The Effects of Sexism on American Women: The Role of Norms vs. Discrimination*

Kerwin Kofi Charles† Jonathan Guryan‡ Jessica Pan§

June 4, 2019

Abstract

We examine the extent and channels through which reported sexism affects American women. Using a sample of internal migrants and an IV approach, we show that higher prevailing sexism where a woman currently lives (residential sexism) as well as where she was born (background sexism) adversely affect her labor market outcomes relative to men, increases marriage, and reduces childbearing age. The pattern of whose sexism (men vs. women, and specific percentiles) matters for which set of outcomes suggest that residential sexism primarily affects labor market outcomes through prejudice-based discrimination, and non-labor market outcomes through the influence of current norms.

*We are grateful to Rebecca Diamond, Raquel Fernandez, Nicole Fortin, and Claudia Goldin, Daniel Rees, and seminar participants at Yale University, University of Chicago, Northwestern University, University of Oxford, University College London, NBER, ASSA Meetings, University of Chicago, UC-Berkeley, Iowa State University, American University, Princeton University, Purdue University, University of Michigan, University of Naples, Federal Reserve Bank of Cleveland for helpful comments and discussions.

†University of Chicago and NBER. kerwin.charles@gmail.com

‡Northwestern University and NBER. j-guryan@northwestern.edu

§National University Singapore and IZA. jesspan@nus.edu.sg

1 Introduction

The average American woman's socioeconomic outcomes have changed dramatically over the past fifty years. In the labor market, her wages and probability of employment have risen substantially compared to the average man's (Blau and Kahn, 2017); at every age, she is less likely to have ever been married (Isen and Stevenson, 2008); and she has fewer children over her lifetime and is older when she bears her first (Bailey et al., 2014). A large literature analyzes the evolution of these average changes, but virtually no work has studied heterogeneity in women's outcomes across markets in the U.S.[1] We document large cross-state differences in women's outcomes, even within the same geographic region of the country. These gaps have persisted for decades and are not accounted for by standard individual controls.[2]

This paper assesses how, and by what mechanisms, cross-market differences in prevailing sexism affect these persistent residual cross-market differences in women's outcomes. We focus exclusively on whites to avoid conflating issues concerning gender with the potentially different set of considerations having to do with race, and control throughout for individual differences in years of schooling and age. In principle, sexism might take many forms but we focus on negative or stereotypical beliefs concerning the ability or appropriateness of women engaging in market work rather than home production. More specifically, our analysis proxies for market-level sexism using a composite index of gender-related attitudes constructed using the General Social Survey (GSS). These questions touch on the extent to which residents in a state believe (i) that women's capacities are inferior to men's; (ii) that the family unit is hurt when women focus on activities outside the home; or (iii) that men and women should occupy specific, distinct roles in society.

Suggestive evidence that prevailing sexism at the state-level matters for women's (relative) outcomes is presented in Figure 1, which plots the conditional female-male labor force participation gap (Panel A) and female age at first marriage (Panel B) against the degree of agreement at the state-level to two questions in the General Social Survey (GSS): (1) married woman earning money

[1] Some factors whose roles in driving average national changes have been studied include the contraceptive pill (Goldin and Katz, 2002; Bailey et al., 2014); technological changes in home and market production (Greenwood et al., 2005; Weinburg, 2000; Black and Spitz-Oener, 2010); shifts in occupational sorting (Hsieh et al., 2013); and women's educational gains (Charles and Luoh, 2003; Goldin et al., 2006; Blau and Kahn, 2006). Two of the small number of papers to have studied cross-market differences are Beaudry and Lewis (2014) who study cross-market wage differences; and Black et al. (2014), who study differences in married women's labor supply across cities.

[2] In other words, differences in labor market composition across states on the basis of age and education do not explain the observed cross-state differences.

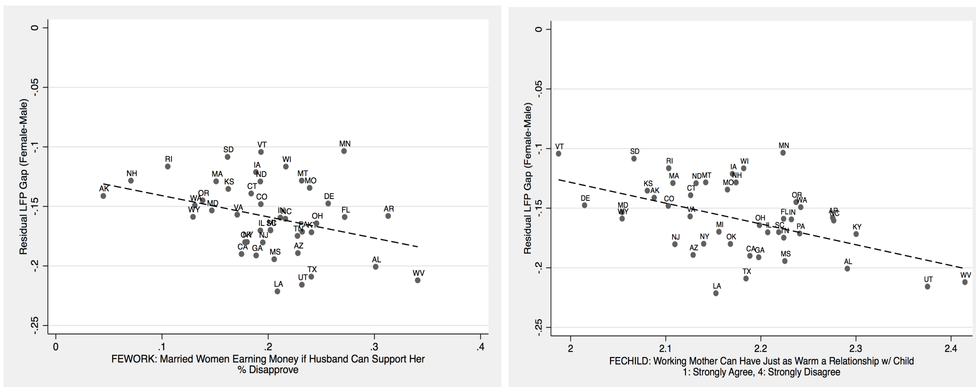
in business or industry if she has a husband capable of supporting her and (2) a working mother can establish just as warm and secure a relationship with her children as a mother who does not work. For both of these questions, there is a strong negative association between women's relative labor supply and age at first marriage, and the degree of disagreement with both statements (implying a higher degree of sexism).

The prevalence of sexist beliefs in the population seems naturally relevant for two forces that the literature has long speculated partly determine women's outcomes. One of these potential forces is gender discrimination. If their sexist beliefs lead some persons to take actions that make it more difficult or less financially rewarding for women to engage in a market activity compared to otherwise identical men, then women would experience a form of "taste-" or "prejudice-based" discrimination (Becker, 1957). Prevailing sexism could also affect a woman's outcomes by operating on or through her preferences. Since at least Smith (1759), economists have understood that an individual's utility function partly depends on the social influences to which they are exposed, and this idea has been formally analyzed in more recent theoretical work (Akerlof and Kranton, 2000; Benabou and Tirole, 2006; Acemoglu and Jackson, 2017). These social influences can take the form of gender norms that prescribe appropriate behaviors. Women might either (un)conciously adopt these norms or adhere to them in social interactions so as to avoid social sanctions. To the extent that these norms are internalized, women's preferences and beliefs may depend not only on norms to which they are currently exposed, but also to exposure they experienced as children or from their parents.

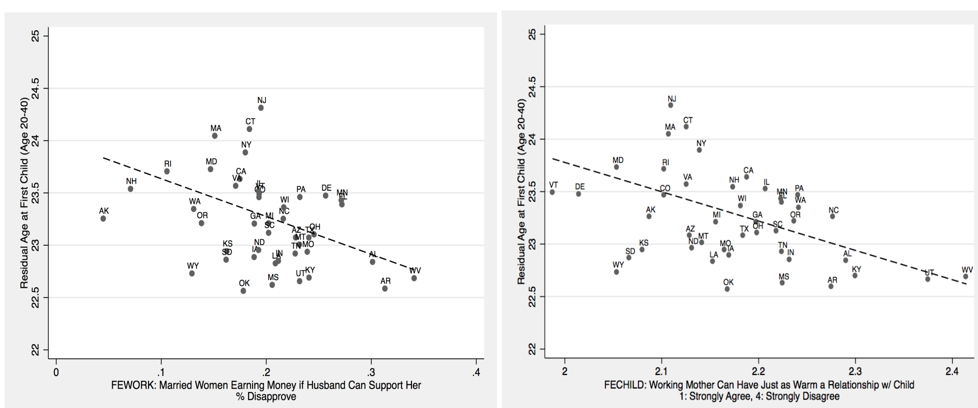
Two main empirical challenges confront efforts to understand the relationship between prevailing sexism and women's outcomes. The first is the extent to which the observed correlation between market-level sexism and women's outcomes reflects a causal relationship. While striking and suggestive, associations like those in Figure 1, could reflect reverse causality or the selective sorting of women (and men) with particular latent productive traits across areas with different levels of sexism. Assuming that one were able to identify the causal effect of sexism, the second empirical challenge is unpacking the roles of norms versus discrimination. Besides the fact that norms and discrimination likely simultaneously affect individuals, thereby making it difficult to isolate their separate effects, each can have effects in one setting (e.g. labor market, housing, education, etc.) that might interact with the effect of the other factor in another context.

Figure 1: Selected GSS Sexism Questions and Women's Outcomes

A. Outcome: Residual Labor Force Participation Gap (Female - Male)



B. Outcome: Residual Female Age at First Marriage



Note: LFP gaps (female-male) are estimated using MORG/ORG CPS data for the sample of whites aged 25 to 64. The female age at first marriage is estimated using Census/ACS data for women between the ages of 20 and 40. LFP gaps are residualized of gender-specific year effects, state fixed effects, individual years of schooling, and age dummies. Female age at first birth is residualized of year effects, individual years of schooling, and age dummies. The x-axis plots the mean responses to two questions from the 1977 to 1988 GSS.

This paper focuses on the effects of sexism that operate through gender discrimination in the labor market and through norms. To disentangle the effect of these two forces, we rely on the sharp set of specific predictions from classical discrimination theory about *which of women's various outcomes* should be especially affected by sexism-based market discrimination and *whose sexist beliefs* among specific persons in the market should drive such discrimination.[3] By contrast, theory places no similar, or at least very different, restrictions on the role of norms on women's outcomes. Differences in the associations predicted by theory allow us to examine the importance of the two factors.

Building on this idea, the paper develops a simple framework that distinguishes between sexism where a woman currently lives ("residential sexism") and the set of norms that she was exposed to in her formative years ("background sexism) for the sample of internal migrants within the U.S. (i.e. persons born in one state but currently living in another state).

The first part of our empirical analysis seeks to estimate how sexism in the market where a woman lives causally affects her outcomes. This effect reflects the influence of norms the woman currently confronts and perhaps internalizes, plus the effect of any sexism-based discrimination she experiences in the labor market. The ideal experiment to identify this causal effect would be to randomly assign to markets with different amounts of sexism, persons who are similar in terms of productive traits and their backgrounds. A particularly important dimension of a person's background that might be expected to affect outcomes is the amount of sexism they faced or were exposed to when growing up. Exposure to those beliefs during her formative years might exert a lasting influences on a person to the extent those views were internalized or affected skills that the person obtained while young. We call sexism where an adult women lives "residential sexism," and the sexism to which she was exposed to in her formative years "background sexism."[4]

We employ an instrumental variables (IV) strategy to approximate this ideal experiment. We first draw a sample of internal migrants in the U.S. from several waves of data from the Census and the American Community Survey (ACS). We then examine how outcomes differ among these adult

[3]Other forms of sexism-based discrimination that operate outside the labor market (e.g. in the housing market or in schools) do not share these predictions. In this paper, we will not be able to say much about these other forms of gender discrimination.

[4]Since we do not know the full history of agents' moves, we use state of birth to proxy for where people spent their early years. Even if individuals moved away from their state of birth at a young age, we argue that sexism in one's state of birth captures the sexism that parents are exposed to, and transmit to their children.

migrants who were born in states with similar sexism (that is, into similar background sexism) but who, because of differences in plausibly exogenous migration costs, are induced to move into states with different levels of residential sexism. The two sources of migration costs that we exploit rely on the geographic distance between markets and the settlement patterns of previous generations of migrants.[5] The instruments strongly predict variation in residential sexism, conditional on background sexism, even for internal migrants from within the same geographic region of the country.

We find that, holding background sexism constant, adults induced to move to more sexist places by the instrumental variables have larger gender gaps in labor force participation, are more likely to marry, and bear children younger. Conditional on employment, the TSLS results show no effects on the gender wage gap. These results are robust to controls for differences across states of birth in the probability of migration, as well as cross-state differences in school quality and women's relative employment opportunities due to historical industry structure. The effect sizes are economically large, ranging from LFP effects that are 1.2 times the standard deviation of cross-state gaps to half of the standard deviation of non-labor market outcomes for a standard deviation higher level of residential sexism.

The sample and specification used for the IV analysis also allow us to speak to the importance of sexist beliefs that people internalized when young on their adult outcomes. We find that among internal migrants who currently live in a given state, gender gaps in labor force participation and wages are larger the higher the level of sexism where they happened to be born. Women who were born in more sexist places are more likely to marry and have their first child at appreciably younger ages. These effects, while economically meaningful, are considerably smaller than the estimated effects of residential sexism. To allay concerns that the effects of background sexism are simply picking up other state-level characteristics, we show that background sexism does not appear to to have a direct impact on men's labor market outcomes. By contrast, higher levels

[5] The variation that under-girds these instruments has been used to study domestic and international migration, but not for the questions analyzed in this paper. See: Card (2001) and Munshi (2003) for influential work linking pre-existing settlement patterns of people from ethnic or national groups to the location decisions of subsequent waves of international immigrants. Several subsequent papers use of the same idea to instrument for immigrant inflows into particular cities (Card, 2001; Cortes, 2008; Cortes and Tesseda, 2011). In terms of domestic migration, Boustan (2010) exploits established settlement patterns among black migrants to instrument for black in-migration to different cities. The idea that distance is a fundamental cost of migration is an assumption in numerous models in urban economics and of migration. Among the papers that have used relative distance to instrument for where movers locate are Boustan et al. (2010) and Ortega and Peri (2014).

of background sexism significantly lower women's participation and wages.[6] Overall, these results indicate that background norms that women were previously exposed to continue to affect their adult life outcomes, even after they have left the market where that particular set of beliefs was most relevant.

The paper turns next to a series of tests intended to determine the relative importance of labor market discrimination versus residential norms for the observed effects of residential sexism on women's (relative) labor market and non-labor market outcomes. We first examine how the estimated effect of residential sexism varies across outcomes and by whether mean sexism is measured among men or among women. We reason that labor market discrimination should account for a larger part of the overall effect of residential sexism for labor market outcomes than for non-labor market outcomes, which it affects only indirectly. In addition, we expect any effect attributable to discrimination to be at least as strongly related to the sexism of men as to that of women, since in the case of gender-based discrimination, men are typically the prejudicial actors toward women in the labor market. For this part of the analysis, we measure labor market outcomes using the Current Population Survey (CPS), and create selection-corrected wages to approximate the offer wage distribution.

Regressing adults' outcomes simultaneously on average sexism among men and among women in the state of residence, we find that selection-corrected wage and labor force participation gaps are strongly related to mean male sexism among men but not to average sexism of women. By contrast, women's marriage rates and timing of first birth systematically vary only with average female residential sexism but not with men's mean sexism. These results are consistent with the idea that the effect of residential sexism on a woman's labor market outcomes is largely due to discrimination from men, whereas its effect on women's non-labor market outcomes derives mostly from her internalization of (or adherence to) the sexist norms where she resides.

To test whether residential sexism's effect on labor market outcomes is truly due to labor market discrimination, we draw on the sharp yet subtle predictions of the taste-based discrimination model (Becker, 1957).[7] Within the context of the employer discrimination model, assume that sexism

[6] While this does not entirely rule out the possibility of that unobserved state-level characteristics may confound the observed effects of background sexism on women's (relative) labor market outcomes, it does require such a confound to affect only women's outcomes but not men's.

[7] This model was subsequently analyzed by Arrow (1973), and more recently by Charles and Guryan (2008).

imposes a psychic cost on a man if he either interacts directly with or encounters a woman at work whose behavior is in contravention of his notion of women's appropriate roles. Prejudicial male employers, therefore, are only willing to hire a woman at a lower wage. Since it is costly when the two types of agents interact, the market acts to separate the prejudiced persons from the objects of their bias. This sorting mechanism guarantees that equilibrium prices will be determined not by the bias of the average, or "representative," prejudiced person but rather by a marginal discriminator whose prejudice is such that he is just indifferent between interacting with the disadvantaged group and not.

Since women constitute close to a half of all workers in most labor markets, the marginal discriminator would be drawn from somewhere close to the center of the male sexism distribution.[8] Gender wage differences across markets should be related to sexism close to the median of the male sexism distribution but not to prejudice at the top and bottom of the distribution, as these persons should be infra-marginal. Median sexism should also affect relative labor supply since any decrease in the market offer wage would reduce the number of persons for whom the wage exceeds their reservation wage.

Consistent with what a model of taste-based labor market discrimination against women by sexist men would predict, we find that the gender gaps in selection-corrected wages and in labor force participation in a market are (a) negatively related to the median of male sexism, but *not* to any other percentile of the distribution of male sexist beliefs; and (b) are not related to any percentile of female sexism. One potential concern is that male sexism is simply picking up the effects of some unobserved features of markets that also affect female wages and labor supply behavior.[9] Two things militate against this possibility in our study. The first is that virtually any such alternative account for what male sexism reflects would seem to imply that all, or at least several, moments of the "sexism" distribution should be systematically related to women's outcomes. Our finding that

[8] Charles and Guryan (2007) show that the relationship between employer distaste and wage gaps can survive perfect competition, as long as prejudicial tastes are portable across labor market roles (i.e. prejudiced employers continue to remain prejudiced when they shut down a firm and become workers). This formulation of the taste-discrimination model combines the employer and employee discrimination models and implies that since the labor market roles of "employer" and "worker" are endogenously chosen by market participants, what matters in determining women's relative labor market outcomes is not the distribution of prejudice among male employers at a point in time, but rather the distribution of prejudice among *all* men in a given market.

