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

Dear Referee #2,

Thank you for your helpful report. We have done our best to reply to each of your concerns, which we feel has dramatically improved the paper.

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 in your report (original comments are in *blue italics*, our replies in regular font).

Thank you again.

Referee #2's Comments

I believe that, using novel data, the manuscript carefully studies important issues in economic history, namely the rise of female labor supply, gender discrimination in the labor market, and the role of legal institutions on human behavior. The findings are intriguing, and the results are presented with thoughtful discussions. My remaining questions and comments for revision are as follows.

Authors' reply:

Thank you so much.

First of all, I am wondering why the county-level aggregate analysis is primarily reported, rather than analyzing rich individual-level data. The prohibition of marriage bars shifts labor demand, and the outcomes can be captured along the supply curve. Ultimately, individuals make decisions about whether to work and what to do. In principle, county-level aggregate analysis using population weights could yield identical results to an uncontrolled individual-level analysis. However, individual-level analysis allows us to control for rich personal and household characteristics available in the census data, addressing omitted variable biases more effectively.

An individual-level analysis is especially desirable when using linked census data. Analyzing individual panel data would reveal detailed information about labor market transition affected by the prohibition between pre-treatment and post-treatment years, after controlling for personal and household characteristics. This approach would also help us capture heterogeneity in the effects of the prohibition on various subgroups, such as those defined by age, county population size, and economic structure.

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 women. 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.[16] 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.

Another question concerns women's mobility in the linked census data. Marriage may have led to geographical relocation across counties or states. Married women might also have moved across counties and states between two adjacent census years. I wonder how the current county-level analysis addresses this issue. I believe this issue could be better addressed by analyzing individual-level data rather than county aggregates.

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, 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

---

[16]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).

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 × 1940	-0.0208	0.7275*	-0.0949	0.0401***	0.0593***
	(0.0135)	(0.4080)	(0.2835)	(0.0093)	(0.0118)
Treated × 1940 × Married in t			1.520***		-0.0430***
			(0.4299)		(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^2	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.

Urban-rural gaps also need to be considered. In the 1930s and 1940s, especially in the South, a considerable share of the school-age population did not attend or graduate from high school. The gap in both the demand and supply of public education may have been significant between rural and urban areas, leading to differences in the demand and supply of school teachers. This is a county-level time-varying factor, not captured by fixed effects. While detailed data on the provision of local public schools may not be available, controlling for urban status should not be challenging. Local industrial structures are frequently used as a proxy for local demand for education and could be easily added to the regressions.

Authors' reply:

Thank you for this recommendation. Your comment prompted us not only to investigate whether urban-rural gaps may be a potential threat to identification, but also whether our effects may have differed by urban/rural status. We have thus added heterogeneity analysis of our main results by individual urban/rural status in Column (7) of Supplemental Table B3, replicated below for your convenience. We were surprised to find that our results are remarkably consistent across teachers in urban and rural areas.

Following your suggestion below on controlling for factors related to the Great Depression, we now also include robustness checks in which we control for county-level shares of manufacturing and agriculture (in addition to other county-level controls), which we discuss in the results section as follows:

> Our main specifications include year and county fixed effects without additional controls. However, particularly due to the Great Depression, there were also county-level time-varying factors that are not accounted for by county fixed effects. To address this, Column (4) of Supplemental Table B3 shows the effect of the prohibitions on the share of teachers that are married women including controls for county-level unemployment rates (to account for the effects of the Great Depression), shares of workers in manufacturing and in agriculture (to account for possible changes in demand for education), and log county population and restricts the sample to only 1930 and 1940 (when data on unemployment is available). We find that our main results do not change with the inclusion of these controls.

