On the Marriage Wage Premium*

Brendon McConnell†
University of Southampton

Arnau Valladares-Esteban‡
University of St. Gallen and Swiss Institute for Empirical Economic Research

November 15, 2019

Abstract

It has long been observed that married men earn higher wages than their single counterparts. In this paper, we document that, in the last decades, an analogous pattern has emerged for women. Married women experienced a wage *penalty* until the 1990s, whereas nowadays there is a sizable premium. To measure the causal effect of marriage on wages presents three main challenges: a significant part of the female population does not participate in employment (sample-selection bias), there might be some variables that are relevant for both wages and the propensity to marry that are not observable (omitted-variable bias), and wages may also affect marriage decisions (simultaneity bias). We apply a variety of techniques, along with a novel instrument based on local social norms towards marriage, to show that marriage has a positive causal effect on wages for both genders, although a sizable part of the observed correlation is spurious. We also show that the effect of marriage on wages is heterogeneous. Further, we present evidence that the main hypotheses discussed in the literature to explain the marriage wage premium for men, household specialization and employer discrimination, have little support in the data.

*We thank Michael Knaus, Michael Lechner, Joan Llull, and seminar participants at the Berlin Applied Micro Seminar, the University of Bristol, the University of Nottingham, the University of St. Gallen, and the University of York for valuable comments. All errors are ours.

†B.I.Mcconnell@soton.ac.uk.

‡arnau.valladares@gmail.com.

1 Introduction

Over the last century, the U.S. has experienced a dramatic shift in the structure of families and the role of women. If one looks at this transformation from the point of view of the family, the patterns of marriage, divorce, fertility and assortative matching have all changed remarkably.[1] Placing the lens on gender, the labor market outcomes of women have evolved significantly. From labor force participation to wages, a wide range of indicators show that the economic role of women in the labor market is more prominent now than ever before.[2] One aspect of this transformation that has received little attention is the evolution of the relationship between wages and marriage.[3] While some authors show that married men earn higher wages than their single counterparts, the so-called *Marriage Wage Premium* (MWP), there is much less work on this relationship for women.[4]

We make three key contributions in this paper. The first is to document the emergence of a MWP for women over the last decades. While until the mid 1980s women experienced a marriage wage *penalty*, from the 1990s there is a sizable marriage wage *premium*. Interestingly, the relationship between marriage and wages for men has remained essentially constant over the same period. The MWP for women is relevant for, at least, two reasons. First, a large deal of the changes in the economic role of women are, in fact, a reflection of the transformation in the economic role of *married* women. As an example, most of the increase in female labor force participation that occurred after WWII can be accounted for by the growth in the employment of married women. Hence, it is crucial to analyze the relationship between wages and marriage for both genders in order to understand the social transformation of the last decades. Secondly, some of the theories that have been proposed to explain the MWP of men rely on intra-household arguments. In particular, the literature has considered the hypothesis that the origin of the MWP of men is related to within-household specialization.[5] The underlying idea is that married men are able to devote more resources to their careers than their single counterparts because their wives specialize in home production. However, the presence of a MWP for both women and men casts doubt on this hypothesis.

Our second contribution is to present evidence on the causal effect of marriage on wages for both men and women. Establishing a causal effect of marriage on wages presents three main challenges. First, there may be unobservable variables that affect both the propensity to marry and wages. That is, the estimated coefficients on marriage in a wage equation may suffer from omitted-variable bias. Second, for women, there is a significant part of the population that does not participate in employment. The underlying economic decision that generates this outcome implies that the sample of observed wages is not a random representation of the population. Hence, the estimated coefficients in the wage equation might suffer sample-selection bias. Thirdly, there might be an issue of reverse causality if wages *also* affect the probability of being married.

[1]See Greenwood, Guner, and Vandenbroucke (2017) and Lundberg and Pollak (2007).

[2]There is a vast literature studying the evolution of the labor market outcomes of women. See, for example, Attanasio, Low, and Sánchez-Marcos (2008), Blau and Kahn (2007, 2017), Goldin (2014), Fernández (2013), and Olivetti (2006).

[3]Some authors study the relationship between marriage and other outcomes. For example, Choi and Valladares-Esteban (2018) or Guner, Kulikova, and Llull (2018).

[4]See Hill (1979), Korenman and Neumark (1992), Juhn and McCue (2016), and Pilossoph and Wee (2019).

[5]See Loh (1996), Cornwell and Rupert (1997), Ahituv and Lerman (2007), and Killewald and Gough (2013).

We tackle all of these issues with different methodologies and different types of data. We start by exploiting between-individual variation using the repeated cross sections of the Current Population Survey (CPS). We apply the bounding technique of Oster (2019) to show that, although unobservable factors appear to bias the estimate of the relationship between marriage and wages upwards, the role of unobservables would need to be *large* to drive this effect to zero. For women, we correct for sample-selection bias using a novel exclusion restriction, the age of the youngest child in the household, which enables us to control for the presence of children in the wage equation. We show that the patterns of the MWP are robust to this correction. We then move to panel data from the National Longitudinal Survey of Youth 1979 (NLSY79) in order to use within-individual variation. We find that individual fixed effects account for part of the relationship between marriage and wages but a positive and significant MWP remains for both genders. Finally, we present a novel variable to instrument for marriage in the wage equation based on local social norms. We use the share of married people who have the same gender, live in the same state, and have the same values for the indicators of college education and presence of children in the household, but are 6 to 15 years older than each individual in our analysis to proxy for the relevant social norms that affect the decision to marry of that individual. The coefficients estimated from the instrumental-variable equations are broadly in line with our previous findings albeit slightly larger for both genders.

Most of the literature on the MWP has focused on establishing whether the positive correlation between marriage and wages for men reflects a causal effect of marriage on wages or this correlation is a consequence of a mechanism by which some men *select* into marriage.[6] Although some authors refer to the latter issue as *selection*, we prefer to think of it as endogeneity.[7] We distinguish between two distinct sources of endogeneity, namely, omitted-variable bias and reverse causality. When we account for omitted-variable bias, both by using bounds and fixed effects, we find that the correlation between wages and marriage decreases with respect to the uncorrected estimate. This result is consistent with the idea that there are some unobservables that positively affect both wages and the propensity to marry. However, when we instrument for marriage, which tackles both omitted-variable bias and reverse causality, we find a slightly higher estimate. Our interpretation of this higher estimate is not that the potential mechanism of reverse causality implies that people with higher wages are less likely to marry but rather that the higher estimate is due to marriage having heterogeneous effects. That is, our instrumental variable estimate reflects the effect of marriage on wages for a set of compliers on the higher end of the treatment-effect distribution. We support this hypothesis by showing that the compliers are likely to be younger and less educated than the whole sample, the average and marginal treatment effects are different, and that the relationship between marriage and wages changes dramatically along the wage distribution. In sum, we find that part of the correlation between marriage and wages is due to omitted-variable bias but there still exist a causal positive effect of marriage on wages, at least, for a sizable fraction of the population.[8]

[6]There are exceptions, for example, Gray (1997) and Maasoumi, Millimet, and Sarkar (2009).

[7]See Korenman and Neumark (1991), Ginther and Zavodny (2001), Stratton (2002), Antonovics and Town (2004), and Krashinsky (2004).