[9] For example, if women in a particular labor market, for whatever reason, do not wish to focus on work, they might be disinclined to invest in the sort of unobserved skills valued in the labor market. Men, living in those communities, when giving responses that we characterize as sexist may be merely reflecting that same community-wide sentiment rather than reporting something that would cause them psychic injury were they to encounter it.

the median sexism – and only the median – matters for women's market outcomes would seem to be inconsistent with any other explanation other than the standard prejudice argument. Moreover, in much of the analysis, we control directly for women's own responses to very same sexism questions put to men. If male responses merely indicate some general community sentiment (e.g. norms) distinct from discriminatory tastes, then these factors should load onto female responses to the questions in the various regressions. The fact that we find strong labor market effects for median male sexism, even after controlling for women's own views about the tasks they should engage in, suggests that the male sexism measures capture something close to the prejudicial feeling we have described.

The results for women's non-labor market outcomes are strikingly different. These outcomes are not related to *any* percentile of male sexism. They also are not related to the median of female sexism, varying systematically instead with other percentiles of female sexism in the market. The unimportance of the median compared to other percentiles of female residential sexism contrasted with the findings for median male residential sexism and labor market outcomes suggests that the latter results arise not because outcomes always load on the "middle" of a sexism distribution but rather from the causal mechanism of discrimination. It is also striking that non-labor market outcomes do not load onto any specific part of the distribution of female residential sexism. This pattern is perfectly consistent with the operation of norms which, unlike the sharp associations predicted by prejudice-based discrimination models, are not predicted to operate chiefly through any particular percentiles. Finally, we discuss several robustness checks that show that the results are not driven by the inclusion of particular states, or particular questions in the construction of the sexism index. We also show that cross-state differences in religiosity are not driving the observed effects of sexism on women's outcomes.

One body of work the paper extends is the previous literature studying gender-related beliefs and norms. A large literature in sociology has stressed the importance of norms for women's outcomes (Kiecolt and Acock, 1988; Burt and Scott, 2002). There is an active literature in economics on gender norms as well. Some of this work has examined how beliefs about gender roles affects women's outcomes across countries (Fortin, 2005) and over time within the U.S. (Fortin, 2010). Other work has examined how female employment in immigrant families' origin countries persist to later generations in the U.S. (Fernandez and Fogli, 2009; Blau et al., 2011). Alesina et al.

(2013) examine the cultural persistence of gender norms over the very long term. How norms affect decision-making within a family and women's labor supply choices have been explored by Fernandez et al. (2004), Bertrand et al. (2015), Bursztyn et al. (2017), and Olivetti et al. (forthcoming). Drawing the formal distinction between background and residential sexism and the mechanisms by which they operate; accounting for the endogeneity of norms one is currently exposed to; separating the effect of men's versus women's beliefs; and evaluating how sexism's influence differs across outcomes and over time are all ways in which our paper extends the previous literature.

We also extend the massive literature on discrimination - particularly work studying taste- (or prejudice-) based discrimination. Charles and Guryan (2008) show that blacks' labor market outcomes are related to different percentiles of white racial animus exactly as taste-based model of discrimination of Becker (1957) would predict. These specific predictions are very different from those we confirm regarding percentiles of sexism and women's outcomes. Our results help to bolster the argument that taste-based discrimination may be an important determinant of outcomes for disadvantaged groups in the economy.[10]

The remainder of the paper proceeds as follows. Section 2 describes the data and presents new descriptive facts that motivate the rest of the analysis. In Section 3, we describe residential and background sexism and outline the empirical framework for estimating their causal effects, discuss the two instrumental variables, and present first-stage results. Section 4 presents estimates of the causal effects of residential sexism and background sexism. Section 5 evaluates the relative importance of discrimination versus norms for the causal effect of residential sexism. Section 6 concludes.

2 Data and Summary Statistics

2.1 Cross-State Differences in Outcomes

We study four socioeconomic outcomes over the period from 1970 to 2017: women's labor force participation and hourly wages relative to those of observationally identical men, and the probability of marriage and age at first child among women aged 20-40. Information on marriage and childbearing

[10] For an alternative view, see Flabbi (2010), who estimates a search model of the labor market with matching, bargaining, employer's prejudice and worker's labor force participation decisions and argues that prejudice is not a relevant factor in explaining the slow-down in gender wage convergence in the U.S. in the 1990s.

comes from the 1980-2000 Decennial Censuses and from the 2012 three-year aggregate American Community Survey (ACS) (2010-2012), which we combine to create a "2012" sample. Because the Census/ACS does not track children across households, we infer how old a woman was when her first child was born from the reported age of her eldest child living in the same household. This part of the analysis is restricted to women aged 20 to 40 so that the oldest child is likely to be still residing in the mother's household. Whenever possible, we follow the convention in labor economics of using the Current Population Survey (CPS) to measure wages and labor force participation.[11] However, as we describe below, much of our analysis can only be done using Census/ACS data. The two data sources yield qualitatively similar results for various descriptive exercises, which is reassuring about the quality of the Census/ACS labor market data. As mentioned previously, we study only adult whites to avoid conflating issues concerning gender with that of race.

The paper focuses on persistent differences across states in women's outcomes since 1970. Table 1 presents cross-state summary statistics of the four outcomes studied over the entire study period, and separately for the early and later halves of that time. The left hand side of the table shows unadjusted versions of the statistics, and the right hand side presents conditional measures that have been regression-adjusted for age and education.[12]

Over the earlier half of our study period, gender wage and labor force participation gaps in the different states averaged 36% and 24.1 percentage points, respectively. Although these gaps are still disturbingly large, they have declined sharply since the mid-1990s, falling to 20.6% and 14.7 percentage points today, respectively. Women's non-labor market outcomes have also changed substantially. For example, whereas the share of women aged 20-40 in a state who have never married averaged 24% across states between the early 1970s and mid-1990s, during the later half of the study period, this mean rose to 41.8%. Similarly, the cross-state mean age of first birth among

[11] As Autor et al. (2008) note, the point-in-time wage and employment information provided by the May/ORG make it a superior source of labor market information, and especially information related to the distribution of wages, compared to the March CPS and Census/ACS, which provide only retrospective annual earnings information. See the Data Appendix for more details on the construction of the hourly wage measure.

[12] Specifically, the unconditional gaps in labor market outcomes are estimated based on a regression of LFP or log hourly wages on controls for gender-specific year effects, state fixed effects, and state*female effects for the time period in consideration. The coefficients on the state*gender dummies measure the unadjusted gaps (female-male) in labor market outcomes. The residual gaps are obtained from a similar procedure that includes additional controls for the number of years of schooling and age fixed effects. The unconditional non-labor market outcomes, which are computed only for women, are estimated based on a regression of a dummy for having never married or age at first birth on controls for year fixed effects and state fixed effects. The coefficients on the state dummies measure the unadjusted non-labor market outcomes. The residual non-labor market outcomes are obtained from a similar procedure that includes additional controls for the number of years of schooling and age fixed effects.

Table 1: Cross-State Summary Statistics of Women's Labor Market and Non-Labor Market Outcomes

	Panel A. Labor Market Outcomes					
	Unconditional			Residual		
	1977-2017	1977-1997	1998-2017	1977-2017	1977-1997	1998-2017
	(1)	(2)	(3)	(4)	(5)	(6)
	Female-Male LFP Gap					
Mean	-0.162	-0.241	-0.147	-0.159	-0.232	-0.148
SD	0.031	0.029	0.038	0.030	0.028	0.037
Max-Min	0.122	0.121	0.141	0.118	0.123	0.135
	Female-Male Wage Gap, Conditional on Working					
Mean	-0.254	-0.360	-0.206	-0.269	-0.361	-0.251
SD	0.036	0.039	0.038	0.033	0.036	0.033
Max-Min	0.191	0.189	0.189	0.172	0.169	0.174
	Panel B. Non-Labor Market Outcomes					
	1980-2012	1980-1990	2000-2012	1980-2012	1980-1990	2000-2012
	Female Share Nevermarried (Age 20 to 40)					
Mean	0.246	0.243	0.418	0.168	0.161	0.316
SD	0.053	0.052	0.054	0.053	0.049	0.059
Max-Min	0.227	0.219	0.234	0.217	0.194	0.243
	Female Age at First Birth (Age 20 to 40)					
Mean	23.58	23.58	24.28	23.67	23.58	24.37
SD	0.655	0.554	0.789	0.428	0.381	0.484
Max-Min	2.66	2.30	3.19	1.75	1.55	1.99
No. of states	44	44	44	44	44	44

Note: LFP and wage gaps (log hourly wages) are for whites aged 25-64 in MORG/ORG CPS. The unconditional gaps are obtained from the coefficients on the state*gender dummies in a regression of LFP or log hourly wages on gender-specific year effects, state fixed effects, and state*female effects. Residualized results are obtained using a similar procedure and include additional controls for years of schooling and age dummies. The non-labor market outcomes are estimated using Census/ACS data on white women aged 20 to 40 at time t. Unconditional outcomes are obtained from coefficients on state dummies in a regression of share nevermarried or age at first birth on year fixed effects and state fixed effects. Residualized versions are based on the same procedure including additional controls for number of years of schooling and age dummies. Data are from the sample of 44 states used in the main analysis (ie, states with necessary information in GSS on sexist beliefs). Means and standard deviations are weighted by the inverse of the variance of the outcome variable.

women aged 20-40 who have ever given birth grew by 0.7 years between the late and early halves of the time frame we study.

While these mean trends may be familiar from previous work, two things in the table are likely not as well-known. First, the close similarity between the regression-adjusted and raw statistics indicate that individual-level controls like age and education explain very little of average cross-state differences in women's outcomes. The other result in the table that has received very little previous attention is the remarkable stability of cross-state differences in all four outcomes. Whether measured in terms of the standard deviation or the max-min gap, neither raw nor regression-adjusted average state-level outcomes have converged over time. Instead, cross-state differences either remained essentially constant or widened slightly.

Not only have the size of mean cross-state gaps changed little over time, but each state's ranking in the cross-state distribution of outcomes has also been relatively constant. States where women's outcomes ranked in 1980s at the top, middle or bottom for any outcomes relative to other states held essentially the same relative positions throughout our period of study.[13] Formally, state×year effects add little further explanatory power in state-level regressions with state and time fixed effects.[14] Figure 1, which plots the outcomes over time after collapsing the states into nine Census divisions, illustrates the lack of both cardinal and ordinal convergence in women's outcomes across states.

2.2 Sexism Data From the GSS

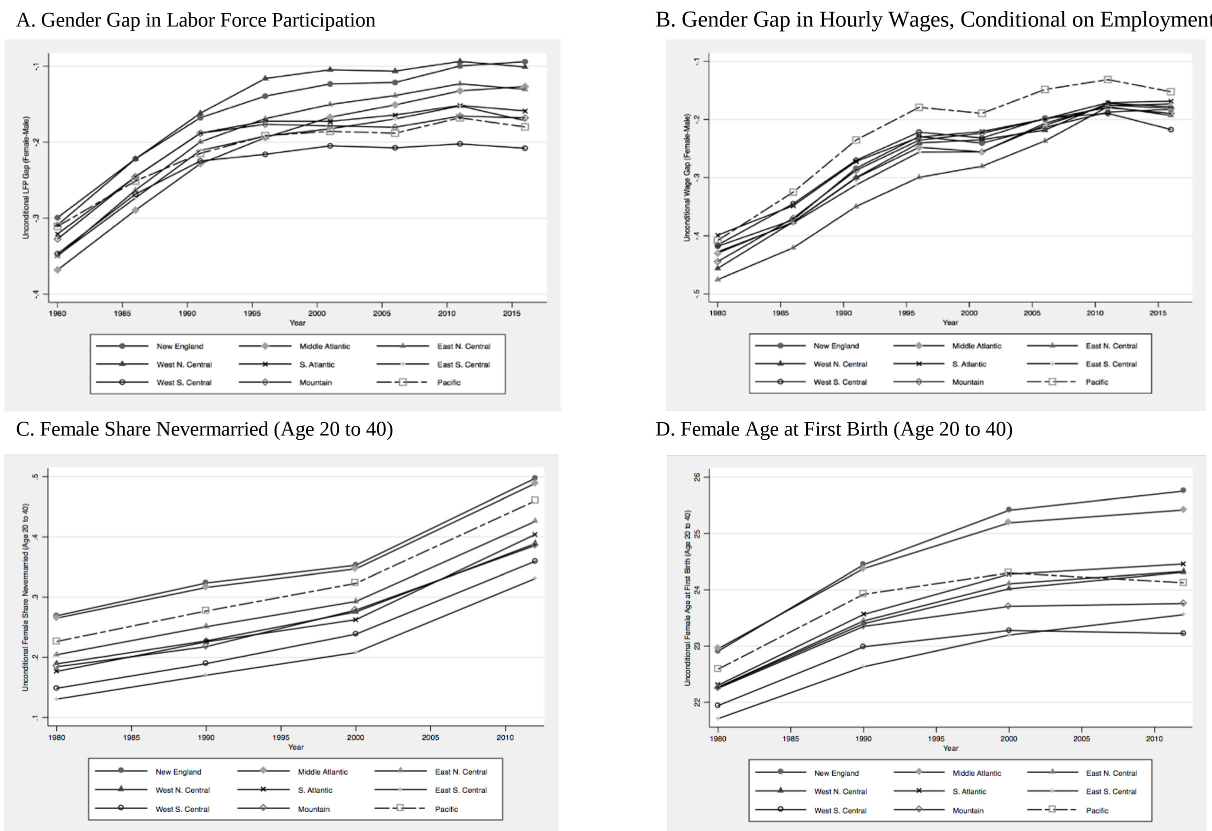
Our information on sexism comes from the General Social Survey (GSS). The GSS is a nationally representative survey that, for several years, has asked respondents various questions concerning their attitudes or beliefs about women's place in society. Our analysis uses responses to the eight questions, which were asked most consistently in the 1977-1998 waves of the GSS.

Two of these questions ("Women should take care of running their home and leave running the country up to men" and "It is much better for everyone involved if the man is the achiever outside the home and women take care of the home and family") elicit beliefs about women's and men's

[13]The correlation coefficient between states' ranking across all states between the first and last year of our study period for residualized labor force gaps, wage gaps, share ever married by age 40 and mean age of first child are, respectively, 0.72, 0.6, 0.93, and 0.84.

[14]See Appendix Table A1 for these results.

Figure 2: Mean Residual Labor Market and Non-Labor Market Outcomes by Census Divisions



Note: LFP and wage gaps are estimated using MORG/ORG CPS data for sample of whites aged 25 to 64. The non-labor market outcomes are estimated using Census/ACS data for women between 20 and 40. Labor market outcomes are residualized of gender-specific year effects, state fixed effects, individual years of schooling, and age dummies. Non-labor market outcomes are residualized of year effects, individual years of schooling, and age dummies. Means are for the nine Census Divisions over time.

appropriate roles inside and outside the home. Two other questions touch on beliefs about women's capacities ("Would you vote a female for President?" and "Are men better suited emotionally for politics than are most women?"). Finally, respondents' stated agreement or disagreement with statements such as "A working mother can establish just as warm and secure a relationship with her children as a mother who does not work," reveals beliefs about whether working mothers can juggle their dual roles effectively."[15]

We drop GSS respondents who are younger than 18, and recode the data so that larger values reflect higher sexism. Using their responses to all of the eight questions, we create a unidimensional index of a person's overall sexism. The sexism index subtracts off from the individual's response to each question the average response from the entire population to the same question in 1977, then divides by the standard deviation in the first year that the question is asked.[16] The one-dimensional aggregate individual-level sexism index is computed by taking the average of the normalized responses in each survey year for each individual.[17]

To obtain aggregate measures of sexism within a state, we first regress the individual-level sexism index on a full set of year dummies. Next, we use the residuals from this regression to create a measure of "average" sexism for each state, which is simply the mean across all years of the residual individual-level sexism index in a state.[18]

Similar to the labor market and non-labor market outcomes that we consider, there is substantial stability in cross-state differences in sexism over time, even as sexism has declined everywhere. As a result, we lose little useful variation by taking means over all years that the state is observed.[19]

[15]See the Appendix Table A2 for GSS variable abbreviation and summary for each of the eight questions about gender beliefs.

