møller is professor at the department of political science, Aarhus University, Denmark. He may be reached at jm@ps.au.dk

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

Historians

Abstract

Political scientists increasingly enlist the work of historians but they often treat this work in a nonchalant or superficial way, which make their evidentiary record questionable. It follows that we need some kind of check on 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 where the use of historians' work as empirical evidence is unconvincing. The more general purpose of the article is to increase the consciousness about what is at stake when political scientists base empirical analysis on evidence gathered by historians.

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

Introduction

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 either by coding historical datasets based on historians' research, which can then be leveraged in quantitative analysis (e.g. Stasavage 2010; Blaydes and Chaney 2013; Abramson and Boix 2019), or by using historical works to create historical narratives, to do qualitative historical comparisons or to carry out within-case process-tracing (see 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, 148-149). 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 I raise is not whether the representations of the past are true or false but whether they would pass muster – at least in a minimum sense where the overall argument is not invalidated by

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

In this article, I first briefly describe what is at stake when political scientists enlist the work of historians. On this basis, I discuss what kind of check we need in order to avoid that the claims about historical evidence made by political scientists rest on feet of clay. I shall argue that what might seem the obvious answer – 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 details of historiography on a particular subject. These criteria include basing historical evidence claims mainly on work by historians rather than by social scientists, on historical work that does not seem blatantly outdated, on an awareness of differences within historiography on a particular subject, and on explicit page numbers in references. I use two recent articles to illustrate the purchase of these criteria.

What do we (really) know about the past?

The first thing political scientist who do historical analysis need to recognize is just 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 – say, the early middle ages

in Europe or ancient China – the only first-hand written sources we have are scattered documents such as law-codes, annals, and charters, which at best present partial information about the societies in question. Often, the bulk of our information comes from either archeology, which is difficult to interpret, or from later second-hand narrative sources, which clearly – often intentionally – misrepresent the past (Loewe and Shaughnessy 1999; Wickham 2005; Li Feng 2008).

To see the problems this create, let us begin with historian Chris Wickham's (2009, 12-18) discussion about the use of such second-hand narrative sources, in this case from the early medieval Europe. Wickham grants that many such 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, 116-7). Moreover, even if they are only partly fictitious, we have little to check their information against. But, Wickham assures us, this does not mean that these are not good sources for understanding the society they claim to write the history of. Their moralizing observations are thus anchored in "recognizable experience". At a minimum, these source therefore 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 as a source of what kind of potential actions and reasoning characterized these societies.

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

Two things follow from this. First, political scientists will not be able to do away with trained historians as gatekeepers to the past, at least if they are dealing with pre-modern 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 cannot be taken at face value but need to be interpreted by someone who understands the biases of different kinds of sources and who can place the interpretation within a particular historical context.

Second, political scientists have to factor in the uncertainty of *any* historical interpretation. Here, the third historical source singled out above, archeology, comes into play. The interpretation of archeological findings obviously also requires expert skills, and will often involve creative leaps in the form of qualified guesses. But another issue concerns the dynamic nature of our archeological knowledge. Wickham (2005, 1) justifies his magisterial attempt to revisit the early middle ages in Europe with the fact that over a couple of decades "what we can say about its archeology has multiplied tenfold – in some countries, a hundredfold". This means that historical interpretations which 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. Political scientist who enlist older work of historians, including work that may only be a few decades old, therefore need to probe whether these interpretations still pass muster among historians or whether new archeological knowledge has made them obsolete.

Finally, two additional issues complicate matters further. First, political scientists face the problem that there will often be competing historical interpretations, based on different perspectives on history. This means that they have to 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 these arguments. 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 tend to find what they are looking for. This further underscores why we need a check on the validity of the evidentiary record in historical analysis.

How to check the reading of the work of historians?

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

It might appear that the solution is simple. Going back to my initial analogy to the methodologist, we surely need to enlist at least one trained historian as reviewer in order to determine whether the reading of the work of historians passes muster. However, this solution has a number of unfortunate side effects. First, as any political scientist who has sent a historical analysis to a trained historian for comments knows, historians operate with different criteria for good research. Whereas political scientists normally try to generalize, compare, and to 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 instructed that a paper should be reviewed as a product of political science, historians would almost certainly be too critical of attempts to use historical material to generalize or make huge comparisons across space and/or time, and the concomitant lack of attention to detail. Second, as

already mentioned, there are different schools within history and representatives from one school (say, historians working in the Marxist tradition) would be overly critical when it comes to interpretations based on the work of other schools (say, historians working in the *Annales* tradition or political historians) (Hexter 1979, 61-145; see 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 importantly, editors and reviewers need to be self-conscious that the use of historians' work – whether for qualitative or quantitative purposes – cannot simply be taken at face value but needs to be questioned just like any other kind 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 limited costs.

Four criteria

What I am thinking about here is not a checklist which papers will either pass or fail. It should be seen as a quickly applied set of criteria that can help reviewers – who are doing a community service for free – and editors identify whether historical claims seem tenuous or not. The metaphor would be a warning signal, not a red light. If there is a warning, it would often make sense to raise this in reviews and, if the decision is R&R, to ask authors to substantiate and/or defend their reading of the work of historians better, if necessary in online appendixes. 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 some kind of 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 obviously not exhaustive, but they are easy to apply and they should flag the blatant cases where an author needs to buttress their historical claims by invoking more relevant and/or unbiased historical data. The first three criteria can be seen as simpler versions of the more demanding criteria presented by Lustick (1996, 615-6) and Møller and Skaaning (2018, 6-13), which cannot be used in their present form because they presuppose precisely what editors and reviewers normally lack, namely a detailed knowledge of the historiography on a particular subject.