[8]Killewald and Lundberg (2017) and Ludwig and Brüderl (2018) argue that there is no causal effect of marriage on wages. Both studies estimate rich fixed-effects models on NLSY79 data. We present a different approach for two reasons. First, we think that an over parametrization of the individual fixed effects might wash away some of the (true) effects of marriage on wages. Secondly, if the effects of marriage start before the exact moment of becoming

Our third contribution is to test the hypothesis discussed in the literature which poses that the MWP for men might be generated or amplified by positive employer statistical discrimination. The idea is that employers might believe that marriage is positively associated with some determinants of productivity which are hard to observe and use marriage as a proxy for those instead. We start by adapting the empirical specification in Altonji and Pierret (2001) for education to the case of marriage. We do not find evidence of any statistical discrimination towards married men despite the fact that the estimates are consistent with the presence of employer learning. In fact, we find that the MWP for men increases over the work life, exactly the opposite that the presence of statistical discrimination implies. For women, we find a similar pattern. Interestingly, the coefficients show that the MWP for women is a composite of a penalty at the beginning of the work life that evolves into a premium as experience increases. This result is relevant because of two reasons. First, it clarifies how a penalty associated with traditional gender norms (wives allocate more resources to non-market work than husbands) and the marriage wage premium for women can coexist.[9] Secondly, the fact that the MWP for women increases with experience in the labor market is additional evidence that the hypothesis of household specialization is not supported by the data. We further specify a framework aimed at decomposing public and private learning in the spirit of Pinkston (2009). Our estimates confirm that there is no evidence to support the hypothesis of positive statistical discrimination towards married individuals. Taken together, we interpret our findings as evidence that statistical discrimination is not a relevant mechanism behind the marriage wage premium of women and men.

The rest of the paper is organized as follows. In Section 2, we describe the data we use and the sample restrictions we impose. In Section 3, we present new evidence on how the conditional correlation between marriage and wages has differentially evolved across genders over the past four decades. Section 4 presents evidence on the causal effect of marriage on wages. In Section 5, we test the extent to which the statistical discrimination hypothesis can explain the marriage wage premium. Finally, Section 6 concludes.

2 Data

We use data from the March Supplement of the CPS from 1977 to 2018 and from the NLSY79 for years 1979 to 2012.[10] The sample restrictions we apply aim both at making our results comparable to the literature and equivalent across the two data sources we use.

2.1 CPS

Our CPS sample consists of white non-Hispanic civilians who are in their prime age (between 25 and 54 years old), not living in group quarters, and for whom we have no missing data on relevant demographic characteristics. We further exclude from the sample self-employed workers, individuals working in the private household sector, and agricultural workers. The group of married individuals consists of people that declare to be married and living with their spouse in

married, for example upon deciding on marriage, then a fixed-effects model amplifies the mismeasurement of the treatment of marriage.

[9] Albanesi and Olivetti (2009) show how traditional gender norms can create a gender wage gap.

[10] The CPS data is made publicly available by Flood, King, Ruggles, and Warren (2015).

the same household. The non-married group is composed only of never married individuals to keep consistency with the literature on the MWP for men.[11]

Using the information on weeks worked last year and usual hours of work per week, we build a variable that proxies the total number of hours worked last year for each individual on our sample. Then, we divide non-allocated total labor income last year, expressed in 1999 US dollars, by the total number of hours worked last year to obtain a measure of hourly wages. As it is common in the literature, we trim the top and bottom 1% of our measure of hourly wages to limit the influence of outliers. We disregard the hourly wage measure of those individuals that report less than 100 hours of work last year and consider them never employed last year.[12] We use the Annual Social and Economic Supplement weights in all CPS-related analysis.

Table 1: Descriptive Statistics, CPS
Means, Standard Deviations in Parentheses

	Men		**Women**	
	1977-1992	**2003-2018**	**1977-1992**	**2003-2018**
Sample Size	276,250	293,075	300,609	323,231
Married	0.808	0.693	0.877	0.784
Employed	0.945	0.893	0.681	0.756
Hourly Wage (1999 Dollars)	19.41	19.85	12.58	15.99
	(9.69)	(11.53)	(6.97)	(9.80)
Age	37.35	39.34	37.55	39.57
	(8.55)	(8.81)	(8.55)	(8.78)
Highest Level of Education:				
HS Dropout	0.134	0.058	0.125	0.042
HS Graduate	0.372	0.295	0.452	0.248
Some College	0.200	0.275	0.198	0.290
College Graduate	0.165	0.254	0.144	0.281
Advanced Graduate	0.129	0.118	0.081	0.139
Number Children, 0-4	0.293	0.244	0.279	0.257
	(0.593)	(0.555)	(0.579)	(0.566)
Number Children, 5-17	0.963	0.778	1.09	0.921
	(1.19)	(1.08)	(1.21)	(1.11)

Our final sample contains 1,193,165 observations, 569,325 men and 623,840 women. For the reasons outlined in Section 3, we further split the sample into two separate time periods, 1977-1992 and 2003-2018. Table 1 presents key descriptive statistics for our working CPS sample. The observed patterns over time, both for men and women, are consistent with well-documented

[11]Separated, divorced, and widowed individuals are excluded from the sample. That is, we focus explicitly on legally married individuals who live in the same household as their spouse. We ignore cohabitation which is not subject to the legal and social obligations of marriage. In Appendix B, we reproduce our main analysis using a sample in which the non-married group is solely composed of separated and divorced individuals. The coefficients estimated with this alternative definition of the non-married group are in line with our main results.

[12]We experimented with restricting the definition of the employed to full-time full-year workers, that is, employed for at least 50 weeks in the past year for 35 or more hours per week. The key results are not substantively different using this alternative specification.

patterns in the US labor market during the last decades. Namely, the decrease in the share of married individuals, the increase in female labor force participation, the increase in educational attainment, and the reduction in the number of children.

2.2 NLSY79

Our NLSY79 sample consists of white non-Hispanic civilians who are between 22 and 55 years old for whom we have no missing data on relevant demographic characteristics. We use solely the male and female cross-sectional sub-samples. These are designed to be representative of the non-institutionalized civilian US population born in the years 1957-1964. We do not weight the analysis based on the NLSY sample.

We only consider individuals with valid marriage histories who enter the sample unmarried and subsequently either remain unmarried or marry and stay married for the years surveyed. We drop respondent-years for periods when individuals are enrolled in formal education, are self-employed, or working for fewer than 10 hours per week. This means we drop respondent-years when individuals are not working. We trim the top and bottom 1% of our measure of hourly wages to limit the influence of outliers. Hourly wages are expressed in 2006 US dollars. Finally, we require all individuals to have at least two observations. As it becomes clearer below, this is in order to be able to compare a consistent sample across different regression specifications.

Table 2: Descriptive Statistics, NLSY79
Means, Standard Deviations in Parentheses

	Men	**Women**
Sample Size	14,736	10,375
Number of Individuals	1,446	1,121
Married	0.458	0.458
Ever Observed Married in Panel	0.778	0.824
Hourly Wage (2006 Dollars)	18.75	15.48
	(10.20)	(7.72)
Job Tenure	4.24	4.08
	(4.49)	(4.46)
Experience	10.76	10.22
	(6.33)	(6.27)
Age	30.80	30.17
	(6.62)	(6.59)
Highest Level of Education:		
HS Dropout	0.079	0.022
HS Graduate	0.434	0.375
Some College	0.188	0.233
College Graduate	0.205	0.259
Advanced Graduate	0.094	0.111
Urban Residence	0.787	0.799
Number Children, 0-17	0.533	0.452
	(0.900)	(0.822)

The NLSY79 survey allows us to construct detailed measures of both experience and tenure with the current employer. We construct such variables prior to the above sample restrictions. Our final sample consists of 1,446 men and 1,121 women which correspond to 14,736 and 10,375 observations respectively. Table 2 presents descriptive statistics related to marriage, wages, tenure, experience, age, education, and the number of children.