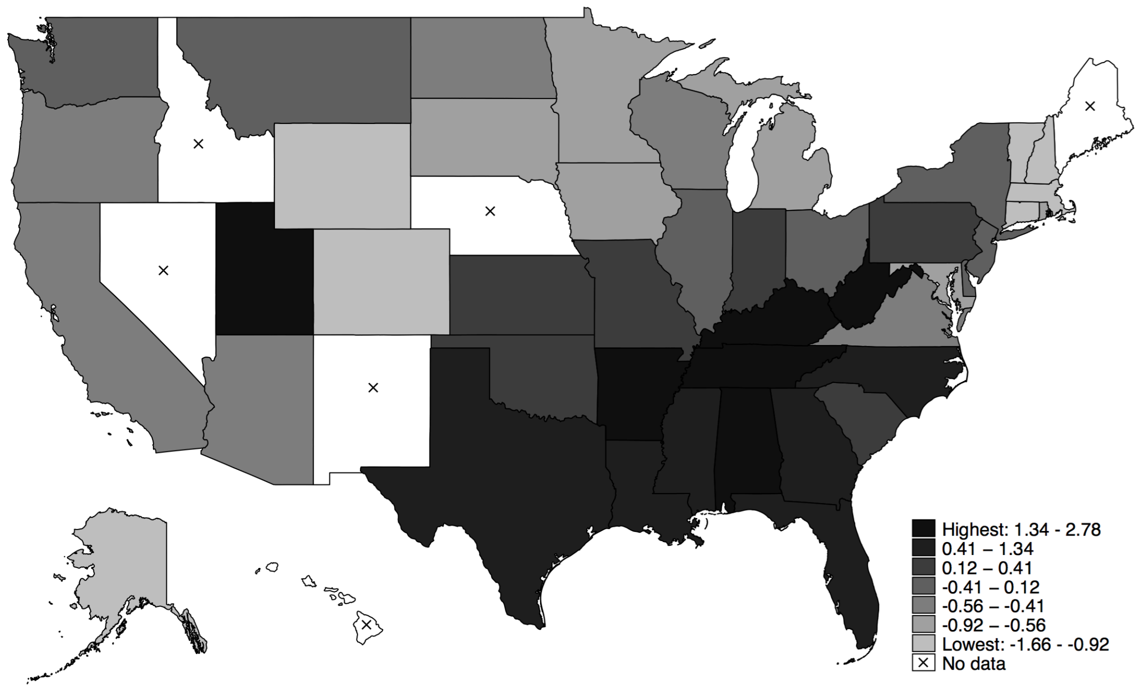
[16]This procedure is similar to that used by Charles and Guryan (2008) in their study of racial prejudice. We normalize by the mean and standard deviation of both male and female responses to the gender-prejudice questions rather than the gender-specific mean and standard deviation to ensure that the sexism index for both men and women are on the same scale.

[17]In additional robustness checks, available upon request, we check that the results are not sensitive to the inclusion of particular questions. Specifically, we create sexism indices dropping from the index one particular question at a time, and use each of these indexes to re-estimate the main specifications. Overall, the results using the leave-one-out indexes are qualitatively similar, suggesting that our main results are not driven by the inclusion of a particular question.

[18]We also construct similar measures for average male and female sexism in a state. For this, we regress the individual-level sexism index on the full set of year fixed effects separately by gender. We then average the residuals from this regression to create a measure of"average" male (female) sexism for each state. We also construct measures of the level of male and female sexism at different quantile points (e.g. the 10th, 50th, and 90th percentiles) of the distribution of the residual individual-level sexism index within a state.

[19]We see this from regressions that relate state-level sexism to state, year and state$\times$year fixed effects. Analysis of the R^2 indicates that state$\times$year fixed effects account for 13-20% of the variation over times in state level outcomes, indicating that there is relatively little *differential* variation in state means over time and thus little convergence. See

Figure 3: Mean Overall Sexism in U.S.



Note: Data are from several years of the General Social Survey (GSS). See text for further details.

This is important for our purposes since one limitation of the attitudinal data is the relatively small sample size. As such, we need to pool the GSS observations for each state over time in order to construct reasonable measures of sexism at the state-level (especially for the smaller states).[20]

Figure 3, which identifies states by their average overall sexism, shows the geographic distribution of sexist beliefs. Sexism is highest in the Southeast and least extreme in New England and the West. The figure shows that there is substantial variation in mean sexism across states *within* each geographic region of the country - a fact which we will exploit in the analysis to follow.

3 Estimating the Effect of Sexism on Women's Outcomes

Before turning to our formal analysis, we present some suggestive evidence of a potential link between measured sexism in one's state of residence and women's outcomes similar in spirit to the

Appendix Table A3 for regression results.

[20]Appendix Table A4 lists the value of the sexism measure and the number of observations used to construct the index for each state.

graphical associations shown in Figure 1. Table 2 presents simple correlations between the average level of sexism in a state and the state-level outcomes we study. Panel (A) shows the results using state-level gaps (female-male) in labor market outcomes and the average women's non-labor market outcomes, which have been adjusted only for gender-specific year effects. Panel (B) shows results after the underlying individual-level data has been residualized of age fixed effects and years of schooling.[21] The results show that higher mean overall sexism in a state is correlated with a larger gender participation gap, higher marriage rates and earlier childbearing among women in the state. Interestingly, the gender wage gap among workers in a state is not correlated with mean sexism in the state. The same pattern is evident in Panel (B) of the table.

These correlations, while striking, do not necessarily imply a causal link between average sexism in a state and women's outcomes. Individuals living in different states could have unobservable traits that are correlated with sexism levels in a state and directly impact labor market and non-labor market outcomes. In addition, cross-state differences in sexism could also capture the effects of other state-level characteristics that differentially affect women's labor market outcomes relative to men and women's non-labor market outcomes. Moreover, it is not clear how these simple associations relate to the mechanisms through which we believe sexism operates. In the next subsection, we define these constructs more precisely and outline the framework and methods we use to estimate and interpret the causal effect of prevailing sexism on women's different socio-economic outcomes.

3.1 Framework and Empirical Approach

3.1.1 Setup

Let br indicate the set of adults who were born in a particular state $b \in [b_1, b_2 \ldots, b_{50}]$ and who currently reside in a given state $r \in [r_1, r_2, \ldots, r_{50}]$.[22] Denote women in a br group by the indicator variable f_{br}. Assume that men's outcomes in the economy are determined exclusively by two (possibly unobserved) sets of factors: θ_b, which are individual productive traits like experience or ability; and φ_r which are characteristics of the market of residence, such as the quality of

[21] See footnote 13 for a description of the procedure used to compute the state-level outcomes.

[22] Our empirical work will use data from only 44 states because of data availability in the GSS. The following states did not have data available to construct the sexism index: Hawaii, Idaho, Maine, Nebraska, Nevada, and New Mexico. In our analysis, we also exclude DC since it is a clear outlier in terms of the key outcomes of interest.

Table 2: Mean Overall Sexism in State and Women's Outcomes

	LFP Gap (Female - Male)	Log Wage Gaps, conditional on working (Female-Male)	Share of Females Nevermarried	Average Female Age at First Child
	(1)	(2)	(3)	(4)
	A. Unconditional Outcomes			
Sexist Beliefs in State of Residence	-0.018***	0.000	-0.046***	-0.520***
	[0.004]	[0.007]	[0.006]	[0.080]
R-squared	44	44	44	44
Observations	0.292	0.000	0.469	0.406
	B. Residualized Outcomes			
Sexist Beliefs in State of Residence	-0.017***	-0.005	-0.040***	-0.279***
	[0.004]	[0.006]	[0.005]	[0.046]
R-squared	44	44	44	44
Observations	0.292	0.017	0.458	0.351
Sample	CPS 1977-2017, Age 25 to 64		Census/ACS 1980-2012, Age 20 to 40	

Note: The table reports coefficients (standard errors) from OLS regressions of various outcomes on mean sexism in a state. The labor market outcomes are estimated from white men and women aged 25-64 in 1977-2013 May/ORG CPS. Non-labor market outcomes in columns (3) and (4) use data on white women aged 20-40 from 1980 to 2000 Census and 2012 ACS. The unconditional outcomes only control for state fixed effects and gender-specific year effects in the construction of the state-level outcomes using the underlying individual-level data. The residualized outcomes additionally control for years of schooling and age dummies. Robust standard errors weighted by the inverse of the variance of the outcome reported in parentheses. ***significant at 1%, **5%, *10%.

public transportation, availability of childcare, or historical industry structure. By contrast, let women's outcomes depend not only on productive individual and market characteristics, but also on prevailing sexism, which affects them through two possible mechanisms: market discrimination and gender norms.

We assume that sexism among the people where she lives or has lived affects an adult woman through two types of mechanisms: norms and discrimination. We use the term "norms" to refer to the set of social conventions, mores and influences with which a person comes into contact. One type of norms, which we call *background norms*, N_b, are the social influences and mores in the place where she grew up and which affect the traits and preferences that a woman brings into her adulthood. Exposure to these norms during her formative years potentially determine a woman's preferences through internalization or unconscious emulation. In addition, through pre-market discrimination, norms where she was born and raised might have led people and institutions there to provide her with different training or education during her youth relative to what was provided to boys, or caused her to make different human capital investment relative to boys.[23]

Apart from the skills and elements of her preferences that were determined earlier in her life partly from exposure to background norms, two other forces determine her adult outcomes. One of these is the influence of what we call *residential norms.* These are social mores and practices in the market where she currently lives. The second force operating on an adult woman of a given level of skill is *market discrimination*, D_r. Following the standard formulation in economics, we define market discrimination as the amount by which others' actions reduce the economic payoff an adult woman gets from market activity relative to that of an otherwise identical man doing the same thing.[24]

We assume that these three forces - the things that determine preferences (and skills) a woman brings into adulthood, the influence of social mores she encounters as an adult, and any market discrimination she faces - are functions of prevailing sexist beliefs in the relevant markets. Specifically,

[23] In the empirical analysis, since we focus on women's outcomes net of education and age, we expect that the effects of background norms that we estimate is potentially a *lower* bound of the true impact.

[24] The paper studies discrimination against adult women in the labor market and not in other settings, such as the housing market.

we assume that

$$
\begin{aligned}
N_r &= \delta^r\ \overline{S}_r + u_1 \\
D_r &= \ \beta\ \overline{S}_r + u_2 \\
N_b &= \delta^b\ \overline{S}_b + u_3
\end{aligned}
\tag{1}
$$

where $\overline{S}_r$ is average overall sexism where lives and $\overline{S}_b$ is mean overall sexism where the woman was born. Residential sexism is an element of the norms that she's currently exposed to and is the basis upon which some persons discriminate against her when she engages in labor market activity. These effects are captured by the terms β and δ^r. Although she no longer lives there, background sexism potentially affects her current outcomes because the norms where she grew up or was exposed to while growing up (e.g. due to parental influences) may have affected enduring aspects of her latent preferences and skills. This effect is captured by the parameter δ^b. The mean-zero terms u_1, u_2, and u_3 represent all other determinants of discrimination and the two types of norms.

Mean outcomes for br persons, Y_{br}, are given by the sum of the "neutral" components θ_b and φ_r, which affect both men and women, plus a portion attributable to discrimination and norms, which affects only women. That is, mean outcomes for men (m) are women (f) can be written as:

$$
\begin{aligned}
Y^f_{br} &= f_{br}(D_r + N_r + N_b) + \theta_b + \varphi_r && (2) \\
Y^m_{br} &= \theta_b + \varphi_r
\end{aligned}
$$

For labor market outcomes (i.e. LFP and wages), we consider gender gaps ($Y^f_{br} - Y^m_{br}$), while we use mean outcomes for women for the non-labor outcomes (i.e. share nevermarried and age at first child). Substituting from (1) and assuming a linear function for $f_{br}(\cdot)$ yields the following baseline specifications for labor market gaps and women's non-labor market outcomes, respectively:

$$
\begin{aligned}
Y^f_{br} - Y^m_{br} &= \alpha + (\beta + \delta^r)\overline{S}_r + \delta^b\overline{S}_b + v_{br} && (3) \\
Y^f_{br} &= \alpha + (\beta + \delta^r)\overline{S}_r + \delta^b\overline{S}_b + \varepsilon_{br}
\end{aligned}
$$

where v_{br} and ε_{br} are random, mean-zero statistical errors. We seek to estimate the causal parameters β, δ^r and δ^b. In some specifications, we include observable controls such as school quality, and women's relative employment opportunities based on historical industry composition, to proxy for state of birth and state of residence characteristics (θ_b and φ_r), and relax the assumption that these components affect only women. Note also that throughout the analysis, we will consider state level outcomes that are net of individual characteristics such as years of schooling and age. This serves to capture state of birth characteristics that may affect individuals' latent productive traits.

3.2 Estimating the Causal Effects of Residential Sexism

To link prevailing sexism in a woman's place of residence to her relative labor market and non-labor market outcomes, we begin by using the framework above to estimate the causal effects of residential sexism ($\overline{S}_r$). Nevertheless, OLS performed on equation (3) is unlikely to yield unbiased estimates of the causal effect of residential sexism due to endogenous migration. In particular, women might choose where to live based on locations' levels of sexism directly or on the basis of other market characteristics that are correlated with market sexism. Residential sexism could therefore be endogenous in a naive OLS regression like (3), so that the observed relationship between higher sexism in a market and the outcomes of women living there reflects both the causal effect of sexism plus the fact that women with particular latent traits are more likely to live in such places in the first place. We use Two Stage Least Squares (TSLS) methods to deal with the possible endogeneity of residential sexism. Using TSLS also addresses any measurement error bias arising from the fact that our analysis treats a person's labor market as their state of residence which almost certainly mis-measures their true labor market - the place where they live, work and socially interact plus the causal effect of internalized residential norms - $(\beta + \delta^r)$. In the later parts of our analysis, we attempt to shed light on the importance of discrimination and internalized norms. For now, our focus is on the more fundamental problem of obtaining a causal estimate of $\overline{S}_r$ in equation (3). Below, we outline our approach for isolating exogenous variation in residential sexism for use in the TSLS analysis.

3.2.1 Determinants of the Migration Decision

The potential endogeneity problem in (3) for which TSLS is intended to correct is that the level of residential sexism observed for migrants born into a given level of background sexism, is a function of the latent productive factors θ_b and φ_r.[25] We assume that, in general, the residential sexism chosen by migrants born in state b, who therefore were exposed to background sexism $\overline{S}_b$, may be written as:

$$\overline{S}_r(b) = \alpha + \gamma \overline{S}_b + \rho_1 Z_1(\Delta_{bj}) + \rho_2 Z_2(\kappa_{bj}) + \xi_{br}, \tag{4}$$

where ξ_{br} is a random error term. The two terms Z_1 and Z_2, are functions that depend on fixed costs of migration that must be incurred by anyone (whatever their individual traits) who moves between locations. They are the source of exogenous variation in the TSLS analysis and are assumed to be independent of θ_b and φ_r.

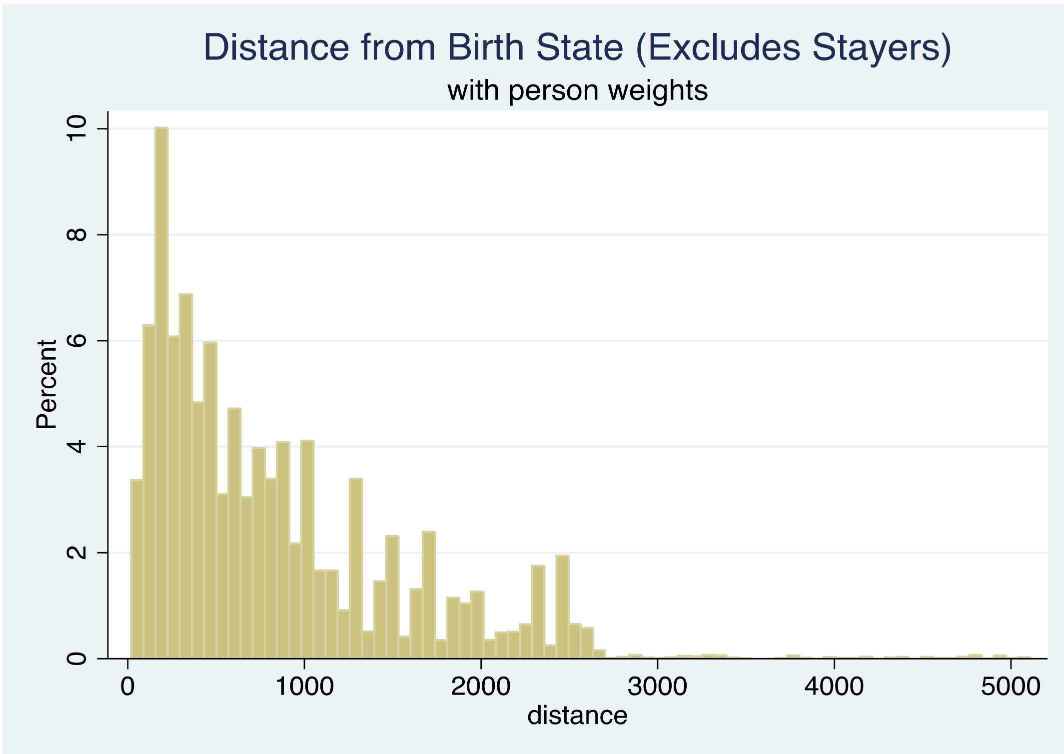
The variable Δ_{bj} is a vector measuring the distance between the state b from which a migrant moves to every other state j (measured as the distance between the states' population-weighted centroids). Relative distance is a fixed migration cost, meant to reflect the fact that it is simply logistically more difficult for people to move farther from their initial locations, irrespective of their individual traits or features of the two markets (see, for example, Boustan et al. (2010) and Ortega and Peri (2014)).

