

RE: The Journal of Economic History - Decision on Manuscript ID JEH-2024-0189

Dear Prof. Moehling,

Thank you for giving us a chance to revise our paper for *The Journal of Economic History*.

Before delving into our point-by-point response to your letter, we wanted to summarize some of the major changes to the draft.

First, we have expanded our discussion of the historical context considerably. Our identification strategy, results, and robustness sections now include careful discussions of how the Great Depression and New Deal may have interacted with the prohibitions we study and the effects we find.

Second, we have updated all our analysis using the linked Census data to use individual-level data, as opposed to county-level aggregates.

In what follows, we respond point-by-point to questions and concerns raised by you and the referees (original comments from you and the referees are in *blue italics*, our reply in regular font).

Thank you again.

## Editor's Comments

*Both reviewers believe that the manuscript has the potential to advance scholarship on the effects of institutionalized discrimination and the rise of female labor participation in the twentieth century U.S. They do, however, raise some concerns and make suggestions for improving the paper.*

*Based on the reviewers' recommendations and my own reading of the paper, I have decided to invite you to revise and resubmit the paper. I believe your paper has promise, but I must make clear that resubmitting your manuscript does not guarantee eventual acceptance. I will send the resubmission to the same reviewers, and I will make my decision based on their evaluation of how you have addressed the concerns raised in this email and their full reports (which are provided below).*

### Authors' reply:

Thank you for the kind words about our paper, and again for giving us the opportunity to engage in yours and the reviewers' comments, which we feel significantly improve the paper.

*Here are the key issues I would like you to address in the revision:*

#### *1. Provide more historical context*

*Both referees make this request. The current manuscript provides a discussion of the history of marriage bars in teaching but not the conditions in which they operated.*

*The glaring omission for me – and noted by both referees – is a discussion of how the Great Depression and the New Deal may have affected the outcomes you study. As Referee 2 notes, there is an extensive literature on how the Great Depression and New Deal affected labor supply, household formation, and fertility decisions. At the very least, you need to discuss whether any of your results, which are based on variation across states, could be explained by the geographic variation in intensity of the economic downturn or the level of New Deal expenditures. I believe it would also be helpful to think through whether being in the midst of the Great Depression dampened or amplified the effects of prohibiting marriage bars. Would we have expected bigger or smaller shifts from single to married women teachers if the country had been in more prosperous times?*

*I am also interested in whether the Great Depression played a role in the politics around marriage bars and their prohibition. For instance, did the economic crisis lead to the intensification of efforts to eliminate marriage bars because of the role women's wages played in the family economy?*

### Authors' reply:

Thank you for this advice. We agree that given the timing of the prohibitions, we certainly should provide more contextual details as to how the Great Depression and the New Deal interact with our findings.

In the updated text, we investigate how the Great Depression and New Deal may have differentially affected our treated and control states, assess how our results may be biased as a result, and include robustness checks that allow us to assess the extent of the bias empirically.

The main new text, which describes the impacts of the Great Depression and New Deal and the approaches we take, can be found in Section 4.3. For your convenience, we replicate the new text below:

Despite the many economic and policy changes that took place between 1910 and 1950 that could have affected women’s marriage and employment outcomes, many of these changes did not differentially impact our treated and control states. For instance, changes in compulsory schooling laws and child labor laws, which could have affected demand for teachers, did not occur in most of our treated and control states between 1910 and 1950.<sup>1</sup> One of the key changes in teacher training in the early 1900s was the shift from 2-year “normal schools” to 4-year college programs, but these changes largely took place before 1930.

However, we do find two notable differences in how our treated and control states were affected by the Great Depression and New Deal. Here, we reference prior work on the effects of the Great Depression to assess the potential biases these differences may cause in our design, and develop strategies that allow us to address these concerns.