Although we apply seemingly similar sample restrictions to construct our CPS and NLSY79 samples, a comparison between Table 1 and Table 2 indicates that there are relevant differences in observable characteristics between the two samples. Namely, the NLSY79 sample is composed of younger and slightly less educated individuals than our CPS sample. Moreover, the CPS sample contains different birth cohorts while the NLSY79 sample is focused on one cohort. Given that an important part of the discussion on the MWP is related to omitted variable bias, in the following section, we discuss how to overcome the potential issues that the differences between the CPS and NLSY79 samples may have on our inference. The concern about equivalence between samples is not only relevant for our analysis. The literature on the MWP has revolved around common questions but some authors have used the CPS while others have used the NLSY79. However, little attention has been devoted to understand whether the differences between the two data sources are responsible for any of the differences in the derived inference.

2.3 Bridging across the Two Samples

We balance two criteria when making sample restriction decisions. First, we are guided by the sample restrictions commonly made in the literature. Secondly, in order to be able to compare results across the two data sources, we make as many common sample restrictions decisions as possible. A constraint to the latter criterion is that we apply different methodologies to the two samples which require potentially divergent data requirements. For example, we correct for sample-selection bias with a sample selection model when using the CPS sample while we rely on individual fixed effects on the NLSY79 sample. To estimate a fixed-effect model imposes significant restrictions on the frequency of observations we require for each person in the sample. As a consequence, the two samples diverge in the labor market attachment of the individuals they consider. Another key difference across the two samples is the period of analysis, which combined with the age restrictions we use, further implies distinct birth cohorts in each sample. The CPS sample includes the years 1997-1992 and 2003-2018 while the NLSY79 sample includes 1979-2012. Moreover, the attrition and the sample restrictions to obtain the NLSY79 sample imply that the median year in that sample is 1990 while, in the CPS sample, the median years are 1985 and 2010 for each period considered. This difference is of direct relevance to the estimation of the wage effects of marriage for women, as shown in the next section.

In order to bridge the differences between the two data sources, we use the data in the CPS to construct a set of samples which contain individuals with similar observable characteristics as the individuals in our NLSY79 sample. We employ different matching approaches to achieve this goal in order to check that the results we obtain are robust to the type of matching technique used. We name the samples from the CPS data which are matched to our NLSY79 sample, pseudo-NLSY samples.[13] We use these pseudo-NLSY samples to run the same main analyse as on our CPS sample. In Section C.2 of Appendix C, we present the results we obtain for all the

[13]The details related to the construction of these samples are discussed in Section C.1 of Appendix C.

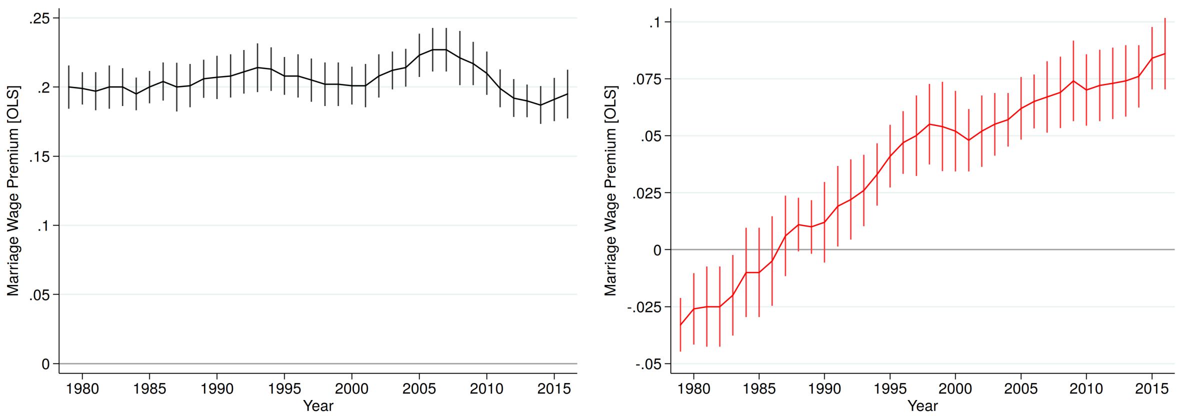
pseudo-NLSY samples. The main conclusion of this exercise is that the estimates are broadly similar both across pseudo-NLSY samples and with respect to the main CPS sample. This result allows us to better discuss the potential divergence in inference that the distinct methodologies used on the CPS and NLSY79 samples yield, as we know the role that the differences between the two samples play.

3 The Correlation between Marriage and Wages over Time

We start by measuring the conditional correlation between being married and hourly wages over time separately for men and women. Using Ordinary Least Squares (OLS) on the CPS sample, we estimate the following linear regression model:

$$y_i = \alpha M_i + X_i^{\prime}\beta + \theta_s + \phi_t + \epsilon_i, \tag{1}$$

where y_i is the natural logarithm of our measure of hourly wages of observation i as described in Section 2.1, M_i is a dummy variable that equals 1 when an individual reports to be married and living with their spouse, X_i is a set of demographic controls which consists of education-category dummies, dummies for the number of own children in the household (separately for ages 0-4 and ages 5-17), and linear and quadratic terms in potential experience. θ_s and ϕ_t are state and year fixed effects respectively. We cluster standard errors at the state level.



(a) Men. (b) Women.

Figure 1: Conditional correlation between being married and wages

Notes: The figures plot $\hat{\alpha}$ from Equation 1 and 95% confidence intervals (based on state-clustered standard errors) as the vertical spikes. Each point centered on t is estimated using observations from year $t-2$ to $t+2$. The dependent variable in all columns is the natural log of wages. Year and state fixed effects are included in all regressions. The following additional controls are included: dummies for highest level of educational attainment, potential experience, potential experience squared, dummies for number of children below the age of 5, dummies for number of children aged 5-17. CPS 1977-2018.

The conditional correlation between marriage and hourly wages has changed dramatically for women (Figure 1b). At the beginning of the period, marriage was associated with a wage penalty of roughly 2.5%. This penalty linearly reduced over time until the mid 1980s. In the late 1980s, a marriage premium emerged and it continued to increase until the end of the sample period. By 2018, marriage is associated with a premium of around 7.5%. This change in the correlation between being married and the wages of women is specially important in the context of the evolution of female labor force participation and the decline of marriage.