Figure 4 assesses the plausibility of this assumption for internal migrants in the U.S. whom we study. Focusing on all adults across multiple Census years living in states different from where they were born, the figure is a histogram of the distances between the migrants' states of residence and states of birth. Since other factors besides relative distance determine migrants destinations, the plotted relationship does not decline perfectly smoothly. Nonetheless, the figure clearly shows that internal migrants in the U.S. are systematically less likely to move to states that are farther away, irrespective of the state from which they are moving. If the probability that a migrant leaving b ended up in j, in fact, depended systematically *only* on how relatively far j is from b compared to other other destinations, plus some random error, the expected level of residential sexism observed among migrants from b would be given by

$$Z_1(\Delta_{bj}) = \sum_{b \neq j} \overline{S}_j / \Delta_{bj}. \tag{5}$$

[25] These are subsumed in the error term in equation (3).

Figure 4: Distance of States Residence from States of Birth Among U.S. Migrants



Note: Data are from the 1980 to 2000 Census and 2010-2012 ACS. The sample is restricted to internal migrants. See text for details.

An equivalent way to think of $Z_1(\Delta bj)$, then, is as the portion of the expected residential sexism for migrants moving from states with sexism $\overline{S}_b$ that is attributable to the exogenous relative spatial location of various states.

The existing literature suggests that where previous generations of migrants from a given initial location have historically settled represents a second fixed cost of migration for the subsequent movers. Various papers, mostly studying international migration, posit that the "enclaves" formed by previous migrants affect the ease with which later movers can adjust to different possible destinations (see, for example (Card, 2001)). These historical migration patterns thus represent a plausibly exogenous driver of the location choices of current migrants, since the factors that drove previous waves of migrants to settle where they did are arguably otherwise unrelated to later migration decisions. To further ensure the exogeneity of the instrument based on migration costs, we use migration patterns of *single men* age 25 and above in 1930 and 1940 to construct the network since where single men decide to migrate is unlikely to be determined by concerns about sexism.

We proxy for existing enclaves that migrants confronted using the share of single men born in state b who lived in a different state j in 1930 or 1940, or κ_{bj}. Using the same logic used for relative location of a birth state, the portion of current location decisions of migrants attributable to exogenous pre-existing migration patterns is given by

$$Z_2(\kappa_{bj}) = \sum_{b \neq j} \overline{S}_j \times \kappa_{bj}. \tag{6}$$

Given the foregoing, equation (4) represents the first stage regression for the endogenous variable $\overline{S}_r$ in equation (3), where $Z_1(\Delta_{bj})$ and $Z_2(f_{bj})$ are instrumental variables whose "strength" is captured by the parameters ρ_1 and ρ_2. Two Stage Least Squares (TSLS) regression performed on this system allows us to estimate the causal effect of residential sexism - $\beta + \delta^r$. To ensure a finer comparison, in all the specifications, we also include fixed effects for region of birth (nine Census divisions). This ensures that the comparison is limited to internal migrants that are born into states with similar levels of background sexism that are in the same Census division, but end up residing in states with different levels of sexism due to differential costs of migrating to different states (as proxied for by relative distance and the presence of historical networks).

3.3 First-Stage Results

The two panels of Figure 5 illustrate the variation that the two instrumental variables isolate. The x-axis of Figure 5a is mean sexism in different states of birth. Although there is substantial variation in background sexism, many states have very similar levels of $\overline{S}_b$. For example, migrants from Massachusetts and Maryland were born into very similar background sexism. The y-axis in the figure plots the value of the instrumental variable based on distance, $Z_1(\Delta_{bj})$.

The figure shows that the component of migrants' destination decision attributable exclusively to the exogenous accident of where their state of birth is located geographically relative to all other states predicts that migrants from Maryland and Massachusetts would move to markets with very different levels of sexism. Looking across the figure, one sees similar variation for different pairs of states with similar background sexism - Pennsylvania/Louisiana, Tennessee/West Virginia and New York/Ohio to name some examples. Figure 5b repeats this exercise for the instrumental variable $Z_2(\kappa_{bj})$, which is based on the amount of historical migration flows (among single men) from the origin state (pre-existing "enclaves"). Here also one sees substantial variation in predicted residential market sexism, conditional on the level of background sexism from which the migrant comes.

Whether $Z_1(\Delta_{bj})$ and $Z_2(\kappa_{bj})$ are valid instruments for the TSLS analysis depends on how much of the variation in $\overline{S}_r(b)$ they explain, holding constant $\overline{S}_b$. The first stage results in Table 3 answers this question. We conduct this analysis on the main sample of internal migrants in the Census/ACS, born in states of birth b and living in different states r. We collapse the more than 6 million individual observations on adult male and female migrants down to (state of birth $\times$ state of residence $\times$ gender $\times$ year) cells, yielding slightly more than 15,000 observations once empty br cells with no reliable sexism data are dropped.[26] The regressions using the collapsed data are weighted using the number of observations in each cell.

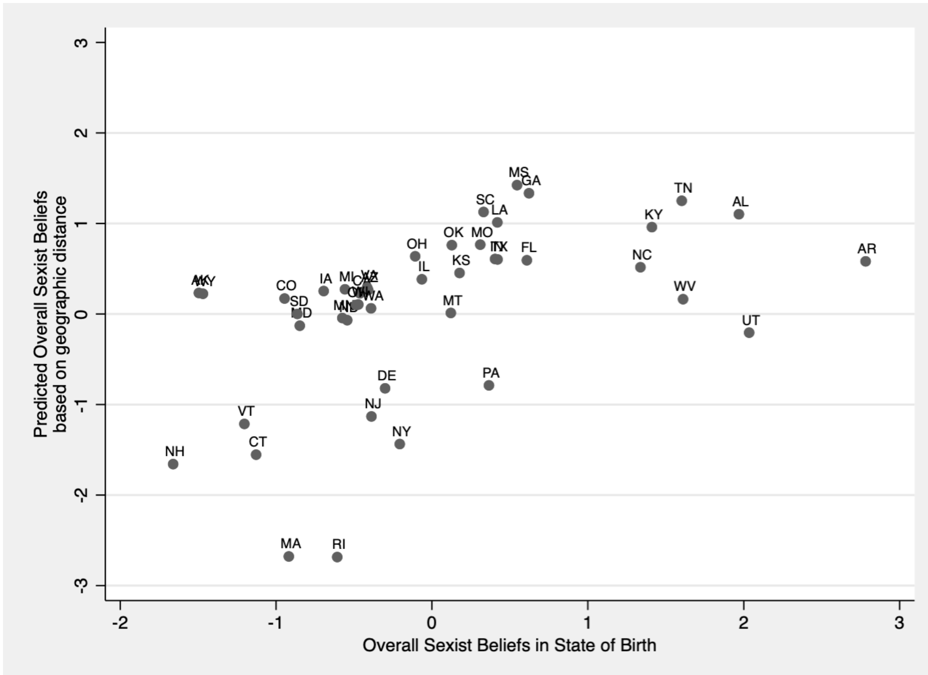
The regressions in Table 3 relate sexism in the market to which a given group of migrants move to the two instrumental variables. All the models in the table control for the migrant group's background sexism, and include fixed effects for the Census region of state b, interactions between region effects and gender fixed effects, and interactions between year fixed effects and gender dummies. Standard errors are clustered at the state of residence level.[27] The results in the first two columns show that each of the IVs is separately strongly related to residential sexism, and explains a sufficiently large portion of the variation in the potentially endogenous regressor to remove any "weak instrument" concern. This is especially true of the "enclave" or "historical settlement pattern" measure, for which the F-stat of 27 far exceeds the conventional threshold of 10. When the two variables are entered simultaneously, only the enclave IV remains strongly significant. This is not surprising; where migrants settled in the past was likely based on temporary factors from that time period *plus* the historically important propensity to not move too far from home. Again, the F-stat on the pair of IVs is comfortably above 10.

One might be concerned that the IVs could be correlated with the labor market traits or family and marriage preferences of migrants. For example, perhaps being (exogenously) surrounded by a

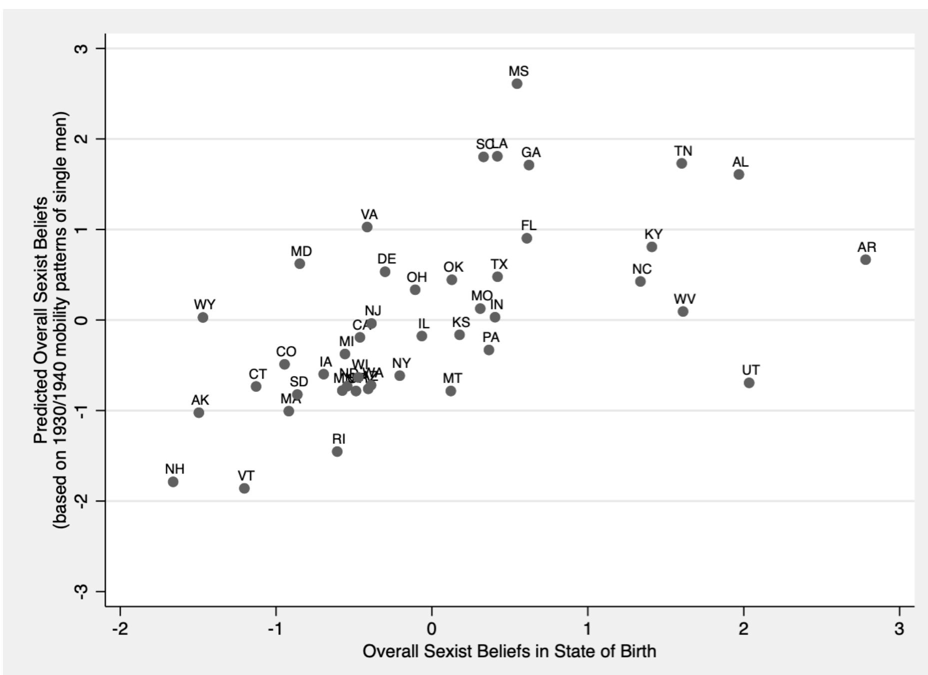
[26] It should be noted that we use different samples for the labor market and non-labor market results. The latter uses only women, so has about half the number of observations after collapsing the individual-level data. We conduct a separate first-stage analyses using the sample on which the subsequent TSLS analysis is performed. The first stage results are very similar across samples.

[27] The results are essentially unchanged if we calculate standard errors using two-way clustering by state of birth and state of residence.

Figure 5: Predicted Residential Sexism based on: (a) Distance Between State of Birth and Other States and, (b) Historical Destination of Previous Migrants from State of Birth



(a) Predicted Residential Sexism Given Relative Location



(b) Predicted Residential Sexism Given Historical Migration

Table 3: First Stage Results: Effect of Relative Distance and Enclave Instruments on Residential Sexism

	LFP sample			
	Outcome: Overall Sexism in the State of Residence			
	(1)	(2)	(3)	(4)
Distance IV	0.269***		0.095	0.096
	[0.070]		[0.068]	[0.068]
Enclave IV		0.259***	0.212***	0.214***
		[0.050]	[0.046]	[0.045]
F-stat of excluded instruments	14.6	27.0	14.0	14.9
Controls:				
Outmigration Rate				X
Overall Sexism in State of Birth*Female	X	X	X	X
Overall Sexism in State of Birth	X	X	X	X
Region of Birth FE*Female	X	X	X	X
Female*Year FE	X	X	X	X
Observations	15,103	15,103	15,103	15,103
R-squared	0.074	0.080	0.081	0.081

Note: The data are from the 1980, 1990, 2000 US Census and 2010-2012 ACS (3-year aggregate). The sample is restricted to whites age 25 to 64 for the labor market outcomes and age 20 to 40 for the non-labor market outcomes (females only) who are not currently living in their state of birth. The unit of observation is at the state of birth*state of residence*gender*year level for the labor market outcomes and at the state of birth*state of residence*year level for the non-labor market outcomes. Sexist beliefs are normalized to have a mean of 0 and standard deviation 1 across the 44 states with available information to construct the index of sexist beliefs in the GSS (excluding DC). The Distance and Mobility IV are standardized to have mean 0 and standard deviation 1 in the sample of states. Regressions are weighted by the number of observations in each cell for each outcome. Standard errors clustered at the state of residence level are in parentheses. ***significant at 1%, **at 5%, *at 10%.

particular level of average sexism in nearby states causes only certain types of people to move in the first place compared to people in places where the nature of sexism in nearby states is different. Our analysis compares migrants who grew up in the same level of sexism. Suppose that persons from states with similar background sexism have the same distribution of preferences regarding the sexism they would like to live under as adults, all else equal. Then, to the degree that the propensity to migrate in the first place is driven by sexism in neighboring states, the fraction of people migrating from two states with similar sexism but with different levels of the IVs should differ systematically. The regressions in the fourth column of the table adds the out-migration rate from each origin state b as a control to account for at least some of the latent traits of migrants that could be correlated with the IVs. The results show that this control does not change the main conclusion of the first stage analysis.[28]

4 Estimates of the Causal Effects of Mean Residential Sexism and Background Sexism

In this section, we present TSLS estimates of the effect of residential sexism on migrants' outcomes using both the relative distance and historical settlement patterns as instruments. In principle, these results account for endogeneity and measurement error problems in naive OLS models, which are also shown in the table for easy comparison. For the labor market outcomes (labor force participation and hourly wages conditional on employment), we estimate variants of equation (3) instrumenting for $\overline{S}_r$ using both predicted sexism based on relative-distance and historical settlement patterns as instruments.[29] Specifically, we estimate 2SLS regressions of the form:

$$y^k_{brt} = \alpha + \delta^r(f_{brt} \times \overline{S}_r) + \delta^b(f_{br} \times \overline{S}_b) + \gamma^r\overline{S}_r + \gamma^b\overline{S}_b + \theta_{ft} + R_{fb} + u_{br} \tag{7}$$

where f is a female indicator and $\overline{S}_r$ and $\overline{S}_b$ are state-level proxies for residential sexism and background sexism, respectively. θ_{ft} and R_{fb} represent the full set of gender-specific year fixed effects and region of birth fixed effects, respectively. As mentioned previously, we collapse the

[28] In the TSLS below, the results are essentially unaffected by which of the four first-stage specifications shown in the four columns of Table 3 the analysis uses. Results for these various exercises available upon request.

[29] TSLS results from models that use only one of the instruments are qualitatively similar and are available on request.

individual observations (adjusted first for individual-level differences in years of education and age) in the Census/ACS down to the state of birth $\times$ state of residence $\times$ gender $\times$ year level. The estimation equation for women's non-labor market outcomes is similar, except that the interactions with the female dummy are omitted. In (7), γ^r is the estimated effect of residential sexism on migrant men's adult outcomes, while δ^r measures the differential effect for women. In this TSLS framework, these causal effects measure how their outcomes are affected for migrants induced to live in a particular level of residential sexism because of the instruments. The regressions control for mean sexism in the state of birth so the estimated residential sexism are from comparisons of persons born into similar background sexism. For this portion of the analysis, standard errors are clustered at the state of residence level.[30] Note that we can also infer the effects of background sexism on men's and women's outcomes from this specification based on the coefficient estimates of γ^b and δ^b.