The first difference is that the average treated county experienced a more severe economic downturn and a more muted recovery than the average control county during the Great Depression.<sup>2</sup> Two control states in particular experienced comparatively mild effects of the Great Depression: in Supplemental Table B2, Column (2) shows that while KY/NC experienced similar decreases in retail sales per capita between 1929 and 1933 as control states TN and WV, SC and VA experienced much smaller declines. At the same time, Column (3) shows that with the exception of WV, control states experienced slightly larger recoveries in retail sales per capita between 1933 and 1939 than KY/NC.

How might these differential trends bias our results? Prior literature finds that the Great Depression delayed women’s decisions to marry ([6]) and contributed to more widespread implementation of marriage bars, to preserve job openings

<sup>1</sup>The only change that occurred was that VA, one of our control states, required one more year of compulsory schooling in the 1920s, which would have led to increased demand for teachers and bias our estimates downwards.

<sup>2</sup>There is data on New Deal spending by county from 1933 to 1939, but given the time period and decennial nature of the census, the data cannot be used to construct time-varying measures of New Deal spending in 1930 and 1940. We proxy for both recession severity and economic recovery using changes in retail sales per capita instead.

for men ([4]). At the same time, [1] find that more severe economic downturns pushed more white women into the labor market, with suggestive evidence that married women were affected as well—although this “added worker” effect was tempered by New Deal programs ([2]).<sup>3</sup>

The direction of potential bias is therefore ambiguous. In the absence of the marriage bar prohibitions, worse economic conditions in treated states relative to control states would have increased the use of marriage bars, delayed marriage among women, and decreased resources for schools, all of which would have reduced the relative likelihood that women got married or taught in treated states compared to control ones. At the same time, the “added worker” effect may have resulted larger inflows of married women into the labor force in treated states compared to control states. Our estimates could thus be biased downward by the reduced likelihood of treated women getting married or teaching in treated states, or biased upward by the “added worker” effect in treated states.

We use several strategies to address these potential biases. Our primary strategy involves comparing treated states to alternate control groups that were more similarly affected by the Great Depression. Supplemental Appendix C outlines these specifications in detail, including a border counties design, where we only compare counties along state borders, and a matched county design, where we match counties based on 1920 and 1930 characteristics. We also include controls for unemployment rates and industry shares in several specifications.

The second notable difference between treated and control states was their exposure to New Deal policies. A large number of control counties in TN, but only a small number of treated counties, were directly impacted by the Tennessee Valley Authority (TVA), a major New Deal program that aimed to provide the first federal electrification program and stimulate the economy in the hard-hit Tennessee Valley. Prior work finds that the TVA had large impacts on the local economy, particularly in shifting work away from agriculture to higher-paying manufacturing jobs ([7]). Combined with the aforementioned studies, these results suggest that including TVA-affected counties in our analysis could bias our estimates upwards, as increasing manufacturing jobs in TVA-affected counties would have decreased the need for married women to work in more control counties than treated ones. To address this potential bias, we conduct robustness checks that exclude the counties affected by the TVA.

We describe the results of the additional specifications we include to account for the effects of the Great Depression and New Deal in Section 5.4:

Finally, we discuss the robustness of our results to the impacts of the Great Depression and New Deal. Our alternate control group designs allow us to address the potential bias caused by treated states being more severely impacted overall by the Great Depression than control states. In our border counties design, we

---

<sup>3</sup>Similarly, while prior work finds that birthrates decreased during the Great Depression ([8]), the New Deal countered these effects by decreasing infant mortality and increasing birthrates ([3]).

compare counties that share similar industry composition and other unobservable characteristics before 1933, and hence are more likely to experience similar effects of the Great Depression and more similar New Deal support. In our matched counties design, we match counties on retail sales per capita in 1929 and growth over time. We also examine how our results are affected by including controls for county unemployment rates and industry composition. Finally, in Column (5) of Supplemental Table B6, we present our results excluding counties in the TVA. As discussed above, our results are robust to all these specifications, suggesting that any biases caused by the Great Depression are minimal.

