Focusing on Fundamentals: A Reply to Koski and Horng
Sarah F. Anzia and Terry M. Moe
EDUCATIONAL EVALUATION AND POLICY ANALYSIS 2014 36: 120
DOI: 10.3102/0162373713518661

The online version of this article can be found at:
http://epa.sagepub.com/content/36/1/120

Published on behalf of

American Educational Research Association

and

http://www.sagepublications.com

Additional services and information for Educational Evaluation and Policy Analysis can be found at:

Email Alerts: http://eepa.aera.net/alerts
Subscriptions: http://eepa.aera.net/subscriptions
Reprints: http://www.aera.net/reprints
Permissions: http://www.aera.net/permissions

>> Version of Record - Feb 7, 2014
What is This?
Koski and Horng argue that our study of seniority-based transfer rules is narrow and that its findings only apply under limited circumstances. They say that their own study, by contrast, takes a broad frame—so broad that, as detailed in their 127-page technical report, they estimated 40 different models. They argue that the bigger picture that emerges from their more extensive analysis is that there is little evidence for our finding that these rules have a negative impact on the percentage of experienced teachers in disadvantaged schools relative to advantaged schools.

As a starting point, we should note that Koski and Horng’s published article does not mention this technical report, its 40 models, or its broad frame. The article simply presents an empirical analysis of seniority-based transfer rules and concludes that there is no impact. In our view, the article is properly assessed based on what it actually says and does. But given that Koski and Horng (in their response) have directed attention to the technical report, we will address their claim that the report provides the bigger picture and bolsters the article’s conclusion of no impact.

The technical report shares the same weaknesses as the article. Consider Table 1 of their response, which lays out the results of all 40 of the report’s models. These models are mainly distinguished by different measures of transfer rights (TLST and TLS1) and different samples on which the estimations were carried out. For reasons that we discussed in our own article, both the measures and the sampling criteria are problematic.

First, consider TLST, their “total” measure of seniority-based transfers: This measure is based on six items, some of which are not measures of the role of seniority in transfers. There is no theoretical justification for this, and it makes TLST a weak index of the key independent variable. Their TLS1 variable is a weak measure, too, because it is based on just one coding item (voluntary transfers). Koski and Horng say that this one-item measure reflects the argument of the original Moe study and the newer Anzia–Moe study as well, but we have never argued for such a measure. We have consistently argued that the best measure is one that combines seniority-based measures of voluntary and involuntary transfers—and in both the Moe and Anzia–Moe studies, those are the measures used.

Now consider the five samples that distinguish the various technical report models in Table 1. As we discussed in our article, it is important not to combine elementary, middle, and high schools in the same sample, because teachers typically do not transfer from one type of school to another. That is why we focused our own analysis on elementary schools—plus the fact that these schools are by far the most numerous type within districts. Koski and Horng agree
that teachers rarely transfer across different school types, and they exclude high schools from their main sample (Sample A) for this reason. Yet, Sample A still includes both elementary and middle schools; it also includes charter schools, which are not covered by district labor contracts. Sample B is the same, except it omits Los Angeles (and charter schools). Sample C focuses just on elementary school districts; but these districts typically contain middle schools as well as elementary schools. A focus on these districts also excludes a large number of elementary schools operating in unified districts; a huge amount of data is thrown away. Sample D includes all unified districts except Los Angeles—and therefore pools different types of schools. Sample E focuses just on high school districts; but some of these districts contain middle schools, and in any event, many of the state’s high schools are to be found in unified districts—meaning that, again, a large amount of relevant data is thrown away.

What this brief discussion illustrates is that Koski and Horng’s broad frame is essentially a proliferation of models that are not well justified on analytic grounds: Every one of their samples is problematic, and each model is based on a weak measure of seniority-based transfers. There is no bigger picture here. And the many insignificant coefficients shown in Table 1 do not justify or bolster the no-impact conclusion of their article.