Tables 4 and 5 present the regression results for the labor market and non-labor market outcomes, respectively. The labor market results in Table 4 indicate that living in a more sexist state has no effect on the likelihood that a man participates in the labor force. This result holds for both the OLS estimates and the baseline 2SLS estimates reported in column (2). By contrast, we find that higher residential sexism significantly decreases the relative labor force participation rate of women compared to men who were born into the same level of background sexism. The estimated effects of residential sexism are economically large – a standard deviation increase in sexism in one's state of residence (approximately the difference between residing in California rather than Mississippi, or in Minnesota rather than Texas) lowers women's labor force participation relative to men by about 3 percentage points. This is approximately 15% of the mean unconditional labor force participation gap across states of residence and 1.2 times the standard deviation of gaps across states.

Column (3) shows that the results remain similar when we include controls for the outmigration rate from each origin state b (interacted with the female dummy), which suggests that differential likelihood of migration (which may influence patterns of selection among migrants) across states, with the same level of background sexism, are unlikely to be driving the results. In column (4), we

[30]The results are virtually unchanged if standard errors are computed using two-way clustering by the state of birth and state of residence.

include additional controls (interacted with the female dummy) for state-level characteristics that might be correlated with predicted residential sexism such as proxies for skill quality in one's state of birth and state of residence (mean National Assessment of Educational Progress – Long Term Trend (NAEP-LTT) reading and math test scores for boys) as well as women's relative employment opportunities due to historical industry structure in one's state of residence.[31,32] The qualitative results remain similar, and if anything, the magnitudes of the point estimates become larger with the inclusion of these additional controls.[33] The results in columns (5) to (8) of Table 4 suggest that most of the impact on residential sexism on women's relative labor market outcomes operates largely through the extensive margin of participation. Conditional on employment, we find little evidence that higher residential sexism increases gender wage gaps or has a significant effect on men's wages, especially once controls for outmigration rates and other state-level characteristics are included (see column (8)).

Turning to the effects of background sexism, we find that higher sexism in one's state of birth is uncorrelated with men's labor force participation rates, but significantly reduces women's labor force participation relative to men (see columns (1) to (4) of Table 4). The magnitude of the estimates are similar across all the specifications and are about one-fifth the size of the estimated impacts of residential sexism. Interestingly, unlike the participation results, we find evidence of a larger adverse effect on women's wages relative to men conditional on employment among migrants who were born into more sexist states. While it is impossible for us to entirely rule out the possibility that the observed effects of sexism on women's *relative* outcomes are driven by unobserved state-level characteristics, it is worth noting that, across all specifications, we find little consistent evidence that background sexism (as well as residential sexism) has a direct negative impact on

[31]The NAEP-LTT is a set of standardized tests given to a random sample of U.S. 9-, 13- and 17-year old students enrolled in public schools. Tests in math or reading have been given almost every other year since 1971. We calculate the normalized scores (i.e. z-scores) for white boys in each state for math and reading over the full sample of years (1971-2004 for reading, 1978-2004 for mathematics). The NAEP-LTT was also administered in mathematics in 1973, but we do not have access to state identifiers for those data.

[32]To capture cross-state differences in historical industry composition which may result in differential female employment opportunities across state, we calculate the predicted female employment share in each Census/ACS decade based on the 1950 industry structure (12 industry groups) at the state-level: $D_{st} = \sum_j s^f_{jt} \frac{E_{1950,js}}{E_{1950,s}}$, where s^f_{jt} is the fraction female in each industry j at time t and $\frac{E_{1950,js}}{E_{1950,s}}$ is the employment share of each industry in each state in 1950. The overall predicted female employment share used in the analysis is the average over the four decade time period from 1980 to 2012.

[33]The reduction in the number of observations in columns (4) and (8) is due to the fact that NAEP math scores are not available for Alaska, New Hampshire, and Vermont.

Table 4: The Effects of Residential Sexism and Background Sexism on Labor Market Outcomes

	Labor Force Participation				Log Wages, conditional on working			
Instruments:	OLS	2SLS (Distance + Enclave IV)			OLS	2SLS (Distance + Enclave IV)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Sexist Beliefs in State of Residence:								
Overall*Female Dummy	-0.011**	-0.035***	-0.034***	-0.065**	-0.004	0.027	0.008	0.002
	[0.004]	[0.012]	[0.011]	[0.032]	[0.008]	[0.024]	[0.022]	[0.046]
Overall	-0.005	0.002	0.004	0.024	-0.067***	-0.060*	-0.043	0.028
	[0.003]	[0.006]	[0.006]	[0.020]	[0.014]	[0.030]	[0.029]	[0.068]
Sexist Beliefs in State of Birth:								
Overall*Female Dummy	-0.008***	-0.008***	-0.008***	-0.009***	-0.014***	-0.014***	-0.016***	-0.015***
	[0.002]	[0.002]	[0.002]	[0.002]	[0.003]	[0.003]	[0.003]	[0.003]
Overall	0.001	0.001	0.001	0.002	0.003	0.003	0.004*	0.006*
	[0.001]	[0.001]	[0.001]	[0.001]	[0.002]	[0.002]	[0.002]	[0.003]
Controls (all interacted with Female dummy):								
Additional controls				X				X
Outmigration Rate in State of Birth			X	X			X	X
Region of Birth FE, Year FE	X	X	X	X	X	X	X	X
Observations	15,103	15,103	15,103	13,107	15,103	15,103	15,103	13,107
R-squared	0.917	0.908	0.909	0.900	0.917	0.908	0.909	0.900

Note: The data are from the 1980, 1990, 2000 US Census and 2000-2012 ACS (3-year aggregate). The sample is restricted to white individuals age 25 to 64 who are not currently living in their state of birth. The unit of observation is at the state of birth*state of residence*gender*year level. Sexist beliefs are normalized to have a mean of 0 and standard deviation 1 across the 44 states with available information to construct the index of sexist beliefs in the GSS (excluding DC). In Columns (2) to (4) and (6) to (8), sexist beliefs in the state of residence (Overall*Female and Overall) are instrumented using predicted sexism based on geographic distance and historical migration (i.e. $Z_1(\Delta_{bj})$*Female, $Z_2(\kappa_{bj})$*Female, $Z_1(\Delta_{bj})$, and $Z_2(\kappa_{bj})$). "Additional controls" include the mean NAEP-LTT reading and math scores of boys in each state of residence and state of birth as well as the predicted female employment share in the state of residence based on the state's 1950 industry structure (all interacted with the female dummy). All the regressions also control for gender-specific region of birth fixed effects and gender-specific year fixed effects. Regressions are weighted by the number of observations in each cell for each outcome. Standard errors clustered at the state of residence level are reported in parentheses. ***significant at 1%, **at 5%, *at 10%.

men's labor market outcomes. This suggests that any such confound would have to have an effect only on women's outcomes but not men's, which seems unlikely.[34]

Table 5 presents the results for migrant women's non-labor market outcomes. The sample is restricted to women between the ages of 20 and 40. We find that, compared to other women who were born into similar levels of sexism, female migrants residing in more sexist states are more likely to have ever married and bear their first child at appreciably younger ages. These patterns are found both in the OLS regressions and in the preferred TSLS models. Background sexism appears

[34]To tease out the effects of background sexism, we can also estimate an alternative, more flexible, specification that includes the full set of state of residence fixed effects interacted with the female dummy. As shown in columns (1) to (4) of Appendix Table A5, the estimated effect of background sexism on women's relative labor market outcomes from this specification are largely similar, albeit slightly smaller and somewhat less precisely estimated (still significant at the 10% level).

Table 5: The Effects of Residential Sexism and Background Sexism on Non-Labor Market Outcomes

	Never Married (Age 20 to 40)				Female Age at First Child (Age 20 to 40)			
Instruments:	OLS	2SLS (Distance + Enclave IV)			OLS	2SLS (Distance + Enclave IV)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Sexist Beliefs in State of Residence:								
Overall	-0.029***	-0.047***	-0.044***	-0.065***	-0.197***	-0.156*	-0.209**	-0.200
	[0.006]	[0.011]	[0.011]	[0.023]	[0.035]	[0.089]	[0.089]	[0.160]
Sexist Beliefs in State of Birth:								
Overall	-0.015***	-0.015***	-0.015***	-0.014***	-0.130***	-0.129***	-0.133***	-0.126***
	[0.002]	[0.002]	[0.002]	[0.002]	[0.018]	[0.017]	[0.017]	[0.016]
Controls:								
Additional controls				X				X
Outmigration Rate in State of Birth			X	X			X	X
Region of Birth FE	X	X	X	X	X	X	X	X
Year FE	X	X	X	X	X	X	X	X
Observations	7,522	7,522	7,522	6,534	7,453	7,453	7,453	6,491
R-squared	0.698	0.675	0.682	0.674	0.481	0.480	0.482	0.502

Note: The data are from the 1980, 1990, 2000 US Census and 2000-2012 ACS (3-year aggregate). The sample is restricted to white females age 20 to 40 who are not currently living in their state of birth. The unit of observation is at the state of birth*state of residence*year level. Sexist beliefs are normalized to have a mean of 0 and standard deviation 1 across the 44 states with available information to construct the index of sexist beliefs in the GSS (excluding DC). In Columns (2), (3), (5), and (6), overall sexist beliefs in the state of residence are instrumented using predicted sexism based on geographic distance and historical migration (i.e. $Z_1(\Delta_{bj})$, and $Z_2(\kappa_{bj})$). "Additional controls" include the mean NAEP-LTT reading and math scores of boys in each state of residence and state of birth as well as the predicted female employment share in the state of residence based on the state's 1950 industry structure. All the regressions also control for region of birth fixed effects and year fixed effects. Regressions are weighted by the number of observations in each cell for each outcome. Standard errors clustered at the state of residence level are reported in parentheses. ***significant at 1%, **at 5%, *at 10%.

to act in the same direction on a woman's marriage and childbearing age; however, the magnitude of residential sexism is typically larger – nearly 1.5 to three times as large as the effects of sexism into which a woman was born as a child. These estimates are largely robust to the inclusion of additional controls for state-level characteristics.[35,36]

A noteworthy feature of the TSLS results for the effects of residential sexism in Table 4 is that the participation point estimates are about three times as large as the corresponding OLS results. Strikingly, there is no similar difference between OLS and TSLS for the non-labor market estimates in Table 5. What type of endogenous sorting would cause OLS estimates of the negative effect of residential sexism on the likelihood of participation, in particular, to be biased towards zero?[37] This type of bias implies that, for some reason, female migrants who are observed living in more sexist states have relatively high levels of latent labor force attachments. Something that might generate this pattern is if women who move to states with higher sexism are systematically more likely to have been made to do so *by their jobs.* Because these women would be disproportionately likely to be working after moving, their presence in the set of all female migrants living in a state imparts an upward bias to estimates of the negative effect residential sexism on labor force participation. By isolating variation in location that is attributable only to exogenous relative distance or historical settlement patterns, the TSLS models hopefully remove this bias.

5 Mechanisms for the Causal Effect of Residential Sexism

We showed that both residential sexism and background sexism appear to have a causal impact of women's relative labor market outcomes and non-labor market outcomes. We argued that the causal effects of background sexism reflect the long-term effect of norms that one was exposed to earlier in life. By contrast, the effect of higher residential sexism could reflect not only the effect of women's adherence to or internalization of norms that surround her where she currently lives,

[35]The estimated effects of residential sexism on female age at first child is no longer statistically significant when additional state-level characteristics are added. Nevertheless, the magnitude of the point estimate remains quite similar (if anything, larger) as controls are added to the baseline TSLS specification.

[36]Columns (5) to (6) of Appendix Table A5 shows that the results are qualitatively similar, albeit slightly smaller in magnitude when we use the alternative specification that includes the full set of state of residence fixed effects to estimate the effects of background sexism on women's non-labor market outcomes.

[37]OLS regressions might also suffer from attenuation bias caused by the fact that the use of the coarse geographic area of the state as the labor market mis-measures the market in which women interact and work. One might have expected any such mis-measurement to also cause substantial attenuation bias in the non-labor market results, yet the OLS and TSLS for these regression are very similar.

but also discrimination imposed upon her in the market. How much, if any, of the effect of higher residential sexism is due to discrimination rather than norms? In this section, we conduct a variety of analyses to shed light on this question.

5.1 Impact of Average Male and Female Sexism

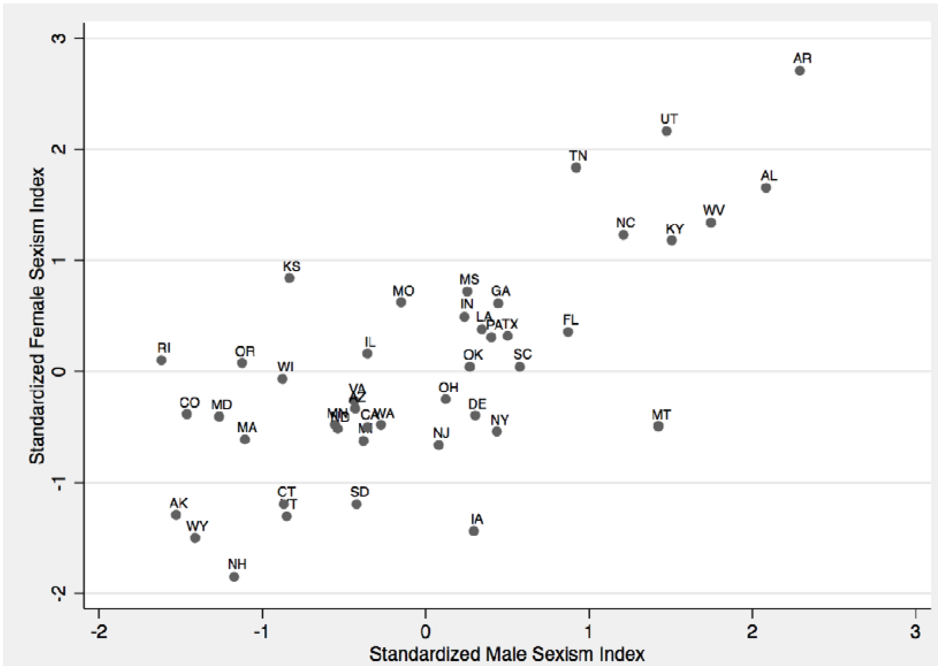
We begin this portion of our analysis with an assessment of the differential effect on the two types of outcomes of average sexism in a market among men and women. Two ideas motivate this exercise. First, the portion of the overall effect of residential sexism attributable to discrimination should be larger for outcomes like wages and labor force participation which labor market discrimination can directly affect, compared to non-labor market outcomes like marriage or childbearing age which are affected only indirectly by market discrimination, if at all. Second, should there be discriminatory action in the labor market, one would expect it to be especially related to sexism among men since they are the prejudicial actors in the labor market.

Using GSS reports of respondent gender, we create separate measures of sexism for men and women in each state. Figure 6 plots the mean of these male and female sexism indices for the different states. The figure shows that although the two measures are very strongly positively correlated, they are far from perfectly so. Men are, on average, more sexist than women in most states, but the extent to which this is true varies substantially from state to state and there are several states for which mean sexism among women is higher than that for men. From the figure, it is clear that there is significant variation in one of the mean sexism measures conditional on information about the other - an important requirement for the analysis we conduct below.

We estimate a series of regressions relating different outcomes in a state simultaneously to average male and average female sexism in the state. This part of our analysis does not require the migrant sample used above, with information about state of birth and state of residence. Therefore, instead of the Census/ACS, we use data from the CPS. As discussed above, the CPS is the preferred source of data on labor market information in the U.S. Data on the two non-labor market outcomes in the table still come from the Census/ACS. All the regressions in the table use residualized outcomes that account for the number of years of schooling and age dummies in the underlying individual-level data used to construct the state-level outcomes.[38]

[38] See footnote 13 for details on the procedure used to create the state-level outcomes.

Figure 6: Sexism Among Men and Among Women in Each State



Note: Data are from the General Social Survey. The standardized female and male sexism indices have mean 0 and standard deviation 1 in the cross-state sample (44 states).

The results in column (1) of Table 6 show that the gender labor force participation gap in a state is significantly larger the more sexist the average man in the state. The coefficient on average female sexism is also negative, albeit not statistically significant. To the extent that female average sexism captures the role of current norms (and other state-level factors that might be correlated with male sexist beliefs), we would expect to observe a negative relationship between female beliefs and the female-male gap in labor market outcomes. What matters for our analysis is whether LFP gaps are at least as strongly related to male sexism as female sexism. Although we cannot reject that the coefficient on average male sexism is equal to (or more positive) than the coefficient on average female sexism, the point estimate on male sexism is larger than on female sexism.

Mirroring the wage results found previously among migrants for overall residential sexism with the Census wage data, the results in column (2) of the table show that the wage gap, conditional on employment, in a state measured using CPS data does not vary at all with mean sexism among either men or women in the state.