*Referee 1 wants to know how teacher training varied over your study period and if and how it varied across states. Training to be a teacher was an investment decision; marriage bars imposed constraints on the returns to that investment. How then did marriage bars impact who trained to be teachers and how long they remained in the profession? Did this impact vary across time and space in a way that may affect your analysis?*

#### **Authors' reply:**

Thank you for raising this issue. We do not find evidence of teacher training varying significantly across states from 1930 onward. The primary change that occurred in teacher training, which we now mention in the identification strategy in Section 4.3, took place before 1930 in both our treated and control states:

One of the key changes in teacher training in the early 1900s was the shift from 2-year “normal schools” to 4-year college programs, but these changes largely took place before 1930.

From this, we do not expect changes in teacher training to affect our analysis.

Interestingly however, the overall shift towards longer, more formal teacher training does help explain one of our descriptive findings. In describing the marriage bar prohibitions in Section 2, we note that the share of married women employed in teaching slowly rose across all treated and control states even prior to 1930. We have added a footnote that gives a new possible interpretation of this shift, which we reproduce here:

One potential contributor to the slight increase in married women teachers across states by 1930 was the shift to longer, more formal accreditation. With most teacher training programs changing from local 2-year normal schools to state-approved 4-year teaching colleges by 1930, schools in need of qualified teachers may have chosen to retain/rehire trained and married women teachers as a result ([5]).

#### *2. Analyze individual-level data*

*As Referee 2 notes, using the individual-level data would allow you “to control for rich personal and household characteristics available in the census data, ad-*

*adjusting omitted variable biases more effectively.” They would also allow you to “capture heterogeneity in the effects of the prohibition on various subgroups, such as those defined by age, county population size, and economic structure.” Particularly for the linked census data, I do not see the value of limiting the analysis to county-level aggregates. Those aggregates are difficult to interpret in any case because they only include the linked individuals.*

### Authors’ reply:

Thank you for this suggestion. We agree that there is more precision and information offered in the individual-level data that could be beneficial to include in the linked analysis in particular.

As a result, we have updated the majority of our analysis to use individual-level data. The only place where we have opted to continue using county-level aggregates is in the cross-sectional analyses, where we examine the effect of the prohibitions on the county-level shares of teachers who are men/women/married women, the total number of teachers in a county, and the county-level shares of married women who are teachers. We continue using county-level outcomes in the cross-sectional analysis because with this analysis, we are interested in assessing the effects of the prohibitions on a region’s teacher workforce, meaning our outcomes of interest are most naturally defined at a regional level.

We do, however, replicate our main cross-sectional results at the individual level by estimating the effects of the prohibitions on the likelihood that an individual teacher is a married woman. The results are shown in Supplemental Table B3. The estimate in Column (1), while less natural to interpret, is very similar to our preferred county-level estimates in both size and magnitude.<sup>4</sup> This individual-level analysis also allows us to explore heterogeneity in our results by race, urbanicity, and age, and to perform robustness checks that include additional controls suggested by the referees, including county population size, local industry shares (to proxy for education demand), and local unemployment rates (to proxy for Great Depression impacts). Although it would be ideal to also control for education and household income, these variables are unfortunately not available until the 1940 Census.

### *3. Describe how you deal with geographic mobility in the linked samples*

*Referee 2 raises this question and is one of the reasons they argue for using the individual-level data. Do you define an individual’s county as where they lived in year  $t-1$  or  $t$ ?*

### Authors’ reply:

Thank you for this clarifying question. We define an individual’s county based on where they live in year  $t - 10$ . We clarify this in Section 4 when we describe our outcome variables,

---

<sup>4</sup>Note that this is exactly equivalent to our county-level specification when individual observations are inversely weighted by the number of teachers in a county, which we demonstrate in Column (2).

and replicate the new text below for your convenience:

The second set of outcomes allows us to understand the mechanisms by which the prohibitions affected the composition of teachers. These outcomes are constructed at the *individual level* using the linked decennial Censuses, allowing us to investigate women's transitions in and out of marriage, teaching, and the workforce. For each woman  $i$  observed in year  $t - 10 \in \{1910, 1920, 1930\}$ , our main outcomes of interest include: an indicator for her marital status in year  $t$ , and an indicator for the interaction between her marital status in year  $t$  and her occupational status in year  $t$  (either teaching, not teaching but in the labor force, or not in the labor force). Treatment status is defined based on the county in which the person lives in year  $t - 10$ .