If the published article is assessed for what it actually says and does, moreover, its no-impact conclusion is unwarranted. As we have argued, their data analysis was based on problematic measures and sampling criteria; but within the context of that analysis, they did not recognize that their findings actually supported, at a high level of statistical confidence, the hypothesis at issue: that seniority-based transfer rules have a negative impact on the proportion of experienced teachers in disadvantaged schools relative to advantaged schools. This was not apparent to readers, because the authors presented their findings with no standard errors or t-scores, only asterisks to indicate significance (or not) based on two-tailed tests. We think it is straightforward that, because the hypothesis being tested asserts a negative relationship, the appropriate test is one-tailed rather than two-tailed. But that aside, providing standard errors or t-scores is essential in any statistical analysis so that readers can judge for themselves how much confidence to have in the estimates. The standard errors from Koski and Horng’s analysis—which we estimated ourselves—show a high level of statistical support for the hypothesis of negative impact.

In their response, Koski and Horng essentially argue that significance tests should always be two-tailed, never one-tailed, regardless of the content of the hypothesis—which does not square with the statistical theory of hypothesis testing. They also argue that if a coefficient fails to meet a predetermined level of significance, then the automatic conclusion is one of no impact—even if the standard errors indicate that the null hypothesis can be rejected at a very high level of confidence. As they see it, judging coefficients is a strict matter of up or down, not a matter of determining what the estimation really has to say about statistical confidence. Most important, Koski and Horng reject the notion that they should present the standard errors or t-scores of their estimates. This is a position that, in our view, cannot be justified. Yet, it is reflected once again in Table 1 of their response, which provides readers with the estimated coefficients from 40 models—with no standard errors or t-scores.

In writing our own article, we sought to highlight these concerns about Koski and Horng’s article, but criticism was not our purpose. Our purpose was to be constructive in helping to build and encourage a new literature. We aimed to bring evidence to bear on the impact of seniority-based transfer rules—and in the process, to clarify the key theoretical and methodological issues involved in carrying out this kind of research, to explain what decisions on these issues seem to make good analytic sense, and to try to move the literature forward on a firm analytic footing. Koski and Horng criticize these efforts as leading to a narrow analysis. But the decisions we made were specifically designed to mitigate basic problems, and in so doing to allow for better models, measures, and tests.

Details aside, three of these analytic decisions stand out. First, we restricted our sample to elementary schools for reasons we have just discussed. Second, we declined to use the Koski–Horng transfer measures—for reasons we have
also discussed—and we built our own analysis around a composite index that, because it captures both voluntary and involuntary transfers and because it is only based on measures that explicitly code the role of seniority, is better suited to the job. Note that these two decisions on samples and measures are efforts to improve upon the Koski–Horng analysis, given their data set and coding scheme (within which we were operating)—but they are hardly permanent. In future work, scholars can surely carry out analyses on samples of middle schools and samples of high schools; and they can create other transfer measures based on new data sets and coding schemes. We are not wedded to the specifics of our own analysis. Our emphasis is on the criteria and logic underlying them—and on mitigating problems.

Our third decision was to make teacher experience the dependent variable, and not to include teacher credentials as well. Our reasons are straightforward. Seniority-based transfer rules are explicitly about experience. They are not about credentials. So if these rules do have an impact, it should show up most directly in how experienced teachers get distributed across schools—and exploring this connection should be the initial focus of research. We can only emphasize, moreover, that this is a very new literature, and already there is a dispute about whether these rules have any impact at all. The way to settle it is to put the focus on fundamentals. Experience is fundamental. Credentials are not.

Attention to fundamentals sheds a still brighter light on Koski and Horng’s own findings. They make much of the fact, for example, that 29 of their 40 technical report models yield insignificant impacts (by their calculations and decision rules), but they do not point out that 20 of the 29 insignificant coefficients arose in their credentials models (with 20 out of 20 producing insignificant coefficients)—whereas a majority of coefficients in their experience models were actually significant. The difference is quite striking and of theoretical interest, yet Koski and Horng do not take note of it in explaining what their statistical findings mean.