In column (3) we study the gender difference not in the wages received by workers but rather

Table 6: Mean Male vs. Mean Female Sexism and Women's Outcomes

	LFP Gap (Female - Male)	Log Wage Gaps, conditional on working (Female-Male)	Selection-Corrected Log Wage Gaps (Female-Male)	Share of Females Nevermarried	Average Female Age at First Child
	(1)	(2)	(3)	(4)	(5)
Sexist Beliefs in State of Residence					
Male average	-0.011*	-0.004	-0.030***	-0.010	-0.093
	[0.006]	[0.006]	[0.010]	[0.011]	[0.091]
Female average	-0.007	-0.001	-0.004	-0.032**	-0.204*
	[0.006]	[0.006]	[0.010]	[0.012]	[0.107]
Sample	CPS 1977-2017, Age 25 to 64			Census/ACS 1980-2012, Age 20 to 40	
Individual level controls in first-stage regression:					
Education, Age dummies	X	X	X	X	X
Statistical test (Ha):	Male average < Female average			Female average < Male average	
p-value (one-tailed)	0.36	0.35	0.05	0.17	0.28
Observations	44	44	44	44	44
R-squared	0.298	0.017	0.254	0.456	0.345

Note: The table reports coefficients (standard errors) from OLS regressions of residual state-level outcomes on various measures of sexist beliefs. The residual female-male employment and wage gaps are estimated using the sample of whites age 25 to 64 from the 1977-2017 May/ORG CPS data and control for the number of years of schooling, age dummies, gender-specific year effects and state fixed effects. The non-labor market outcomes are estimated using the sample of white females age 20 to 40 from the 1980 to 2000 Census and 2010-2012 ACS (3-year aggregate) and control for the number of years of schooling, age dummies and year fixed effects. Sexist beliefs are normalized to have a mean of 0 and standard deviation 1 across the 44 states with available information to construct the index of sexist beliefs in the GSS. Regressions are weighted by the inverse of the variance of the outcome variable. Robust standard errors are reported in parentheses. ***significant at 1%, **5%, *10%.

in the wages that men and women could *command* in the labor market. Our examination of the "offer" (or "potential") wage gap follows a long tradition that attempts to account for missing wage data among those who do not work - a particularly important problem among women. The offer wage is what a person would be paid if they *did* work, but people only work if their offer wage is such that their utility from work exceeds their reservation utility from non-work. Presumably, discriminatory treatment against women lowers their offer wages compared to those of men, which also lowers their relative likelihood of working.

We measure the gender difference in offer wages in a state as the median difference in selection-corrected wages between men and women. Our selection correction is based on a simple imputation procedure whereby non-fulltime employed women and men in the top quartile of the gender-specific education distribution are assigned wages above the median and women and men in the bottom quartile of the gender-specific education distribution are assigned wages below the median. Our procedure follows the approach of Brown (1984), Neal (2004) and Olivetti and Petrongolo (2008).[39] Using these selection-corrected wages, we estimate median regressions, controlling for education, age fixed effects, gender-specific year effects, state effects and state×female effects. The coefficients on the state×gender dummies measure the median offer wage gaps in a state.

The results in column (3) of Table 6 show that when related simultaneously to mean sexism among men and women living there, the selection-corrected wage gap in a state loads exclusively on the mean male sexism and exhibits no statistically significant relationship with average residential female sexism. This difference is not a matter of statistical imprecision; the estimated effect for male sexism is more than five times as large as that for residential female sexism. Here, the one-tailed test indicates that the effect of male average sexism on female-male selection-corrected wage gaps is significantly more negative than the effect of female average sexism at the 5 percent level.

Columns (4) and (5) show results for the two non-labor market outcomes. In a striking reversal of the participation and wage results, we find that both the likelihood of ever having married and the age of first child bearing are associated exclusively with mean sexism among women in the market, but not at all with mean sexism among men. Again, these effects are not the result of one parameter being estimated with particularly large standard errors. Rather, the two point estimates are of very different sizes, with the estimated effect for mean female sexism being roughly three times as big as the insignificant associations for male sexism.[40]

[39]The advantage of this imputation approach is that it only requires us to make assumptions on the position of the imputed wage observations relative to the median (i.e. whether the unobserved offer wage is above or below the median). As noted in Olivetti and Petrongolo (2008), this procedure does not require assumptions on the actual level of missing wages, nor does it require us to identify an instrument that satisfies the exclusion restriction as required in the valid implementation of the Heckman two-stage sample selection correction model. Note also that this selection correction procedure does not impute wages for women with missing wages in the middle two quartiles on the education distribution under the assumption that one cannot, with reasonable confidence, predict whether the wages these women could command in the market would be above or below the median. Our results are robust to alternative selection-correction imputations such as imputing wages for women and men in the top and bottom ten percent of the education distribution, as well as imputing wages for non-employed women and men. These results available upon request.

[40]Nevertheless, we do not have sufficient power to reject that the coefficient on the female average is larger than or equal to the coefficient on the male average at conventional levels of significance.

5.2 Percentile Tests

The results in Table 6 are consistent with the interpretation that the effect of residential sexism on labor market outcomes is chiefly due to discrimination from sexist men, while residential sexism's effect on non-labor outcomes reflects the influence of internalized sexist norms from other women. We use perhaps the main implication of Becker (1957) to further probe the validity of this conclusion.

Becker's model of prejudice- or taste-based discrimination implies that if there were gender discrimination in the labor market based on men's sexist beliefs, market forces would sort women away from interactions with the most sexist men. Women's equilibrium outcomes would be determined by the sexism of the most sexist men with whom they were forced to interact after this sorting – the "marginal discriminator." Since women constitute roughly half of the population, this cost-minimizing sorting should result in the marginal discriminator having sexism somewhere close to the middle of the male sexism distribution and *not* far away from the median, such as around the 10th or 90th percentiles. Further, if gender discrimination is due to men's sexism, labor market gaps should not be systematically related to *any* moment of the female residential sexism distribution.

An important critique of the Becker taste-discrimination model is that in the long-run, competitive pressures should drive out prejudiced employers. Charles and Guryan (2007) show that the relationship between employer distaste and wage gaps can survive perfect competition, as long as prejudicial tastes are portable across labor market roles (i.e. prejudiced employers continue to remain prejudiced when they shut down a firm and become workers). This formulation of the taste-discrimination model combines the employer and employee discrimination models and implies that since market participants endogenously choose to become "employers" or "workers," what should matter in determining women's relative labor market outcomes is the distribution of prejudice among *all* men in a given market rather than just male employers.

We test these implications empirically by estimating a series of regressions relating labor market outcomes in a state to the 10th, median and 90th percentiles of the distributions of male and female sexism in the state. Table 7 shows the results. Column (1) shows that when regressed on all three percentiles of male sexism simultaneously, the labor force participation gap is strongly statistically

Table 7: Relationship Between Alternative Percentiles of Sexist Beliefs and Women's Labor Market Outcomes

	Labor Force Participation Gap (Female - Male)			Selection-Corrected Log Wage Gaps (Female-Male)		
	(1)	(2)	(3)	(4)	(5)	(6)
Male Sexist Beliefs in State of Residence:						
10th Percentile	-0.000		0.001	-0.002		0.000
	[0.008]		[0.007]	[0.009]		[0.011]
Median	-0.017**		-0.014*	-0.030**		-0.028**
	[0.007]		[0.007]	[0.013]		[0.014]
90th Percentile	0.001		0.001	-0.004		-0.005
	[0.006]		[0.007]	[0.011]		[0.011]
Female Sexist Beliefs in State of Residence:						
10th Percentile		-0.005	-0.003		-0.010	-0.007
		[0.008]	[0.011]		[0.013]	[0.014]
Median		-0.011	-0.007		-0.007	0.005
		[0.009]	[0.013]		[0.018]	[0.017]
90th Percentile		-0.000	0.004		-0.013	-0.002
		[0.005]	[0.006]		[0.010]	[0.011]
p-value: Ha: Male p50 < p10	0.11		0.12	0.07		0.08
p-value: Ha: Male p50 < p90	0.06		0.14	0.10		0.16
Observations	44	44	44	44	44	44
R-squared	0.275	0.243	0.330	0.267	0.143	0.270

Note: The residual female-male employment and wage gaps are estimated using the sample of whites age 25 to 64 from the 1977-2017 May/ORG CPS data and control for the number of years of schooling, age dummies, gender-specific year effects and state fixed effects. Sexist beliefs are normalized to have a mean of 0 and standard deviation 1 across the 44 states with available information to construct the index of sexist beliefs in the GSS (excluding DC). Regressions are weighted by the inverse of the variance of the outcome variable. Robust standard errors are reported in parentheses. ***significant at 1%, **5%, *10%.

Table 8: Relationship Between Alternative Percentiles of Sexist Beliefs and Women's Non-Labor Market Outcomes

	Share of Females Nevermarried (Age 20 to 40)			Average Female Age at First Birth (Age 20 to 40)		
	(1)	(2)	(3)	(4)	(5)	(6)
Male Sexist Beliefs in State of Residence:						
10th Percentile	-0.016		0.002	-0.088		0.037
	[0.010]		[0.012]	[0.084]		[0.088]
Median	-0.009		0.008	-0.128		-0.007
	[0.014]		[0.013]	[0.097]		[0.101]
90th Percentile	-0.016		-0.024	-0.073		-0.133
	[0.014]		[0.015]	[0.098]		[0.107]
Female Sexist Beliefs in State of Residence:						
10th Percentile		-0.015	-0.027**		-0.144	-0.199*
		[0.012]	[0.012]		[0.107]	[0.115]
Median		-0.011	0.005		-0.011	0.084
		[0.013]	[0.016]		[0.120]	[0.125]
90th Percentile		-0.021**	-0.021**		-0.182**	-0.181**
		[0.009]	[0.010]		[0.072]	[0.082]
R-squared	44	44	44	44	44	44
Observations	0.356	0.449	0.516	0.284	0.358	0.407

Note: The non-labor market outcomes are estimated using the sample of white females age 20 to 40 from the 1980 to 2000 Census and 2010-2012 ACS (3-year aggregate) and control for the number of years of schooling, age dummies and year fixed effects. Sexist beliefs are normalized to have a mean of 0 and standard deviation 1 across the 44 states with available information to construct the index of sexist beliefs in the GSS (excluding DC). Regressions are weighted by the inverse of the variance of the outcome variable. Robust standard errors are reported in parentheses. ***significant at 1%, **5%, *10%.

related to the median, but not the 10th or 90th percentiles. [41]

By contrast, column (2) shows that when regressed only on percentiles of female residential sexism, a state's participation gap does not vary in a statistically significant way with *any* part of the female sexism distribution in the state, although the point estimate for the median is bigger than those for the other two percentiles. The regression in column (3) includes the three percentiles for both female and male sexism. In this full regression, the participation gap is related only with median male sexism - the point estimate for which is twice as large as that for the statistically insignificant estimate for the median female.

The next three columns in the table show how the offer wage gap in a state is related to the different percentiles of male and female sexism. These results mirror the participation findings and are, in fact, slightly stronger. The selection-corrected wage gap in a state does not systematically vary with any percentile of female sexism. However, the wage gap is related in a strong and highly statistically significant way only with the median of the male residential sexism. The point estimate for the male sexism is more than ten times larger than the estimated association for any other percentile point in the table.[42]

It is worth noting that the fact that including the female quantiles has little effect on the pattern or significance of the results for male sexism in the participation or selection-corrected wage regressions suggests that the patterns we observe are unlikely to be driven by some general community sentiment that is distinct from sexism or some unobserved state level characteristic that is correlated with male sexism and has an independent effect on women's relative labor market outcomes. If male responses merely capture such general sentiments or unobserved heterogeneity across states, we would expect these factors to similarly load onto female responses to the questions in the various regressions. The fact that we find strong effects for median male sexism, even after controlling for women's own views about the tasks they should engage in, suggests that the male sexism measures capture something close to the discriminatory tastes that we have described. As further evidence that the observed patterns are not entirely driven by omitted state-level characteristics, in Appendix Table A6, we show that the results remain robust to the inclusion of controls

[41]The formal test of whether the coefficient on male median sexism is more negative than the coefficient on male sexism at the 10th or 90th percentile yields p-values of 0.11 and 0.06, respectively.

[42]The formal test of whether the coefficient on male median sexism is more negative than the coefficient on male sexism at the 10th or 90th percentile yields p-values of 0.07 and 0.10, respectively.

for cross-state differences in women's employment opportunities as predicted by historical industry structure as well as proxies for school quality (NAEP-LTT reading and math test scores for boys).

While taste-based discrimination models make sharp predictions about the relative importance of different quantiles of the relevant sexism for labor market outcomes, there is no such prediction from accounts describing the operation of norms. If they are strongly affected by norms, non-labor market outcomes could thus vary systematically with any part or parts of female sexism distribution and be consistent with norms influencing outcomes.

Table 8 shows the results of the same percentile exercise for the two non-labor market outcomes that we do for wage and employment gaps in Table 7. The difference between the two sets of outcomes is striking. All of the specifications in Table 8 find that neither the female marriage rate in the state nor the average age at which they have their first child is related in a statistically significant fashion with any percentile of male sexism. Both outcomes are, however, strongly related to female sexism - although in neither case is the association with the median. Instead, as best illustrated in the specifications in columns (3) and (6), which include all the percentiles of both male and female sexism, the two outcomes are significantly related to the top (90th percentile) and bottom (10th percentile) of the distribution of female sexism but not with the median.

That women's non-labor market outcomes are not related to median female residential sexism but are related to the tails of that distribution is consistent with norms accounting for sexism's effect on these outcomes. As noted, existing accounts about social norms make no sharp theoretical prediction about which percentiles of underlying beliefs are particularly important for the operation of these effects. Moreover, the results in Table 8 highlight that state-level outcomes do not always, for some mechanical reason, load most strongly onto median sexism in a horse-race with other percentiles. This raises confidence that the finding in Table 7 concerning the male median and labor market outcomes truly captures the effect of marginal discriminator as predicted by theory.

Finally, we assess the extent to which we are able to separately identify the impact of sexism from cross-state differences in religiosity.[43] Moreover, we can use cross-state differences in self-reported religiosity as a "placebo test" for the percentile analyses to ascertain whether the importance of the median male for labor market outcomes is unique to proxies for discriminatory tastes (i.e. sexism). We measure religiosity using the following question in the General Social Survey: "How

[43]The cross-state rank correlation between male (female) sexism and religiosity is 0.53 (0.60), $p = 0.000$.

often do you attend religious services?" Individuals respond to the question using a 0 (never) to 8 (more than once a week) scale. As shown in Appendix Table A7, our main findings remain virtually unchanged with the inclusion of the religiosity control. The results of the placebo test, presented in Appendix Table A8 also reveals no particular relationship between the median of the male religiosity distribution and women's relative labor market outcomes, in stark contrast to the results for male sexism.

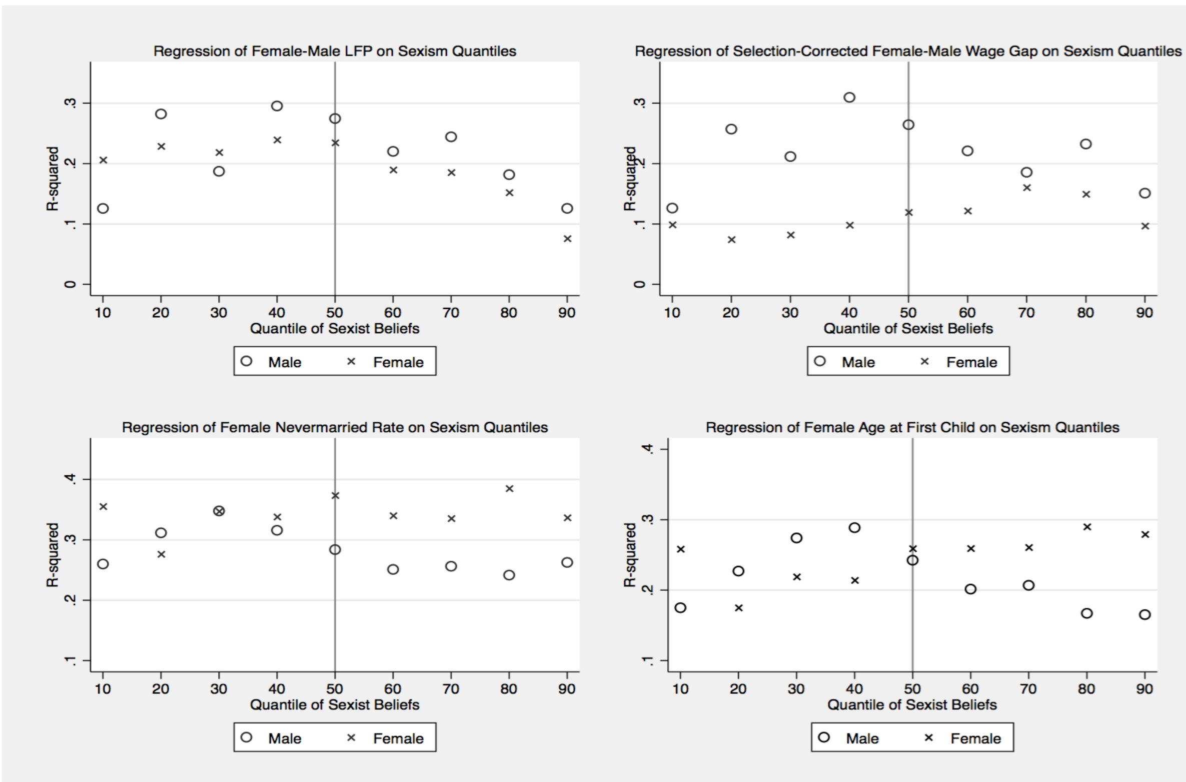
5.3 Robustness

Before concluding this section, we present some results intended to assess the sensitivity of the results in Tables 7 and Table 8 to two features of our estimation.