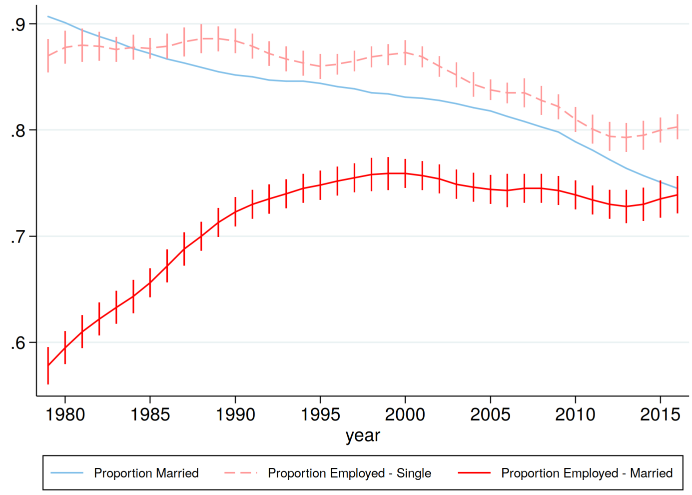


Figure 2: Employment rates and share of married women. CPS 1977-2018.

Notes: The employment rates are from a regression of employment on a constant, conditional on marital status. 95% confidence intervals (based on state-clustered standard errors) are the vertical spikes. Marriage rates are from a regression of employment on a constant. Each point centered on t is estimated using observations from year $t-2$ to $t+2$. CPS 1977-2018.

As seen in Figure 2, the emergence of a marriage wage premium for women has been coupled with a secular decline in the share of married women and the convergence in the employment rates of married and single women. That is, the labor market has moved between two significantly different scenarios. Before the 1990s, marriage was the norm, female employment when married was less common than when single, and there was a wage penalty for married women and a premium for men. From the beginning of the 2000s, marriage is less prevalent, female employment when married is almost as high as when single, and marriage is associated with a wage premium for both genders. Throughout the paper, our analysis distinguishes systematically between these two periods in order to shed light on the possible factors that affect the returns to marriage for both genders. In particular, we consider the first and the last 16 years of our sample, that is, the 1977-1992 and the 2003-2018 periods.[14]

For men, despite the remarkable changes in family structure, the decrease in the marriage rate, the increase of divorce, the surge of assortative mating, and the dramatic change in the role of married women in the economy, the conditional correlation between being married and hourly wages has remained remarkably stable over the past four decades. As shown in Figure 1a, the wages of married men are around 20% higher than those of their single counterparts.

4 Towards Measuring the Causal Effect of Marriage on Wages

To measure the causal effect of marriage on wages presents three main issues: omitted-variable bias, sample selection for women, and reverse causality. We tackle all these issues in this section

[14]These two periods are also consistent with changes in the labor force participation of single women related to fertility as discussed in Kleven (2019).

using different methodologies and exploiting both between-individual variation (using the CPS sample) and within-individual variation (using the NLSY79 sample). In Section 4.1, we use the bounding technique in Oster (2019) to assess how unobservables might affect the relationship between marriage and wages. We compute a measure of how important the effect of unobservable characteristics needs to be in order to drive to zero the correlations estimated in the OLS regressions of Section 3. In Section 4.2, we present a new exclusion restriction for the two-step Heckman (1979) correction for the wages of women which allows to control for children in the wage equation. Both sections 4.1 and 4.2 rely on the between-individual variation found in the repeated cross sections of the CPS. In Section 4.3, we use the within-individual variation of the panel structure from the NLSY79 to estimate the married coefficient in a wage equation with individual fixed effects. Finally, in Section 4.4, we present a new instrument for marriage based on social local norms that we apply to the CPS data. For women, we combine the instrumental variable approach with the selection correction of Section 4.2.

4.1 Bounds on the Married Coefficient

We use the extension in Oster (2019) of Altonji, Elder, and Taber (2005) to bound the effect that unobservable characteristics might have on the estimated married coefficient. The key idea underlying this approach is that, with a set of assumptions, we can use the relationship between marriage and the observables to infer something about the relationship between marriage and the unobservables. That is, we use the extent of selection on observables to bound the impact of selection on unobservables.

4.1.1 Empirical Specification

Let us start by assuming that the true data-generating process is given by

$$y = \alpha M + X\beta + \theta + \phi + W_2 + \nu, \tag{2}$$

where $\alpha M + X\beta + \theta + \phi$ is analogous to Equation 1, in matrix notation, and W_2 is an index that represents the role of unobservable variables in the wage equation. We assume that W_2 is orthogonal to the observable covariates. Given that the components of W_2 are indeed not observable we cannot estimate Equation 2. In order to bound the effect of not being able to include W_2 when estimating α, we consider the following two specifications:

$$y = \alpha M + \nu, \tag{3}$$

$$y = \alpha M + X\beta + \theta + \phi + \nu. \tag{4}$$

The estimate of α in Equation 3, which we denote as $\mathring{\alpha}$, measures the unconditional correlation between marriage and wages. When we estimate Equation 3, when can also compute how much of the dispersion in wages is explained by the dispersion in marriage. Let us denote the R^2 of estimating Equation 3 as $\mathring{R}^2$. Analogously, α in Equation 4 captures the conditional-on-covariates correlation between marriage and wages, we denote the estimates from this equation $\tilde{\alpha}$ and $\tilde{R}^2$. Intuitively, one can see that by comparing how much $\tilde{\alpha}$ changes with respect to $\mathring{\alpha}$, mediated by $\tilde{R}^2$ and $\mathring{R}^2$, it is possible to infer the role of observables in determining the relationship between wages and marriage.

In order to establish a bound on the effect of unobservables on α, Oster (2019) exploits the idea that if we know the R^2 of estimating Equation 2, we can approximate the estimate of α in Equation 2 as

$$\alpha^* \approx \tilde{\alpha} - \delta\left[\mathring{\alpha} - \tilde{\alpha}\right] \frac{R^2_{max} - \tilde{R}^2}{\tilde{R}^2 - \mathring{R}^2}, \tag{5}$$

where α^* and R^2_{max} are the estimate of α and the R^2 of Equation 2, respectively. The parameter δ defines the relative role of observables and unobservables in determining the *true* relationship between wages and marriage. We follow Oster (2019) and consider $R^2_{max} = min(1.3 * \tilde{R}^2, 1)$. Further, as Altonji et al. (2005) and Oster (2019), we select $\delta = 1$ for define the bounds.

Finally, note that the idea reflected in Equation 5 allows for two measures. First, by considering that the effect of observables and unobservables is equal, that is $\delta = 1$, we can compute bounds from the estimate of α from Equation 4 (which is identical to Equation 1). We refer to these bounds as the Oster bounds. Secondly, we can ask what value of δ is required to drive the estimate of α from Equation 1 to zero. In other words, how much larger the influence of unobservable variables needs to be, relative to that of the observables, in order for the married coefficient we estimate from Equation 1 to be zero. We report and discuss both measures in the following section.

4.1.2 Results

Table 3: OLS with Oster Bounds, CPS

Dependent Variable: Log(Hourly Wage) in 1999 Dollars

	Men		**Women**	
	1977-1992	**2003-2018**	**1977-1992**	**2003-2018**
Married, Unconditional	0.311***	0.345***	-0.140***	0.091***
	(0.010)	(0.009)	(0.009)	(0.007)
Married	0.202***	0.208***	-0.006	0.074***
	(0.005)	(0.005)	(0.005)	(0.005)
Oster Bounds	[0.124, 0.202]	[0.102, 0.208]	[-0.006, 0.056]	[0.067, 0.074]
δ Required for Coefficient of 0	1.869	1.573	0.107	4.854
Unadjusted R^2	0.229	0.266	0.189	0.227
$R^2_{\mathbf{max}}$	0.298	0.345	0.245	0.295
Adjusted R^2	0.229	0.265	0.188	0.226
Observations	261,737	267,259	204,261	245,038

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. Standard errors are reported in parentheses, where these are clustered by state. The dependent variable in all columns is the natural log of wages. Year and state fixed effects are included in all regressions. The following additional controls are included: dummies for highest level of educational attainment, potential experience, potential experience squared, dummies for number of children below the age of 5, dummies for number of children aged 5-17. In brackets we report bounds on the OLS estimate accounting for selection on unobservables using the Oster [2019] method: the bounds are set assuming the coefficient of proportionality is zero or one. Below the bounds we report the coefficient of proportionality (δ) that is required for the implied point estimate to be zero.