We agree with Koski and Horng that the larger goal is to understand how these seniority rules might affect teacher quality. But this is just another reason not to put the focus on credentials. Basic teaching credentials are only tenuously connected to quality, at best, and cannot be regarded as a well-supported proxy. Experience, however, is not only fundamental to the theory but also a better indicator of quality. The evidence shows that teachers in their first year or two on the job do not perform as well, on average, as their more experienced colleagues. Experience is properly the main focus of analysis.

Much remains to be done in exploring the connection between seniority rules and teacher quality, and more generally, in determining how contract rules affect the behavior of teachers, the organization of schools, and the quality of government service provision. Our own work is just the beginning of what we hope will prove a dynamic program of research going forward, involving many scholars. Our core theme, both here and in the article, is less about specific findings than about that process. It is about the importance of thinking clearly, systematically, and rigorously about the fundamentals that need to undergird this line of research. And it is about encouraging progress based on analytic focus and coherence.

An Addendum on Cohen-Vogel, Feng, and Osborne-Lampkin (2013)

As our article was being copy edited, we learned that a new article on seniority-based transfer rules by Cohen-Vogel et al. (2013) had just been published in this same journal. This new article cites our own and was written in knowledge of its contents, because we sent it to the authors long ago. We did not know about their finished paper, however, until it was actually published. Had we known, we would have included it in our assessment of the literature. The best we can do here, given the restrictions we are operating under, is to offer some very brief comments. We have posted a fuller set of comments online (http://gspp.berkeley.edu/directories/faculty/sarah-anzia).

The authors argue, based on Florida data, that the Koski–Horng conclusion of no impact is confirmed. It is possible, of course, that seniority-based transfer rules actually do not have any effect in Florida. But that would surprise
us—and we have certain concerns about their analysis. The first is that it suffers from the same weaknesses of measurement and sampling as Koski and Horng’s analysis: They use the same problematic measures of transfer rights (TLST and TLS1), and their sample apparently pools together elementary schools, middle schools, high schools, and charter schools in carrying out the estimation. Second, while they frame their analysis as a replication of Moe and Koski–Horng, their own models are actually quite different from, and far more complex than, the Moe and Koski–Horng models—and in ways that likely pose problems for their estimation. Third, while their models mainly generate insignificant coefficients (which does not surprise us, given the models’ complexity), the coefficients that do emerge as significant do not have a clear substantive or theoretical interpretation. And finally, the Cohen-Vogel et al. analysis does not explore whether seniority rules have an impact in large districts but not in small ones—a key finding of our own article, well grounded in theory, which is important to consider before embracing a blanket conclusion of no impact.

We do not know what the Florida evidence would show given a different research design. That said, we take Cohen-Vogel, Feng, and Osborne-Lampkin’s interest in the study of seniority-based transfers as a very positive development and an inspiration for new research. We hope that many more scholars will be drawn to this line of inquiry in future years—and that the field of education will be enlivened by a large and growing literature on the study of collective bargaining and its impacts on America’s schools.

Declaration of Conflicting Interests
The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding
The author(s) received no financial support for the research, authorship, and/or publication of this article.

Reference

Authors
SARAH F. ANZIA is an assistant professor of public policy at the Goldman School of Public Policy at the University of California, Berkeley. Her research focuses on American political institutions and the role of organized interest groups in elections and policy-making in the United States, particularly in state and local government.

TERRY M. MOE is the William Bennett Munro Professor of political science at Stanford University and a senior fellow at the Hoover Institution. His research interests include the politics and reform of American education, teachers unions, and collective bargaining, as well as the broader study of public sector unions, public bureaucracy, and American political institutions.

Manuscript received November 26, 2013
Accepted November 26, 2013