The first of these features is the set of percentiles we use to measure different points in the distributions of sexism. The tests we conduct are based on the predictions of a Becker model of discrimination that, since women represent roughly half the labor market, the sorting mechanism that determines equilibrium outcomes in the presence of taste-based discrimination should cause the marginal discriminator to be drawn from center of the sexism distribution rather than the tails. There being no rule about precisely which percentiles represent the "middle" versus the "tails" of the distributions, our main results settle upon the 10th, 90th and the median as natural candidates. Plus, these are the percentile points examined by Charles and Guryan (2008) in their work showing that black-white wage gaps are related, as a taste-based model would predict given blacks' small share of the overall population, *only* with the left-tail of the distribution of racial animus.[44] Yet, there are two reasons to suppose that the marginal discriminator should have sexism slightly less than the 50th percentile. For one thing, whereas women now constitute close to half of the labor force, over the entire period we study, the labor force was roughly 40% female. Further, Becker's model argues that because of the cost savings that non-discriminating firms enjoy, they should grow slightly larger than their discriminating counterparts, with the result that they should account for a disproportionate share of hiring of persons from the disadvantaged group (women in our case).

[44] Using measures of racial prejudice from the GSS and computed in a very similar fashion to our sexism measures, Charles and Guryan (2008) find that black-white wage gaps are related to the 10th percentile and not to the median and 90th percentile of white prejudice precisely as the prejudice model predicts for racial minorities. That the data appear to be consistent with the very different predictions about which quantile points in the relevant distribution are marginal and infra-marginal in the case of race and gender suggests that prejudice – both racial and gender – are important determinants of relative wages.

Figure 7: R-Squared Statistics from Bivariate Regressions of Each Outcome on Separate Quantiles of Male and Female Residential Sexism



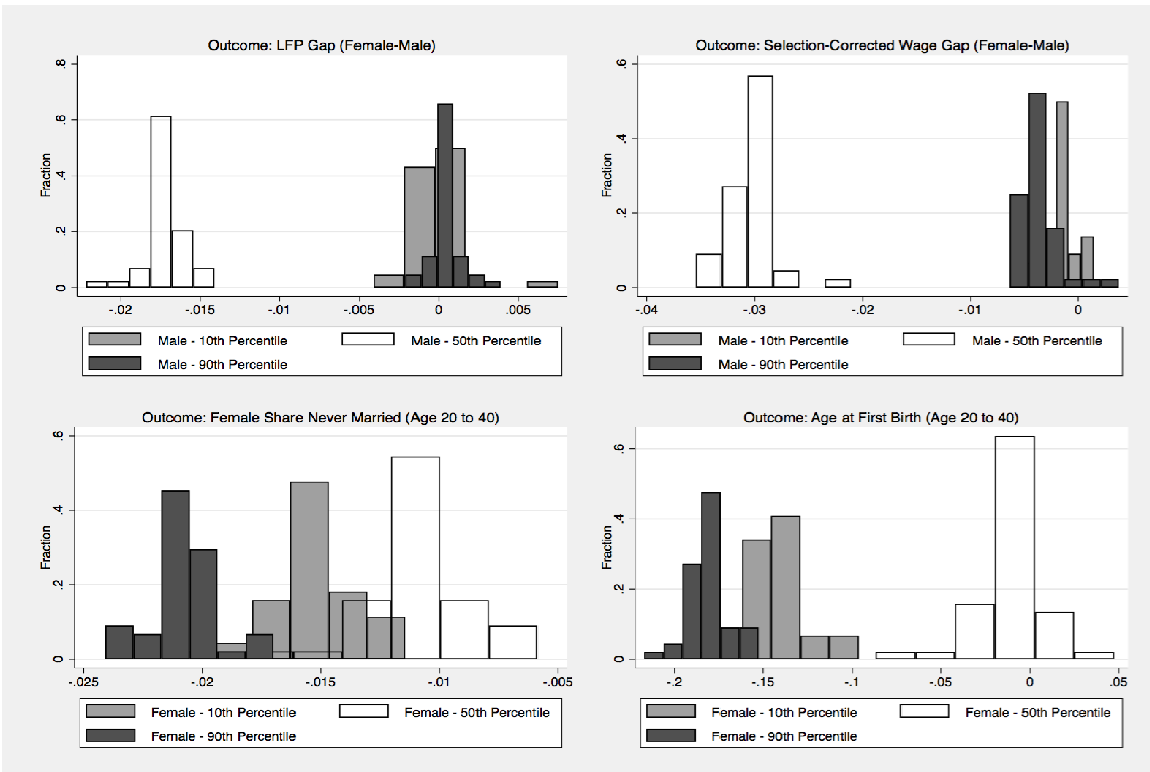
Note: Each circle or cross reports the R-squared from a bivariate regression of the outcome indicated in the title of each graph on the 10th, 20th, ..., 90th percentile of the male (circle) or female (cross) residential sexism distribution within each state.

To examine whether our results hinge in an important way on the percentile points we have chosen to examine, we conduct a robustness exercise in which we estimate a series of simple bivariate regressions relating each outcomes to different percentiles of male and female sexism, ranging from the 10th to the 90th percentile. Figure 7 plots the R-squares from each of these eighteen pairwise associations. The figure shows the point in male and female residential sexism distributions where the association with the particular outcome is strongest. Two things stand out clearly in the figure. One is that for the labor market outcomes the pairwise associations for males yield the highest R-squared, whereas for the non-labor market outcomes associations, R-squared statistics are highest with percentiles of female sexism. The other noteworthy pattern in the figure is which percentiles show the biggest R-squared. For labor market outcomes, the association is highest somewhere between the 40th and median of male sexism. In sharp contrast, for neither non-labor market outcomes is the association maximized for median female sexism. It is worth emphasizing that for these results no special restrictions or specifications have been imposed on the data. We therefore find it reassuring that these associational findings are so strongly consistent with our main results.

The other robustness exercise we conduct is motivated by the fact that state-level variation plays an important role in our analysis. We are therefore naturally concerned that a small set of states - or even one particular state - might drive our main results. To assess how important an issue this is, we estimate the main regressions in Tables 7 and 8 many times, each time dropping observations from a particular state from the estimation sample. We then plot histograms of point estimates forthcoming from this series of "leave-out" regressions. Since there are 44 states in our sample, we run 43 regressions for each outcome. Having already shown the men's sexism drives labor market outcomes, the baseline regressions for labor market results only includes the three percentiles of male sexism. For non-labor market outcomes, the baseline regression is only estimated on the three different percentiles of female sexism.

Figure 8 presents the histograms of point estimates across the various regressions. The three colors in the figure represent point estimates for the three different percentiles from the series of regressions. The figure shows that for the labor market outcomes, the point estimates for median male sexism from the series of regressions are clustered around a negative value. The point estimates for the other two percentile of male sexism all cluster around 0 across all of the regressions. The results for the non-labor market outcomes are very different. For these outcomes, the histograms

Figure 8: Sensitivity of Regression Estimates to Omitting Observations from Different Individual States: Coefficient on Male and Female Sexism from 43 Regressions



Note: The residual female-male employment and wage gaps are estimated using the sample of whites age 25 to 64 from the 1977-2017 May/ORG CPS data and control for the number of years of schooling, age dummies, gender-specific year effects and state fixed effects. The non-labor market outcomes are estimated using the sample of white females age 20 to 40 from the 1980 to 2000 Census and 2010-2012 ACS (3-year aggregate) and control for the number of years of schooling, age dummies and year fixed effects. Data on sexism are from the General Social Survey. The standardized female and male sexism indices have mean 0 and standard deviation 1 in the cross-state sample (44 states). The histograms show the distribution of coefficient estimates from 43 separate (leave one state out) regressions of columns (1) and (4) from Table 7 and columns (2) and (5) from Table 8 for the labor market and non-labor market outcomes, respectively.

show that point estimates for median female sexism are clustered around 0, while those for the other two percentiles both cluster around negative values. These results provide strong graphical evidence that the main results we show above are not driven by any particular state, but instead represent something fundamental about the economy overall.

6 Conclusion

This paper studies how prevailing sexist beliefs about women's appropriate roles and abilities affect American women's socioeconomic outcomes. Studying adults who live in one state but who were born in another, and using an instrumental variables approach, we show that both sexism in a woman's current state of residence and her state of birth adversely affect her labor market outcomes relative to men, increase her likelihood of marriage, and to have children at an earlier age. We argue that sexism where a woman was born, which we call background sexism, affects her outcomes even though she is currently living in another market, through the influence of norms and beliefs that she internalized during her formative years.

We present various pieces of evidence which, collectively, suggest the sexism where a woman lives (residential sexism) affects her non-labor market outcomes through the influence of prevailing sexist beliefs of other women where she lives. By contrast, residential sexism's effects on her labor market outcomes seem to operate chiefly through the mechanism of market discrimination by sexist men. These results extend the massive existing literature on discrimination. "Taste-based" models – one of the two main classes of theoretical accounts of market discrimination along with statistical discrimination – posit that discriminatory actions in the market are undertaken by persons holding negative or aversive sentiments towards the disadvantaged group in question. Our finding that quantiles of male (but not female) sexism in a market are associated with worse labor market outcomes for women in the very specific way that a model of gender discrimination arising from underlying sexist sentiments would predict, provides one of the few pieces of direct evidence that prejudice-based discrimination, undergirded by prevailing sexist beliefs may be an important driver of women's outcomes in the U.S.

References

Acemoglu, Daron and Matthew O. Jackson, "Social Norms and the Enforcement of Laws," *Journal of the European Economic Association*, 2017, *15* (2), 245–295.

Akerlof, George A. and Rachel E. Kranton, "Economics and Identity," *The Quarterly Journal of Economics*, August 2000, *115* (3), 715–753.

Alesina, Alberto, Paola Giuliano, and Nathan Nunn, "On the Origins of Gender Roles: Women and the Plough," *The Quarterly Journal of Economics*, May 2013, *128* (2), 469–527.

Arrow, Kenneth, "The Theory of Discrimination," in Orley Ashenfelter and Albert Rees, eds., *Discrimination in Labor Markets*, Princeton, NJ: Princeton University Press, 1973, pp. 3–33.

Autor, David H., Lawrence F. Katz, and Melissa S. Kearney, "Trends in U.S. Wage Inequality: Revising the Revisionists," *Review of Economics and Statistics*, 2008, *90*, 300–323.

Bailey, Martha, Melanie Guldi, and Brad J. Hershbein, "Is There a Case for a 'Second Demographic Transition?' Three Distinctive Features of the post-1960 U.S. Fertility Decline," in Carola Frydman Leah P. Boustan and Robert A. Margo, eds., *Human Capital in History: The American Record*, Chicago, IL: The University of Chicago Press, 2014, pp. 273–312.

Beaudry, Paul and Ethan Lewis, "Do Male-Female Wage Differentials Reflect Differences in Return to Skill? Cross-City Evidence from 1980-2000," *American Economic Journal: Applied Economics*, 2014, *6* (2), 178–194.

Becker, Gary, *The Economics of Discrimination*, The University of Chicago Press, 1957.

Benabou, Roland and Jean Tirole, "Incentives and Prosocial Behavior," *American Economic Review*, August 2006, *96* (5), 1652–1678.

Bertrand, Marianne, Emir Kamenica, and Jessica Pan, "Gender Identity and Relative Income Within Households," *Quarterly Journal of Economics*, 2015, *130* (2), 571–614.

Black, Dan A., Natalia Kolesnikova, and Lowell J. Taylor, "Why Do So Few Women Work in New York and So Many in Minneapolis? Labor Supply of Married women Across U.S. Cities," *Journal of Labor Economics*, 2014, *79.*

Black, Sandy E. and Alexandra Spitz-Oener, "Technological Change and the Skill Content of Women's Work," *Review of Economics and Statistics*, 2010, *92* (1), 187–194.

Blau, Francine D. and Lawrence Kahn, "The U.S. Gender Pay Gap in the 1990s: Slowing Convergence," *Industrial Relations and Labor Relations Review*, 2006, *60* (1), 45–66.

_ and _ , "The Gender Wage Gap: Extent, Trends, and Explanations," *Journal of Economic Literature*, 2017, *55* (3), 789–865.

_ , Lawrence M. Kahn, and Kerry L. Papps, "Gender, Source Country Characteristics, and Labor Market Assimilation Among Immigrants," *Review of Economics and Statistics*, 2011, *93* (1), 43–58.

Boustan, Leah P., "Was Postwar Suburbanization White Flight? Evidence from the Black Migration," *Quarterly Journal of Economics*, 2010, *125* (1), 417–443.

—, Price V. Fishback, and Shawn Kantor, "The Effect of Internal Migration on Local Labor Markets: American Cities During the Great Depression," *Journal of Labor Economics*, 2010, *28* (4), 719–746.

Brown, Charles, "Black-White Earnings Ratios since the Civil Rights Act of 1964: The Importance of Labor Market Dropouts," *Quarterly Journal of Economics*, February 1984, *99* (1), 33–44.

Bursztyn, Leonardo, Thomas Fujiwara, and Amanda Pallais, "Acting Wife: Marriage Market Incentives and Labor Market Investments," *American Economic Review*, Nov 2017, *107 (11)*, 3288–3319.

Burt, Keith B. and Jacqueline Scott, "Parent and Adolescent Gender Role Attitudes in 1990s in Great Britain," *Sex Roles*, 2002, *46*, 239–45.

Card, David, "Immigration Inflows, Native Outflows, and the Local Labor Market Impacts of Higher immigration," *Journal of Labor Economics*, January 2001, *19* (1), 22–64.

Charles, Kerwin Kofi and Jonathan Guryan, "Prejudice and the Economics of Discrimination," *NBER Working Paper 13661*, 2007.

— and —, "Prejudice and Wages: An Empirical Assessment of Becker's The Economics of Discrimination," *Journal of Political Economy*, 2008, *116* (5), 773–809.

— and Ming-Ching Luoh, "Gender Differences in Completed Schooling," *Review of Economics and Statistics*, 2003, *85* (3), 559–577.

Cortes, Patricia, "The Effect of Low-Skilled Immigration on United States Prices: Evidence from CPI Data," *Journal of Political Economy*, 2008, *116* (3), 381–422.

— and Jose Tesseda, "Low-Skilled Immigration and the Labor Supply of Highly Skilled Women," *American Economic Journal: Applied Economics*, 2011, *3* (3), 88–123.

Fernandez, Raquel, Alessandra Fogli, and Claudia Olivetti, "Mother and Sons: Preference Development and Female Labour Force Dynamics," *Quarterly Journal of Economics*, 2004, *119*, 1249–1299.

— and —, "Culture: An Empirical Investigation of Beliefs, Work, and Fertility," *American Economic Journal: Macroeconomics*, 2009, *1* (1), 146–177.

Flabbi, Luca, "Prejudice and Gender Differentials in the U.S. Labor Market in the Last Twenty Years," *Journal of Econometrics*, 2010, *156*, 190–200.

Fortin, Nicole, "Gender Role Attitudes and Women's Labour Market Outcomes Across OECD Countries," *Oxford Review of Economic Policy*, 2005, *21*, 416–438.

—, "Gender Role Attitudes and Women's Labor Market Participation: Opting-Out, AIDS, and the Persistent Appeal of Housewifery," Technical Report 2010.

Goldin, Claudia and Lawrence F. Katz, "The Power of the Pill: Oral Contraceptives and Women's Career and Marriage Decisions," *Journal of Political Economy*, 2002, *110*, 730–770.

—, —, and Ilyana Kuziemko, "The Homecoming of American College Women: The Reversal of the College Gender Gap," *The Journal of Economic Perspectives*, 2006, *20* (4), 133.