Table 3 presents the Oster bounds and the δ required to drive the estimated correlation between marriage and wages from Equation 1 to zero, for both men and women in the two sample periods we consider. For men, the results are roughly similar between sample periods. Namely, the Oster bounds indicate that if unobservable variables have a symmetric effect on the relationship between wages and marriage as that of observables, the estimated coefficient decreases by around half. In the period 1977-1992, the coefficient moves from 0.202 down to 0.124, a 39% decrease. For the years 2003-2018, the coefficient changes from 0.208 down to 0.102, a 51% decline. Importantly, the bounds do not include zero. That is, unobservable variables need to have a bigger impact on the observed relationship between wages and marriage than observables for the married coefficient to be zero. In particular, unobservables need to affect the estimated coefficient more than 1.5 times than observables. Both the wider bounds and the lower δ *required for a coefficient of zero* in the 2003-2018 period compared to the 1977-1992, indicate that the role of observable variables in explaining the relationship between wages and marriage for men has decreased over time.

For women in the 1977-1992 period we find that the unconditional penalty for married women (-0.140) can be fully explained by differences in observable variables between married and single women (the married coefficient is -0.006). This is consistent with the trend we observe in Figure 1b. That is, the 1977-1992 comprises a period of time in which the marriage wage penalty of women is decreasing and starts to become a premium. The bounds indicate that unobservables might drive the estimated coefficient to a *higher* value, from -0.006 to 0.056. This pattern diverges from that observed for men and women in the 2003-2018 period. The fact that unobservable variables drive up the coefficient is consistent with the hypothesis that the emergence of a MWP for women is related to the change in the selection-into-employment pattern, from negative to positive, uncovered by Mulligan and Rubinstein (2008). We discuss selection in our sample in Section 4.2. In the 2003-2018 sample period, the MWP for women is clearly positive. Similarly to what we compute for men, the bounds indicate that, assuming a symmetric effect to observable variables, unobservables drive the estimated coefficient from 0.074 to 0.067, a 10% decrease. In other words, unobservable variables need to be more than four times more relevant in the wage regression for the estimated married coefficient to be zero (the δ *required for a coefficient of zero* is 4.854).

Our interpretation of the results of the results is twofold. First, the main patterns observed in Figure 1 are unlikely to be reversed by unobservable variables. That is, the existence of a MWP for men and the emergence of a MWP for women is robust to correcting for omitted-variable bias. Secondly, the emergence of a MWP premium for women appears to be related to the role of unobservable variables in the wage equation. This is consistent with the patters of selection-into-employment described in the literature, which we confirm in Section 4.2. The idea is that while pre-1990s married women are negatively selected into employment, the progressive increase in the labor force participation of women implied that more productive women joined the labor force driving up the wages for this group. We look at whether this trend alone can explain the MWP for women in the next section.

4.2 The Role of Selection into Employment

For women, the association between marriage and hourly wages observed when estimating Equation 1 might be biased due to the fact that a sizable proportion of women does no participate in employment and, therefore, their wages are not observed. Moreover, as pointed out by Mulligan and Rubinstein (2008), the pattern of selection into employment has changed substantially over the last decades. In particular, Mulligan and Rubinstein (2008) find that the selection of women into full-time full-year employment evolved from negative in the 1970s to positive in the 1990s. Hence, it is crucial to address the selection bias induced by participation in the labor market in order to correctly estimate the association between marriage and hourly wages for women.

4.2.1 Empirical Specification

We implement two specifications of the classic sample selection model. The bivariate normal-maximum likelihood and the Heckman two-step correction. Our specification for the former is given as:

$$E_i = \mathbb{1}\{\gamma_1 Z_{E,i} + \gamma_2 M_i + X_i^{'}\gamma_3 + \theta_{1s} + \phi_{1t} + \epsilon_{1i} > 0\} = \mathbb{1}\{Z_i^{'}\gamma + \epsilon_{1i} > 0\}, \tag{6}$$
$$y_i = \alpha M_i + X_i^{'}\beta + \theta_{2s} + \phi_{2t} + \epsilon_{2i}, \tag{7}$$

where $\mathbb{1}$ is an indicator function for employment, that is, when an individual works, $E_i = 1$ and y_i is observed. $Z_{E,i}$ is the variable that captures the exclusion restriction. We assume that the structure of the error terms is given by:

$$\begin{pmatrix} \epsilon_{1i} \\ \epsilon_{2i} \end{pmatrix} \sim N \begin{pmatrix} \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & 1 \end{pmatrix} \end{pmatrix}. \tag{8}$$

We specify Heckman's two-step approach as:

$$E_i = \mathbb{1}\{\gamma_1 Z_{E,i} + \gamma_2 M_i + X_i^{'}\gamma_3 + \theta_{1s} + \phi_{1t} + \xi_i > 0\} = \mathbb{1}\{Z_i^{'}\gamma + \xi_i > 0\}, \tag{9}$$
$$y_i = \alpha M_i + X_i^{'}\beta + \theta_{2s} + \phi_{2t} + \sigma_{12}\lambda(Z_i^{'}\gamma) + \epsilon_i, \tag{10}$$

We start by estimating Equation 9 using a probit. Then, we use the estimated coefficients to compute $\lambda(Z_i^{'}\gamma) = \phi(Z_i^{'}\gamma)/\Phi(Z_i^{'}\gamma)$. We estimate Equation 10 by OLS. Given that $\lambda(Z_i^{'}\gamma)$ is constructed using estimated values of γ, we bootstrap to obtain the standard errors.

In both specifications, we impose that the exclusion restriction $Z_{E,i}$ appears in the employment equation (Equations 6 and 9) but not in the wage equation (Equations 7 and 10). The exclusion restriction we use is the age of the youngest own child in the household. Specifically, the exclusion restriction is composed of a series of dummy variables for the age of the youngest own child in the household: one dummy for each age from 1 to 17, with under 1 as the base category, a dummy for children above 18, and a dummy for no children. The rational of this choice is as follows. Given the extensive literature on the motherhood penalty and the fatherhood premium coupled with the positive correlation between marriage and having children, it is important to control for children when estimating the relationship between wages and marriage.[15] A particularly common exclusion restriction in the literature is to use a dummy variable for the presence of own children in the household.[16] However, this option is not compatible with controlling

[15] See Angelov, Johansson, and Lindahl (2016), Chung, Downs, Sandler, and Sienkiewicz (2017), Killewald (2013), Kleven, Landais, and Sgaard (2019), or Kuziemko, Pan, Shen, and Washington (2018).