Table B3. Heterogeneity in estimated effects of prohibitions on becoming a married woman teacher

Dependent Variable:					Pr(Married Teacher in t)				
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Variables									
Treated ×1940 (γ_{1940}^{DD})	0.0350***	0.0347***	0.0240***	0.0313***	0.0375***	0.0388***	0.0346***	0.0405***	0.0397***
	(0.0073)	(0.0066)	(0.0079)	(0.0064)	(0.0075)	(0.0072)	(0.0072)	(0.0077)	(0.0068)
Treated ×1940 × *Nonwhite*						-0.0332**		-0.0426**	
						(0.0129)		(0.0171)	
Treated ×1940 × *Urban*							-0.0023	-0.0107	
							(0.0112)	(0.0121)	
Treated ×1940 × *Nonwhite* × *Urban*								0.0306	
								(0.0232)	
Treated ×1940 × (*Age* − 30)									-0.0025***
									(0.0005)
Inverse Weighted by County	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	437,111	437,111	322,511	221,846	343,350	437,111	437,111	437,111	437,111
Adjusted R²	0.09100	0.10761	0.09672	0.03215	0.10726	0.12169	0.10846	0.12249	0.12136

Notes: Estimation of Equation (2) for the sample of all teachers in our balanced sample of counties in treated (KY, NC) and neighboring Southern control states (VA, SC, TN, WV). The outcome variable for all regressions is an indicator for whether an individual is a married woman. Column (1) estimates Equation (2) for all teachers with no weights or controls. Column (2) replicates the county-level regression in Column (1) of Table 4 by inversely weighting each observation by the number of teachers in a county and year. All remaining columns use county-year inverse population weights for comparability with our main county-level results. Column (3) restricts the sample to the set of counties with at least ten Black teachers in 1930 and 1940 (see Section 3.3 for more details). Column (4) includes county-level controls for the share of workers in manufacturing, the share of workers in agriculture, unemployment rate, and log population. Note that the sample size is smaller because full-count Census data on employment status was only available in 1930 and 1940. Column (5) excludes counties in the Tennessee Valley Authority (see Section 4.3 for further discussion). Columns (6) and (7) include indicator for whether a teacher is non-white (*Nonwhite*) and whether a teacher lives in an urban area (*Urban*) respectively as additional interaction terms. Column (8) includes *Nonwhite*, *Urban*, and the interaction between the two as additional interaction terms. Column (9) includes an individual's age relative to 30 (the mean age for teachers in treated counties in 1930) as an additional interaction term. All regressions include county and year fixed effects and use 1910-1950 full-count cross-sectional decennial Census data unless otherwise stated fullcount-census. Standard errors are clustered at the county level.

Additionally, it should be noted that during the period between 1930 and 1940, or the treatment period defined in the analysis of the manuscript, the Great Depression swept across the United States. The nationwide common impact of the Great Depression and subsequent public interventions can be separated using a difference-in-differences approach. However, state-specific and county-specific impacts may remain and complicate the analysis. Considering the local variation in public intervention in the labor market, identifying the causal effects of the prohibition on labor supply using a state-level difference-in-differences approach with county-level aggregate data could be challenging. I do not have any immediate suggestions to address this issue. A good starting point might be a careful discussion of the possible relationship between the regression results and the Great Depression or New Deal. If easily accessible county-level data capturing local variation in New Deal policies exist, conducting robustness checks using such data would be desirable. To address the heterogeneous impacts of the Great Depression on local labor markets, local unemployment rates could be added as control variables.

Authors' reply:

Thank you for raising this issue and for your thoughtful suggestions. 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 and find ways to address these potential relationships.

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.[17] 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 dif-

---

[17]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.

ferences 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.[18] 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 ([5]) and contributed to more widespread implementation of marriage bars, to preserve job openings 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]).[19]

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 expo-

---

[18]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.

[19]Similarly, while prior work finds that birthrates decreased during the Great Depression ([7]), the New Deal countered these effects by decreasing infant mortality and increasing birthrates ([3]).

sure 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 ([6]). 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 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.

Finally, in addition to the research papers by Claudia Goldin and Martha Bailey cited in the manuscript, there is a wealth of economic history literature studying the rise of female labor force participation during the same period. For example, Bellou and Cardia (2021) examines the rise of female employment during the Great Depression. Costa (2000) provides a good review of the long-run change in U.S. female labor supply compared to international counterparts. Finegan and Margo (1994) argues that the labor supply of married women in this period was closely associated with the spouse's employment. Hill (2015)'s findings on the relationship between economic conditions and marriage during the Great Depression also warrant attention.

Authors' reply:

Thank you for pointing this out and providing these references. We have expanded our literature review as a result, particularly when discussing the Great Depression and New Deal as shown above, to include the literature on female labor force participation during the

Great Depression; indeed, the references you mentioned ended up being some of the most relevant for our study.

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] 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.

[6] 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.

[7] 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.