Additional linked outcomes include a woman's occupational score in year  $t$  and indicators for whether, in year  $t$ , she has any children, is in the labor force, or is living in a different state.

Your question on mobility further prompted us to look into whether the prohibitions had any effect on women's mobility. Given we find that the prohibitions increased the employment of married women in teaching, we expected to see the prohibitions cause unmarried women who stayed unmarried to move more (say, out of state for other jobs), or to cause unmarried women who got married to move less. We find evidence of the former effect, which we now include in Section 5. We replicate the text and corresponding Supplemental Appendix Table B5 below for your convenience:

Similarly, although the prohibitions increased movement out of state on average across all incumbent unmarried women teachers (Column 4), we find that this difference is almost entirely driven by those who remained unmarried (Column 5). These results are consistent with incumbent teachers (in particular unmarried women, who experienced no positive retention effects from the prohibitions) pursuing other jobs or wanting to teach in states that were unaffected by the prohibitions.

**Table B5.** Estimated effects of marriage bar prohibitions on fertility, occupational scores, and mobility

Dependent Variable:	Has child in $t$		Occupational Score in $t$		Moves state in $t$	
Model:	(1)	(2)	(3)	(4)	(5)	
Treated $\times$ 1940	-0.0208 (0.0135)	0.7275* (0.4080)	-0.0949 (0.2835)	0.0401*** (0.0093)	0.0593*** (0.0118)	
Treated $\times$ 1940 $\times$ Married in $t$			1.520*** (0.4299)		-0.0430*** (0.0131)	
Dep. Var. 1930 Treated Mean	0.4819	10.39	10.39	0.1421	0.1421	
County fixed effects	Yes	Yes	Yes	Yes	Yes	
Observations	59,542	59,542	59,542	59,542	59,542	
Adjusted R <sup>2</sup>	0.05953	0.09277	0.46260	0.01673	0.04023	

*Notes:* All columns show results from estimating Equation (2) for women in linked Sample 1, defined in the notes of Table 3. The outcome in Column (1) is an indicator for whether the woman has child in year  $t$ . The outcome in Column (2) is the woman's occupational score in year  $t$ . The outcome in Column (4) is an indicator for whether the woman is in the LF in year  $t$ . The outcome in Column (6) is an indicator for whether the woman lives in a different state in year  $t$  compared to year  $t - 10$ . Columns (3), (5), and (7) add to the previous column an interaction term for whether the woman is married in year  $t$ . Standard errors are clustered at the county level. All specifications include county fixed effects.