Greenwood, Jeremy, Ananth Sheshadri, and Mehmet Yorukoglu, "Engines of Liberation," *The Review of Economic Studies*, 2005, *72*, 109–133.

Hsieh, Chang-Tai., Erik Hurst, Charles I. Jones, and Peter J. Klenow, "The Allocation of Talent and U.S. Economic Growth," Working Paper 18693, National Bureau of Economic Research January 2013.

Isen, Adam and Betsey Stevenson, "Women's Education and Family Behavior: Trends in Marriage, Divorce and Fertility," in "Topics in Demography and the Economy National Bureau of Economic Research, Inc.," National Bureau of Economics Research, 2008.

Kiecolt, K. Jill and Alan C. Acock, "The Long-Term Effects of Family Structure on Gender-Role Attitudes," *Journal of Marriage and the Family*, 1988, *50*, 709–17.

Munshi, Kaivan, "Networks in the Modern Economy: Mexican Migrants in the U.S. Labor Market," *Quarterly Journal of Economics*, 2003, *118* (2), 549–599.

Neal, Derek, "The Measured Black-White Wage Gap Among Women is Too Small," *Journal of Political Economy*, 2004, *112* (2), S1–S28.

Olivetti, Claudia and Barbara Petrongolo, "Unequal Pay or Unequal Employment? A Cross-Country Analysis of Gender Gaps," *Journal of Labor Economics*, oct 2008, *Vol. 26, No. 4*, 621–654.

_ , Eleanora Patacchini, and Yves Zenou, "Mothers, Peers, and Gender Identity," *Journal of the European Economic Association*, forthcoming.

Ortega, Francesc and Giovanni Peri, "Openness and Income: The Role of Trade and Migration," *Journal of International Economics*, 2014, *92* (2), 231–251.

Smith, Adam, *The Theory Of Moral Sentiments*, A. Millar, 1759.

Weinburg, Bruce, "Computer Use and the Demand for Female Workers," *Industrial and Labor Relations Review*, 2000, *53* (1), 290–308.

Data Appendix

A. Construction of Hourly Wages using the CPS Data

We use data from the May CPS for 1977 to 1978 and the CPS Merged Outgoing Rotation Groups (MORG) for 1979 to 2017. The main analysis sample includes whites between the ages of 25 to 64. Following the procedure used by Autor, Katz, and Kearney (2008), hourly wages are defined as the log of reported hourly earnings for those paid by the hour and the log of usual weekly earnings divided by hours worked last week for non-hourly workers. The wage sample excludes those who are self-employed. Top-coded earnings are multiplied by 1.5. Hourly earnings that fall below $1.675/hour in 1982 dollars are dropped, as are hourly wages exceeding 1/35th the top-coded value of weekly earnings. All the earnings measures are deflated to 2009$ using the price deflator for personal consumption expenditures (PCE) from the BEA. We exclude allocated earnings observation in all years, except where allocation flags are not available (Jan 1994 to Aug 1995). Between 1989 and 1993, nonflagged allocated observations are identified and dropped by using the unedited earnings values provided in the source data.

B. Construction of Hourly Wages using the Census/ACS

We use the Census IPUMS 5% data for the 1980, 1990, and 2000 Census and the 2012 (2010-2012) three-year aggregate data from the ACS. The sample includes whites between the ages of 25 to 64 who are not living in group quarters (i.e. non-military and non-institutionalized). Hourly wages are calculated as total annual wage and salary income divided by the product of weeks worked and usual hours worked in the previous year. The wage sample excludes the self-employed. Following Autor, Katz, and Kearney (2008), we drop the bottom 1% of hourly earners and multiply hourly wages of top-coded earners by 1.5. The maximum hourly wage is limited to 1.5 times the maximum annual income amount divided by 1,750 (35 hours per week for 50 hours per year). The earnings measures are deflated to 2000$ using the price deflator for personal consumption expenditures (PCE) from the BEA.

C. Coding of Education in the CPS and Census/ACS Samples

We follow steps outlined in Autor, Katz, and Kearney (2008) to create a comparable measure of the number of years of schooling across years in the CPS. In 1992, the CPS changed the education question. Figures from Park (1994) are used to assign years of completed education to each worked based on gender, race, and highest degree held. For the Census and ACS samples, we impute the number of years of education based on the variable indicating the individual's highest year of school or degree completed (*educd*).

Table A1: Convergence in Mean State Level Labor and Non-Labor Market Outcomes Over Time

	Unconditional		*Residualized*	
	State FE only	State FE and Year FE	State FE only	State FE and Year FE
	(1)	(2)	(3)	(4)
	Female-Male LFP Gap			
R-squared	0.185	0.952	0.194	0.944
	Female-Male Wage Gap, Conditional on Working			
R-squared	0.129	0.949	0.166	0.929
No. of Obs (44 states by 8 time periods)	352	352	352	352
	Female Share Nevermarried (Age 20 to 40)			
R-squared	0.287	0.985	0.067	0.997
	Female Age at First Birth (Age 20 to 40)			
R-squared	0.410	0.959	0.248	0.976
No. of Obs (44 states by 4 time periods)	176	176	176	176

Note: Regressions use data from the 1977-2017 May/ORG Current Population Surveys (CPS) for labor market outcomes and from the 1980 to 2012 Census/ACS for non-labor market outcomes. The table reports the R-squares from separate regressions of state-level labor market and non-labor market outcomes on state fixed effects and year fixed effects. For the labor market outcomes, we group the CPS years into 8 time periods; for the non-labor market outcomes, the Census/ACS years are grouped into 4 time periods.

Table A2: GSS Questions Used to to Construct Sexism Index

FEWORK	Do you approve or disapprove of a married woman earning money in business or industry if she has a husband capable of supporting her?
FEHOME	Do you agree or disagree with this statement? Women should take care of running their home and leave running the country up to men.
FEPRES	If your party nominated a woman for president, would you vote for her if she were qualified for the job?
FEPOL	Tell me if you agree or disagree with this statement: Most men are better suited emotionally for politics than are most women.
FECHILD	A working mother can establish just as warm and secure a relationship with her children as a mother who does not work.
FEPRESCH	A preschool child is likely to suffer if his or her mother works.
FEHELP	It is more important for a wife to help her husband's career than to have one herself.
FEFAM	It is much better for everyone involved if the man is the achiever outside the home and the women takes care of the home and family.
	Years where all questions overlap: (1977, 1985, 1986, 1988, 1989, 1990, 1991, 1993, 1994, 1996, 1998)

Table A3: Convergence Over Time in Average State-Level Sexism

	State FE only (1)	State FE and Year FE (2)
	Overall Sexism	
R-squared	0.498	0.869
	Average Male Sexism	
R-squared	0.535	0.794
	Average Female Sexism	
R-squared	0.473	0.870
No. of Obs (44 states by 3 time periods)	107	107
Includes:		
State Fixed Effects	yes	yes
Year Fixed Effects	no	yes

Note: The sexism data is from the 1977 to 1998 waves of the GSS. The table reports R-squares from separate regressions of average state-level sexism on state fixed effects and year fixed effects. The three time periods are (1) 1977 to 1988 (2) 1989 to 1993 (3) 1994 to 1998. The number of observations do not add up to 44*3 = 132 due to missing observations in some states in some time periods. We restrict the analysis to state*time period cells with at least 15 male and female respondents.

Table A4: Sexism Measures, by State

	All		Men		Women	
State	N	Index	N	Index	N	Index
Alabama	331	1.97	131	2.10	200	1.65
Alaska	44	-1.50	22	-1.53	22	-1.31
Arizona	264	-0.41	121	-0.43	143	-0.35
Arkansas	110	2.78	51	2.30	59	2.70
California	1569	-0.46	719	-0.35	850	-0.51
Colorado	484	-0.95	199	-1.46	285	-0.39
Connecticut	228	-1.13	105	-0.86	123	-1.21
Delaware	53	-0.30	18	0.31	35	-0.40
Florida	703	0.61	304	0.87	399	0.34
Georgia	352	0.62	162	0.44	190	0.61
Illinois	615	-0.06	286	-0.35	329	0.15
Indiana	421	0.41	177	0.24	244	0.49
Iowa	79	-0.69	37	0.29	42	-1.44
Kansas	222	0.18	84	-0.82	138	0.83
Kentucky	156	1.41	63	1.52	93	1.17
Louisiana	206	0.42	94	0.35	112	0.37
Maryland	249	-0.85	109	-1.27	140	-0.42
Massachusetts	350	-0.92	140	-1.11	210	-0.62
Michigan	759	-0.56	332	-0.38	427	-0.63
Minnesota	318	-0.57	144	-0.55	174	-0.49
Mississippi	147	0.55	57	0.26	90	0.72
Missouri	424	0.31	179	-0.14	245	0.61
Montana	60	0.12	21	1.43	39	-0.50
New Hampshire	85	-1.66	39	-1.16	46	-1.86
New Jersey	433	-0.39	192	0.08	241	-0.67
New York	1103	-0.21	450	0.44	653	-0.56
North Carolina	574	1.34	258	1.21	316	1.23
North Dakota	166	-0.54	70	-0.53	96	-0.52
Ohio	668	-0.11	295	0.13	373	-0.26
Oklahoma	210	0.13	91	0.28	119	0.04
Oregon	215	-0.49	92	-1.12	123	0.07
Pennsylvania	753	0.37	323	0.41	430	0.30
Rhode Island	82	-0.61	30	-1.60	52	0.09
South Carolina	224	0.33	98	0.58	126	0.03
South Dakota	52	-0.86	25	-0.41	27	-1.19
Tennessee	485	1.60	195	0.93	290	1.84
Texas	941	0.42	411	0.51	530	0.31
Utah	104	2.03	49	1.48	55	2.14
Vermont	80	-1.20	33	-0.84	47	-1.30
Virginia	603	-0.42	248	-0.43	355	-0.28
Washington	288	-0.39	139	-0.27	149	-0.50
West Virginia	138	1.61	62	1.74	76	1.32
Wisconsin	365	-0.47	164	-0.87	201	-0.08
Wyoming	70	-1.47	43	-1.41	27	-1.52

Note: The sexism measure for each sample (e.g. all, men, women) are standardized to have a mean of 0 and standard deviation of 1 in the sample of 44 states. Larger values indicate higher measured sexism. See text for details on the construction of the sexism indices.

Table A5: Effect of Background Sexism on Labor Market and Non-Labor Market Outcomes, Including State of Residence Fixed Effects

	Labor Force Participation		Log Wages, conditional on working		Never Married (Age 20 to 40)		Age at First Birth (Age 20 to 40)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	All Migrants				Female Migrants Only			
Sexist Beliefs in State of Birth:								
Overall*Female dummy	-0.006*	-0.007*	-0.012***	-0.011***				
	[0.003]	[0.004]	[0.003]	[0.004]				
Overall	0.000	-0.001	0.002	0.002	-0.012***	-0.010***	-0.114***	-0.111***
	[0.001]	[0.001]	[0.003]	[0.005]	[0.003]	[0.003]	[0.021]	[0.028]
Controls (all interacted with Female dummy for Cols (1) to (4)):								
Additional controls		X		X		X		X
State of Residence FE	X	X	X	X	X	X	X	X
Region of Birth FE, Year FE	X	X	X	X	X	X	X	X
Observations	15,103	14,083	15,063	14,055	7,522	7,016	7,453	6,962
R-squared	0.936	0.939	0.929	0.931	0.832	0.839	0.574	0.587

Note: The data are from the 1980, 1990, 200 US Census and 2010-2012 ACS (3-year aggregate). The sample is restricted to whites age 25 to 64 for the labor market outcomes and age 20 to 40 for the non-labor market outcomes (females only) who are not currently living in their state of birth. The unit of observation is at the state of birth*state of residence*gender*year level for the labor market outcomes and at the state of birth*state of residence*year level for the non-labor market outcomes. Sexist beliefs are normalized to have a mean of 0 and standard deviation 1 across the 44 states with available information to construct the index of sexist beliefs in the GSS. "Additional controls" include the mean NAEP-LTT reading and math scores of boys in each state of residence and state of birth s well as the predicted female employment share in the state of residence based on the state's 1950 industry structure (all interacted with the female dummy for the labor market outcomes). All the regressions also control for state of residence fixed effects, region of birth fixed effects and year fixed effects (all interacted with the female dummy for the labor market outcomes). Regressions are weighted by the number of observations in each cell for each outcome. Standard errors clustered at the state of birth are reported in parentheses. ***significant at 1%, **at 5%, *at 10%.

Table A6: Robustness to Controls for State-Level Differences in School Quality and Industry

	Labor Force Participation Gap (Female - Male)			Selection-Corrected Wage Gaps (Female-Male)		
	(1)	(2)	(3)	(4)	(5)	(6)
Male Sexist Beliefs in State of Residence:						
10th Percentile	0.000	0.002	0.011*	-0.001	-0.000	0.024*
	[0.008]	[0.007]	[0.006]	[0.009]	[0.011]	[0.013]
Median	-0.020***	-0.017**	-0.016**	-0.034**	-0.034**	-0.041***
	[0.007]	[0.007]	[0.006]	[0.013]	[0.013]	[0.012]
90th Percentile	0.004	0.004	0.002	0.002	0.001	0.015
	[0.006]	[0.006]	[0.006]	[0.013]	[0.014]	[0.012]
Observations	41	41	41	41	41	41
R-squared	0.282	0.297	0.545	0.255	0.256	0.512
Controls:						
Education Quality (NAEP test scores for boys)			X			X
Predicted Female Employment Share			X			X
Overall Average Female Sexism		X	X		X	X

Note: The residual female-male employment and selection-corrected wage gaps are estimated using the sample of whites age 25 to 64 from the 1977-2017 May/ORG CPS data and control for the number of years of schooling, age dummies, gender-specific year effects, and state fixed effects. Sexist beliefs are normalized to have a mean of 0 and standard deviation 1 across the 44 states with available information to construct the index of sexist beliefs in the GSS (excluding DC). Regressions are weighted by the inverse of the variance of the outcome variable. The sample in column (1) is restricted to states with non-missing information on NAEP math test scores and predicted female employment share. The proxies for education quality at the state-level are NAEP-LTT mean math and reading test scores for boys. See text for details on how the predicted female employment share is computed. Columns (2) and (3) include controls for overall average female sexism. Robust standard errors are reported in parentheses. ***significant at 1%, **5%, *10%.

Table A7: Sexism and Women's Outcomes, Controlling for Religiosity

	LFP Gap (Female - Male)	Selection-Corrected Log Wage Gaps (Female-Male)	Share of Females Nevermarried	Average Female Age at First Child
	(1)	(2)	(3)	(4)
	A. Male vs. Female Sexism			
Sexist Beliefs in State of Residence				
Male average	-0.015***	-0.034**		
	[0.005]	[0.013]		
Female average			-0.024***	-0.218**
			[0.008]	[0.086]
Religiosity	-0.003	0.001	-0.022***	-0.076
	[0.006]	[0.012]	[0.007]	[0.063]
	B. Percentiles of Male Sexism			
Male Sexist Beliefs in State of Residence:				
10th Percentile	0.001	-0.001		
	[0.007]	[0.010]		
Median	-0.017**	-0.030**		
	[0.007]	[0.014]		
90th Percentile	0.003	-0.003		
	[0.006]	[0.011]		
Religiosity	-0.005	-0.002		
	[0.006]	[0.013]		
Observations	44	44	44	44

Note: See notes to Table 6. The specifications are similar to those reported in Tables 6, 7, and 8, but include an additional control for religiosity at the state-level as measured in the GSS ("how often do you attend religious services," 0 (never) to 8 (more than once a week)).

Table A8: Placebo Test using Religiosity (Percentile Tests)

	LFP Gap (Female - Male)	Selection-Corrected Log Wage Gaps (Female-Male)	LFP Gap (Female - Male)	Selection-Corrected Log Wage Gaps (Female-Male)
	(1)	(2)	(3)	(4)
Religiosity in State of Residence:	Male		Female	
10th Percentile	0.005	0.003	0.007**	0.000
	[0.004]	[0.005]	[0.003]	[0.007]
Median	-0.008	-0.005	-0.007	0.003
	[0.007]	[0.011]	[0.008]	[0.012]
90th Percentile	-0.010*	-0.024***	-0.015**	-0.034***
	[0.006]	[0.008]	[0.006]	[0.009]
Observations	44	44	44	44
R-squared	0.228	0.171	0.328	0.253

Note: See notes to Table 7 for details on the sample and how the outcomes are constructed. The specifications are similar to those reported in columns (1) and (2) of Table 7, except that the percentiles of male and female sexism are replaced with percentiles of male and female religiosity at the state-level. Standard errors clustered at the state of birth are reported in parentheses. ***significant at 1%, **at 5%, *at 10%.