[16] We experimented with this exclusion restriction while not controlling for children in the wage equation. In line with the existence of the motherhood penalty and the fatherhood premium, we find a lower MWP for women and a higher MWP for men.

for children in the wage equation. Moreover, intuitively, if we think about the constraints that affect the employment decisions of women, it seems clear that the time a mother needs/wants to devote to children is decreasing with the age of the child. For example, a newborn requires more time, i.e., is more likely to affect the employment margin, than a teenager.

To the best of our knowledge, our paper is the first to use the age of the youngest own child as an exclusion restriction. We note two relevant points. First, the age of the youngest child dummies are jointly significant in the probit employment equation. Secondly, the set of controls (X_i) in the wage equation includes dummies for children aged 0-4 and 5-17. Hence, because we already control for the presence of children, we think it is safe to exclude $Z_{E,i}$ from the wage equation. The implicit assumption is that what affects wages is whether there are young (0-4) and/or older (5-17) children in the household but not the age of the youngest, which is only relevant for the employment decision. In the following section we present the results of correcting for sample selection using the two specifications we consider.

4.2.2 Results

Table 4: Heckman Two-Step and Maximum Likelihood Models, CPS
Dependent Variable: Log(Hourly Wage) in 1999 Dollars

	Heckman Two-Step		**Full Maximum Likelihood**	
	1977-1992	**2003-2018**	**1977-1992**	**2003-2018**
Married	-0.003	0.073***	-0.004	0.074***
	(0.005)	(0.005)	(0.005)	(0.005)
Inverse Mills Ratio	-0.070***	0.072***	-0.037***	0.024***
	(0.019)	(0.025)	(0.012)	(0.007)
Adjusted R^2	0.189	0.227		
Observations	204,261	245,038	204,261	245,038

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. Standard errors are reported in parentheses, where these are clustered by state. The dependent variable in all columns is the natural log of wages. Year and state fixed effects are included in all regressions. The following additional controls are included: dummies for highest level of educational attainment, potential experience, potential experience squared, dummies for number of children below the age of 5, dummies for number of children aged 5-17. In columns 1 and 2, we implement Heckman's two-step method. In this case we bootstrap standard errors, allowing for clustering at the state level, and using 200 iterations. In columns 3 and 4 we estimate both stages jointly via Maximum Likelihood. The exclusion restrictions are a series of dummies for age of youngest child in the household, where age less that 1 is the base category, and a dummy for 18 and over and no children are also included.

Table 4 presents the estimated married coefficients (α in Equations 7 and 10) along with the Inverse Mills Ratio associated with the first stage (Equations 6 and 9) for the two periods we consider in the CPS sample. The results in Table 4 point out two relevant implications. First, the married coefficients estimated by OLS in Table 3 are robust to correcting by selection into employment. The OLS married coefficient for 1977-1992 is -0.006 while the selection-corrected coefficient is -0.003 when computed with Heckman's two-step procedure and -0.004 when computed with maximum likelihood. None of the coefficients is statistically different from zero. For 2003-2018, the OLS estimate is 0.074 while selection-corrected coefficients are 0.073 and 0.074. All coefficients are statistically significant at the 1% level. That is, selection-into-employment plays a negligible role in the determination of the returns to marriage of women. Secondly, qualitatively, the selection patterns in our data are in line with those described by Mulligan and

Rubinstein (2008) although they focus on full-time and full-year workers while our definition of employment includes more work arrangements and we use a different exclusion restriction. In particular, our two specifications indicate that in the 1977-1992 sample period, women are negatively selected into employment while they are positively selected into employment in the 2003-2018 period.

4.3 A Fixed Effects Framework

We now turn to the NLSY79 sample in order to exploit within-individual variation to estimate the returns to marriage. For the purpose of maximizing comparability between the estimates in sections 4.1, 4.2, and 4.4, which use the CPS sample, and those presented in this section, we report the estimates of a Fixed Effects (FE) model along with Pooled OLS. However, a caveat is in order. As discussed in Section 2.3, the requirements of the methodologies we use for the CPS and NLSY79 imply that the two samples differ in some relevant descriptive statistics. Hence, the differences in the coefficients from the Pooled OLS regressions using the NLSY79 sample are not directly comparable with those obtained from the CPS data in sections 4.1, 4.2, and 4.4. Comparable coefficients estimates can be found in Table C2 of Appendix C, where we reproduce the NLYS79 sample using CPS data.

4.3.1 Empirical Specification

We estimate a FE specification of the form:

$$y_{it} = \alpha M_{it} + X_{it}^{'}\beta + \eta_i + \epsilon_{it}, \tag{11}$$

where, analogously to Equation 1, y_{it} is the natural logarithm of hourly wages, M_{it} is a dummy variable which takes value 1 when an individual reports to be married and living with their spouse, X_{it} is a vector of controls which includes dummies for levels of education, categories for the number of children, experience, tenure, and an urban residence indicator. The coefficient η_i is an individual-specific time-invariant fixed effect that may be correlated with M_{it} and X_{it}. The key assumption required to consistently estimate the coefficients α and β is that the covariates M_{it} and X_{it} are strictly exogenous. Formally this can be written as

$$E[\epsilon_{it}|M_{i1},\ldots,M_{iT},X_{i1},\ldots,X_{iT},\eta_i] = 0, \text{ for all } t = 1,2,\ldots,T. \tag{12}$$

For the marriage indicator, the strict exogeneity assumption implies that

$$E[M_{it}\epsilon_{is}] = 0, \text{ for all } s, \text{ and } t. \tag{13}$$

With the strict exogeneity assumption in hand, it is useful to take stock of the challenges we face in estimating the causal effect of marriage on wages, in terms both of endogeneity concerns and sample selection bias for women. Firstly, if the underlying source of the endogeneity of marriage is set of factors that are constant over time, then the use of fixed effects corrects for the influence of such time-invariant factors. However, the FE model is not able to estimate the *true* returns of marriage when these omitted factors change over time. Moreover, the FE specification cannot address the issue of simultaneity bias. As an example, if past wage fluctuations drive future marriage decisions, then we can see from Equation 13 that the assumption of strict exogeneity is not met.

Secondly, we do not explicitly consider selection-corrected panel data models.[17] However, we argue that the concern about sample-selection bias in our FE setup is bound to be minor. If the origin of the selection mechanism is constant over the sample period, then it is already captured by the individual fixed effects. In addition, in Section 4.2, we show that the extent to which sample-selection bias affects the coefficients obtained by the uncorrected OLS estimates is negligible.

4.3.2 Results

Table 5: Panel Data, NLSY1979

Dependent Variable: Log(Hourly Wage) in 2006 Dollars

	Men		**Women**	
	Pooled OLS	**Fixed Effects**	**Pooled OLS**	**Fixed Effects**
Married	0.123***	0.051***	0.072***	0.030**
	(0.018)	(0.012)	(0.019)	(0.013)
R-Squared	0.338	0.315	0.318	0.268
Observations	14,557	14,557	10,330	10,330

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. Standard errors are reported in parentheses, where these are clustered by individual. The dependent variable in all columns is the natural log of wages. Columns 1 and 3 present pooled OLS estimates, Columns 2 and present fixed effects estimates. The following controls are included: dummies for highest level of educational attainment, dummies for deciles of both actual experience and tenure, dummies for number of children and a dummy for urban residence.