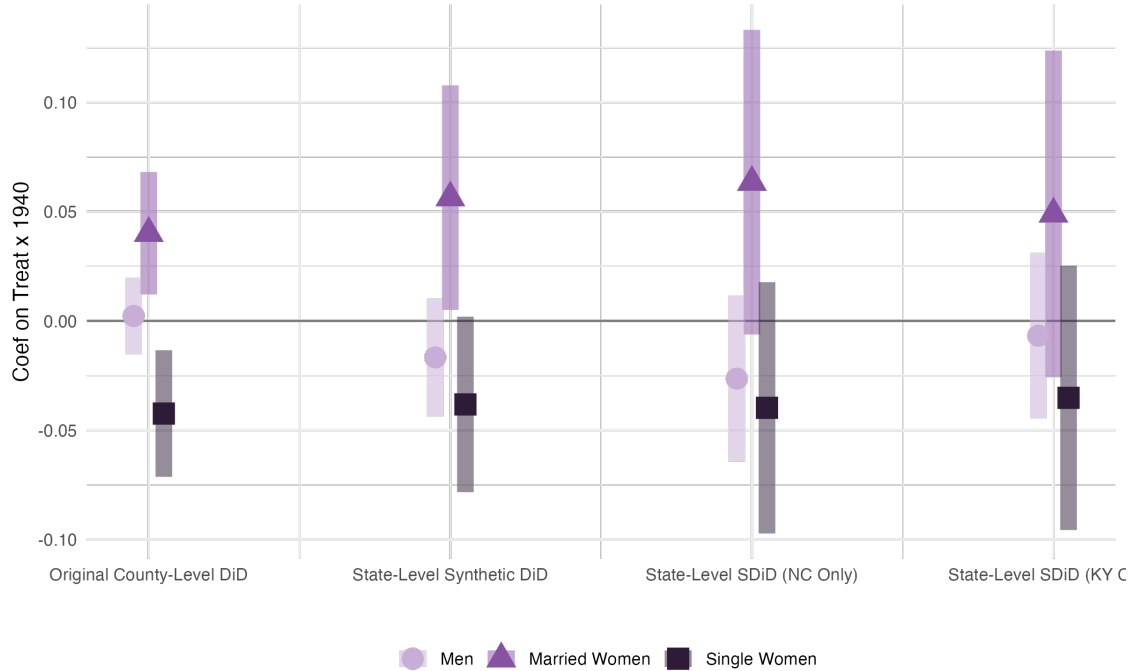


4. Discuss whether the effects of the prohibition of marriage bars differed between NC and KY

*Based on your description on page 11 of the laws, my expectation is that the effects would differ across these states. NC passed a law in 1933 that was according to your description, “broad in its application.” KY passed a law in 1938 that prohibited “rules preventing marriage of any school teacher who has had five years or more teaching experience.” I would expect a much smaller effect for KY given its narrower scope and the fact that it had only been in effect for 2 years in 1940.*

**Authors’ reply:**

Thank you for this suggestion. To address the question of whether the different applications of the prohibition laws across states may have led to different effects, Supplemental Appendix Figure D2 shows the main difference-in-difference estimates separately by state. We reproduce the figure below for your convenience.



**Figure D2.** Estimates of the effect of the prohibition of marriage bars in teaching on the gender composition of teachers. The first column uses the standard difference-in-differences setup from our main specification at the county level, with standard errors clustered at the county level. The other columns use state-level synthetic difference-in-differences, with standard errors computed using a ‘placebo’ method. The second column includes both KY and NC as treated units, while the third and fourth only include NC and KY respectively. 95% confidence intervals are shown.

The results are qualitatively the same by state, although the estimates are less precise given the sample becomes quite small. In line with what one might expect, the effect is slightly weaker in KY, where the laws were not in effect as long as they were in NC by 1940.

We also provide a footnote in the empirical strategy section that mentions this difference between states:

In our state-level analysis in Supplemental Appendix D, we also show results for NC and KY separately. Our findings are not being disproportionately driven by one state; in fact, point estimates are remarkably similar for both states.

*5. Provide more discussion of your findings in terms of retention versus hiring practices*

*I share Referee 1's confusion about whether your results indicate changes in retention or hiring practices: "I had a difficult time understanding why many of the results were attributed to hiring of married women as opposed to retention of married women. On page 22, you find that retention is important. Similarly, are you sure that unmarried teachers are being "let go" (p.24) as opposed to not hired?"*

**Authors' reply:**

Thank you for this suggestion. We have clarified the language throughout our results section to better distinguish between the extent to which our results indicate changes in retention and hiring practices. To summarize: our linked analysis allows us to distinguish between retention and hiring channels in two steps. First, we isolate the group of women who could *only* have contributed to the increase in married women teachers through retention: that is, the women who were previously unmarried teachers. For these women, we estimate the effects of the prohibitions on the likelihood they became married teachers. We then do the same for the group of women who could *only* have contributed to the increase in married women teachers through hiring: that is, the women who were already married but not in the labor force. Comparing results between the two groups allow us to assess the role of retention and hiring practices in explaining our main results.