I have already made it clear why we need the second and third criteria. The fourth – specific page numbers when referencing historical material – is particularly important where the claims that are invoked are not the central arguments or findings of the source but peripheral to it or maybe even made in an implicit way. In this case, the lack of page numbers make it extremely cumbersome for readers to try to check these claims – or to read up on them for one's own use – and it is also something of 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 of diminishing the risk of confirmation bias by steering clear of a theoretically polluted selection or even reading of the evidence that historians have initially produced (Møller and Skaaning 2018, 8-11). This criterion is not meant to imply that the work of historians is unbiased: it, too, is affected by implicit theories, and historians often move in herds or

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 pre-selected evidence to provide support for *one* central theoretical claim (Gaddis 2002, 56-58), meaning that it is likely to be less polluted by the theoretical vantage point. Moreover, this criterion can also be seen as a prudent way of checking whether historical claims first encountered by reading other social scientists can be corroborated based on work by historians, including more recent work.

I proceed to briefly illustrate how the four criteria can be used. To do this, I use examples from two recent public choice articles on the "Rise of Europe": Salter (2015) and Salter and Young (2018).[1] 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 lying behind the development of the European "nonpredatory governance" (726). The main aim of his article is to flesh out this theoretical argument. But Salter (2015, 728-9) makes clear that it is based on empirical underpinnings, that is, we find something resembling the "shareholder state" in medieval Western Europe.

What kind of historical evidence does he base this claim on? Salter (2015, 729) cites the following authorities: "Anderson (1991); Baechler (1975); Benson (1990); Berman (1983); Raico (1994); Stark (2011, Chaps. 14–6)". Only one of these is a general historian (Baechler 1975), another is a legal historian (Berman 1983). Both of these works are clearly dated, and throughout the article, there is no attempt to provide page references for the historical claims that are being made; despite the fact that the main point of these books is surely not that medieval monarchies resembled shareholder states. Nor 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 as granted that the medieval political order did in fact resemble his ideal type where claims to political power and revenue derived from property rights.

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" (173). To answer this, they revisit and elaborate the German historian Otto Hintze's (1962 [1930]) seminal argument "that two-chamber representative assemblies were more effective at resisting absolutist tendencies than three-chamber assemblies" (173). When presenting Hintze's claim, Salter and Young (2018, 178, fn. 14) use political scientist Thomas Ertman's (1997, 20-22) version of it,[2] and when detailing the functioning of two-chamber and three-chamber assemblies, respectively, they mainly refer to outdated historical evidence, including "Bloch 1968a [1939], pp. 145–162; Vinogradoff 1968 [1922]" (180) and "Lord 1930, p. 37; Myers 1975, p. 64" (fn. 20). Though they 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 "take as historical givens the two types of assembly structures" (174).

In both cases, one wonders why, during the review process, the authors were 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? Having read the papers, we have almost no clue about the validity of the evidentiary record. As the historian J. H. Hexter's (1979, 243-248) once critiqued one of his colleague's work, 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 have made 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 some sort of check on the validity of this data during review processes. However, the reason I try to push in what many may see as an open door is that this is not part of best practice today. As mentioned above, when participating in review processes either as author or reviewer, it has often 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. This article is mainly meant 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. I have presented four criteria that are easily applied, and I have used two recent publications to illustrate how they can be used to push authors to do a better job when using historical evidence.

References

Abramson, Scott, and Carles Boix. 2019."Endogenous Parliaments". Forthcoming in *International Organization*.

Baechler, Jean. 1975. *The Origins of Capitalism*. Barry Cooper, trans. Oxford: Basil Blackwell.

Berman, Harold J. 1983. *Law and Revolution: The Formation of the Western Legal Tradition*. Cambridge: Harvard University Press.

Blaydes, Lisa, and Chaney, Eric 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.

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.

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, UK: Nuffield College.

Halsall, Guy 2005. "The sources and Their Interpretation." In *The New Cambridge Medieval History: Volume I 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. London, UK: Collins.

Hintze, Otto. (1962[1930]). "Typologie der ständischen Verfassungen des Abenlandes." In Otto Hintze, *Staat und Verfassung vol. I*. Göttingen: 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.

Li Feng. 2008. *Bureaucracy and the State in Early China: Governing the Western Zhou*. Cambridge, UK: Cambridge University Press.

Loewe, Michael, and Edward Shaughnessy. 1999. "Introduction." In *The Cambridge History of Ancient China*, edited by Michael Loewe and Edward L. Shaughnessy. Cambridge, UK: Cambridge University Press.

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.

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.

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, eds. Tim Büthe and Alan Jacobs.

Newsletter of the American Political Science Association Organized Section for Qualitative

and Multi-Method Research 13 (1).

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.

Endnotes

---

[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 map on to the European space empirically is probably more valid than Hintze's own – rather sweeping – attempt. For example, Ertman (1997, 21) corrects 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) have chosen a good source to rely on. But they make no attempt to argue that this is the case, and without reading Hintze in the original they are of course unable to show (or even know) how valid Ertman's interpretation is.