In Table 5, we present the coefficient associated to marriage for both the FE model and the Pooled OLS. The Pooled OLS estimates indicate that marriage is associated to a 12.3% premium for men and a 7.2% premium for women.[18] The coefficients from the FE model are significantly smaller than those from the Pooled OLS estimation for both sexes. For men, the FE coefficient is 0.051 which is around 59% smaller than the Pooled OLS coefficient. Similarly, for women, the FE coefficient (0.030) is 58% smaller than that estimated in the Pooled OLS. That is, around three fifths of the observed correlation between marriage and wages can be accounted for individual fixed-effects.

Taken together with the results from the bounding exercise of Section 4.1, we interpret the coefficients of the FE model as evidence that, for both men and women, a sizable part of the correlation between marriage and wages is due to the omission in the canonical wage equation of relevant variables that are positively associated with both wages and the propensity to get married. At the same time, both the Oster bounds and the FE estimates indicate that there is a significant part of this correlation that survives the omitted-variable-bias corrections. Hence, we have established that there is a positive relationship between marriage and wages that cannot be

[17]See Dustmann and Rochina-Barrachina (2007) for a discussion of some of the alternative approaches to correct for selection in panel data models.

[18]In Table C2 of Appendix C.2 we present the results of the bounding exercise in Table 3 of Section 4.1 on the set of samples from CPS data matched to the NLSY79 sample. The magnitude of the OLS married coefficients is similar to the Pooled OLS. The interpretation of the Oster bounds is equivalent to the that of the fixed effects.

accounted for by omitted-variable bias nor sample-selection bias. In the following section we use an instrumental variable framework to shed further light on whether marriage has a causal effect on wages by using a methodology that jointly tackles omitted-variable bias, sample-selection bias, and the concern of reverse causality.

4.4 An Instrumental Variable Approach

4.4.1 The Instrument

In this section, we present a novel instrument to estimate the causal effect of marriage on wages. We think of the decision to marry as being not only a product of economic factors, marriage market conditions, preferences, and chance but also social norms. At the same time, we assume that the social norms that affect marriage decisions do not affect individual productivity and, thus, wages. To measure the prevalence of local social norms on the propensity to marry, we proceed as follows. For each individual in our sample, we compute the (CPS-weighted) share of married people of the same sex, who live in the same state, are observed in the same survey year, hold the same coarse level of education, and have children (or not) but are 6 to 15 years older.[19]

The intuitive idea is that, because social norms are persistent over time, the marriage patterns of older cohorts that have similar characteristics to the current cohort are a consequence of the social norms that are relevant for the marriage decisions of the current cohort. Therefore, the marriage rate of the older cohort is a proxy for the social norms that determine the propensity to marry of the current cohort.

4.4.2 Empirical Specification

For both men and women, we run the following two-stage least squares (2SLS) specification:

$$M_i = \pi_1 Z_{M,i} + X_i^{'}\pi_2 + \theta_{1s} + \phi_{1t} + \mu_i, \tag{14}$$

$$y_i = \alpha M_i + X_i^{'}\beta + \theta_{2s} + \phi_{2t} + \epsilon_i. \tag{15}$$

Equation 14 is the first stage which models marriage as dependent on the covariates used in previous specifications and the instrument $Z_{M,i}$. Equation 15 describes the second stage. It specifies how the logarithm of wages, y_i, depends on marriage and the same covariates as in Section 3.

For women, we also run a selection-corrected version of the 2SLS specification described in Equation 14 and Equation 15:

$$E_i = \mathbb{1}\{\kappa_1 Z_{E,i} + \kappa_2 Z_{M,i} + X_i^{'}\kappa_3 + \theta_{1s} + \phi_{1t} + \xi_i > 0\} = \mathbb{1}\{Z_i^{'}\kappa + \xi_i > 0\}, \tag{16}$$

$$M_i = \pi_1 Z_{M,i} + X_i^{'}\pi_2 + \theta_{2s} + \phi_{2t} + \pi_5\lambda(Z_i^{'}\kappa) + \mu_i, \tag{17}$$

$$y_i = \alpha M_i + X_i^{'}\beta + \theta_{3s} + \phi_{3t} + \sigma_{13}\lambda(Z_i^{'}\kappa) + \epsilon_i. \tag{18}$$

[19] When we define the reference cohort we balance two criteria. First, we require that the reference cohort is old enough to minimize competition in the (age-based) marriage market. Second, the reference cohort needs to be close enough to the individual in the sample so that the social norms that define the marriage decisions of the reference cohort affect the marriage decisions of the individuals in our sample. We match on education and the presence of children both because people are more likely to base decisions on those who are similar to themselves and also to proxy for homophilic social networks.

Because we treat marriage as an endogenous variable, differently from our approach in Equation 9 of Section 4.2, we use the instrument $Z_{M,i}$ instead of the dummy for marriage M_i in the employment equation (Equation 16). We start by estimating the employment decision (Equation 16) using a probit. Then, we recover the estimated coefficients to compute $\lambda(Z_i'\kappa) = \phi(Z_i'\kappa)/\Phi(Z_i'\kappa)$. Finally, we estimate the two systems of equations, Equations 14-15 and Equations 17-18 using 2SLS. We bootstrap the standard errors in the selection-corrected 2SLS procedure.[20]

We think the treatment of marriage may have heterogeneous effects and, thus, consider the coefficient estimates from the IV specifications a measurement of a local average treatment effect (LATE).[21] In the next section, we highlight the main features of the supporting evidence we provide in Appendix E for the assumptions that identify a well-defined LATE. First, we require that local social norms significantly affect marriage decisions (First stage). Second, we need that local social norms are (conditionally) randomly assigned across individuals (Conditional independence). Third, the impact of local social norms on marriage has to be monotonic (Monotonicity). Lastly, we require social norms impact wages only through the marriage channel (Exclusion restriction).

4.4.3 Support for the Identifying Assumptions

First Stage. We provide evidence supporting the relevance of the instrument in three places. Figure E1 shows the first stage graphically, as well as presenting information on the distribution of the instrument. For both men and women in both periods, there is evidence of a strong relationship between local social marital norms and individual marriage decisions (conditional on the other relevant covariates discussed in Section 4.4.2). In addition, the first column of Table E1 presents the first stage coefficient for each of the key sample-specification couplets. Finally, we present first-stage F-statistics at the base of the regression results in Table 6, presented below. All pieces of evidence provide strong support for the relevance of the instrument.

Conditional Independence. Table E2 examines the stability of the first stage parameter as we condition on an extra set of covariates. These variables are only available for a subset of the 2003-2018 time period, hence, we do include these in our main specification. They are, however, variables that can plausibly impact marriage decisions and productivity. To the extent that local social norms are conditionally randomly assigned, adding these variables to the first stage should not appreciably impact the point estimate on the instrument.[22]. There is no impact on the first stage coefficient of including these additional regressors, which we interpret as supportive evidence of the conditional independence assumption.

Monotonicity. Allowing for the possibility of heterogeneous treatment effects of marriage requires us to make the additional assumption of monotonicity. In this context, this means that any individual getting married when local social norms are weak also marries when they are strong. It also implies that individuals not marrying when social norms towards marriage are

[20] The instrument for marriage is estimated prior to running the 2SLS procedure. It should be noted that in the case of a generated instrument (which enters only the first stage and the selection equation), we do not need to adjust the standard errors of the 2SLS estimates as it is the case with a generated regressor in the wage equation.