We quote the results of this analysis below:

Our results suggest that the increase in married women in teaching was driven entirely by changes in extensive margin labor supply, both by increasing *retention* of incumbent teachers who would otherwise have exited the labor force, and by increasing *hiring* of women who would have otherwise remained out of the labor force. We find no evidence of negative effects on the LFP of women who were already in the labor force or who would have entered the labor force even absent the prohibitions.<sup>5</sup>

Furthermore, our estimates suggest that new hiring contributed more to the increase in married women teachers than the retention of incumbent teachers did. To see this, we scale the estimated effects in Table 3 by the total number of treated white women in each group in 1930. Although the effect size in percentage points

<sup>5</sup>See Supplemental Table B4 for the estimated null effects on married and unmarried non-teachers in the labor force in  $t - 10$ .

is largest for unmarried women teachers (2.2 p.p.), there were fewer than 8,000 such women in our treated linked sample in 1930, suggesting the prohibitions led to the retention of an additional 168 married women teachers. In contrast, there were more than 434,000 married women outside of the labor force in our treated linked sample in 1930, meaning our effect size of 0.06 p.p. translates to an increase of 258 married women teachers due to new hiring. We conclude that roughly 60% of the overall increase in married women teachers was due to the hiring of new married women.<sup>6</sup>

*6. Drop, or better defend, the back of the envelope calculations presented on pages 20-21 and Appendix C*

*I found these calculations unconvincing. In my opinion, they require some heroic assumptions: (1) that KY and NC are representative of the experience in all states across the country; (2) that the teaching labor market is representative of the labor market for all white collar female workers; and (3) that your estimate of the “treatment” effect of prohibiting marriage bars was not dampened or amplified by being observed during the Great Depression.*

**Authors’ reply:**

Thank you for this suggestion. We agree that the back of the envelope calculations do not significantly strengthen the paper without making strong assumptions. We have now removed them from the paper.

---

<sup>6</sup>While women who were unmarried and outside the labor force in  $t - 10$  also contributed to the increase, recall that we cannot disentangle what shares of these women became married women teachers through hiring or retention channels. However, we can apply a similar argument: of the unmarried women out of the labor force in  $t - 10$ , a much larger number will get married by  $t$  than become teachers by  $t$ ; hence, if the percentage point effect sizes are similar in magnitude, then the effect of the prohibitions on unmarried women outside of labor force will likely be driven by hiring effects.

## References

- [1] Andriana Bellou and Emanuela Cardia. The great depression and the rise of female employment: A new hypothesis. *Explorations in Economic History*, 80:101383, 2021.
- [2] T Aldrich Finegan and Robert A Margo. Work relief and the labor force participation of married women in 1940. *The Journal of Economic History*, 54(1):64–84, 1994.
- [3] Price V Fishback, Michael R Haines, and Shawn Kantor. Births, deaths, and new deal relief during the great depression. *The review of economics and statistics*, 89(1):1–14, 2007.
- [4] Claudia Goldin. Marriage bars: Discrimination against married women workers, 1920's to 1950's. Working Paper 2747, "National Bureau of Economic Research", October 1988.
- [5] Charles Athiel Harper. A century of public teacher education: The story of the state teachers colleges as they evolved from the normal schools. 1939.
- [6] Matthew J Hill. Love in the time of the depression: The effect of economic conditions on marriage in the great depression. *The Journal of Economic History*, 75(1):163–189, 2015.
- [7] Patrick Kline and Enrico Moretti. Local economic development, agglomeration economies, and the big push: 100 years of evidence from the tennessee valley authority. *The Quarterly journal of economics*, 129(1):275–331, 2014.
- [8] Jessamyn Schaller, Price Fishback, and Kelli Marquardt. Local economic conditions and fertility from the great depression through the great recession. *AEA Papers and Proceedings*, 110:236–240, 2020.