[21] See Imbens and Angrist (1994).

[22] A selection on observables idea not too distant from the bounding exercise in Section 4.1.

strong does not marry when they are less pronounced. A growing literature on judge severity instruments (Dahl, Kostøl, and Mogstad (2014); Bhuller, Dahl, Løken, and Mogstad (Forthcoming); Bald, Chyn, Hastings, and Machelett (2019)), which shares the same IV setting that we do of binary endogenous regressor, continuous instrument, notes that monotonicity implies we should see a non-negative first stage coefficient for any sub-sample. Table E1 presents the first stage coefficient for a variety of different sub-samples. In all cases, the coefficient is non-negative, lending support for the monotonicity assumption.

4.4.4 Results

Table 6: IV and IV-Heckman Models, CPS
Dependent Variable: Log(Hourly Wage) in 1999 Dollars

	Men		**Women**			
	IV		**IV**		**IV-Heckman**	
	’77-’92	**’03-’18**	**’77-’92**	**’03-’18**	**’77-’92**	**’03-’18**
Married	0.266***	0.217***	-0.059*	0.084***	-0.048	0.087***
	(0.009)	(0.019)	(0.032)	(0.030)	(0.030)	(0.034)
Inverse Mills Ratio					-0.059***	0.087***
					(0.020)	(0.023)
First-Stage F-Statistic	2325.2	1111.6	416.9	460.4	421.0	465.0
Adjusted R^2	0.226	0.266	0.190	0.224	0.190	0.224
Observations	240,204	254,562	181,094	230,181	181,094	230,181

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. Standard errors are reported in parentheses, where these are clustered by state. The dependent variable in all columns is the natural log of wages. Year and state fixed effects are included in all regressions. The following additional controls are included: dummies for highest level of educational attainment, potential experience, potential experience squared, dummies for number of children below the age of 5, dummies for number of children aged 5-17. The instrument in all specification in the proportion of individuals' reference group that are married. Columns 1-4 present IV estimates, Columns 5 and 6 present selection-corrected IV estimates. In the latter case, the exclusion restrictions for the employment equation are a series of dummies for age of youngest child in the household, where age less that 1 is the base category, and a dummy for 18 and over and no children are also included. In this case we bootstrap standard errors, allowing for clustering at the state level, and using 200 iterations.

In Table 6, we present the estimates from the instrumental variable (IV) specification for men and women along with the its selection-corrected version for women. All the patterns discussed in the previous sections are confirmed by the results from specification Namely, the MWP for men is sizable and has remained fairly constant over the last decades. For women, the relationship between marriage and wages has evolved from a penalty to a considerable premium albeit smaller than that of men.

The coefficients estimated using the IV framework are of the same order of magnitude as the correlations computed in Section 3. However, both the Oster bounds from Section 4.1 and the estimates from the FE model in Section 4.3 show that a considerable fraction of the correlation between marriage and wages can be accounted for by omitted-variable bias while the IV estimates, which correct both for omitted-variable bias and simultaneity bias, are *larger* than the conditional correlations from Section 3. Assuming that both the Oster bounds and the FE estimates truthfully indicate the direction of the omitted-variable bias, two reasons can rationalize the larger coefficients for marriage from the IV estimates. First, the simultaneity bias implies that individuals with, everything else equal, higher wages (for reasons different to marriage) are less likely to get married. Hence, once we correct for simultaneity bias, the true effect of marriage on wages is *bigger* than that observed when estimating the conditional correlations from Equation 1. Note that for this to be the case, the *negative* simultaneity bias needs to be big enough to counterbalance the effect of the omitted-variable bias which works on the opposite direction.

Secondly, the true data generating process is one in which there are heterogeneous treatment effects for marriage. That is, because our IV estimates reflect the returns to marriage of the compliers, those for which the decision to marry is affected by the instrument, if the effects of marriage are heterogeneous, the effect computed out of the compliers is not necessarily equal to the average treatment effect. Hence, our interpretation of the IV estimates is that they reflect the causal effect of marriage on wages for a subgroup of the population with *high* returns to marriage.

4.4.5 The Compliers

The IV estimates represent the average causal effect of marriage on wages for the compliers. Whilst it is infeasible to identify the individual compliers, we can characterize the compliant sub-population by calculating its size and certain observable characteristics of this group. We detail the procedure we follow in Section E.2 of Appendix E. Our approach follows that of Dahl et al. (2014) as we face a similar situation of an endogenous binary variable and a continuous instrument.

We derive two main conclusions from analyzing the compliant population. First, as Table E3 shows, the share of compliers in the population is sizable. The different approaches to compute the share of complier population indicate that, for men in the 1977-1992 period compliers make out around 52% to 63% of the total population. In the case of men in 2003-2018, compliers represent 42% to 54% of the population. For women, in 1977-1992 the share is around 29% to 40% while it is about 20% to 28% in the 2003-2018 period. Secondly, the results on Table E4 indicate that, both for men and women in 1977-1992, the complier population tends to be younger and less educated than the whole population. Interestingly, this pattern is somewhat more nuanced in the 2003-2018, when the compliers are more still more likely to be younger but less so than in the previous period and are somewhat more educated.

4.5 The Heterogeneous Effect of Marriage

As discussed in Section 4.4.4, we interpret the fact that the IV-estimated coefficients are higher than the OLS coefficients as a consequence of marriage having heterogeneous returns. In section we discuss two set of results that support this idea.

In Appendix A, we compute unconditional quantile regressions (UQR), Oster bounds on the quantile estimates, and selection-corrected quantile estimates for women. The coefficients show that the relationship between marriage and wages is of a very different magnitude across the (unconditional) wage distribution. In particular, for both men and women in both sample periods, the MWP premium decreases along the wage distribution. For men, in Figure A1a, this pattern is particularly stark. Both in the 1977-1992 and 2003-2018 periods, the coefficient associated to marriage for the 10th quantile indicates that married men earn 35% to 39% higher wages than their single counterparts, while the coefficient is less than 10% at the top of the distribution. In the case of women, both in Figure A1b and Figure A1c, we observe a similar pattern albeit less pronounced. In the 1977-1992 period, when the average returns to marriage for women are close to zero, the bottom half of the wage distribution displays a positive premium, while the top half exhibits a penalty. In 2003-2018, when the average reflects a MWP for women, the marriage coefficient in the lower end of the wage distribution indicates that married women earn wages that are more than 10% higher than those of their single counterparts, while at the other side of the distribution marriage is associated with a premium of less than 5%.

In Appendix E.4, we use our instrument to compute marginal treatment effects (MTEs). The results show that, for both genders and in both periods, the marginal treatment effects are different than the average treatment effect. That is evidence of marriage having heterogeneous effects. Interestingly, the MTE curves tend to slope upwards in almost all cases (with the exception of men in the 1977-1992 period). This suggests negative selection on gains. That is, those with the lowest resistance to marriage also have the lowest gains, whereas gains are large for those with high resistance.

5 Testing the Statistical Discrimination Hypothesis

Several key papers in the literature on the MWP hypothesize that a mechanism behind the higher wages of married men is positive employer discrimination.[23] The idea is that marriage might be positively related to variables that are relevant for productivity but are hard to observe by employers while marital status is not.[24] However, to the best of our knowledge, there is no systematic test of this hypothesis in the literature. We do so in this section.

We use two key frameworks in the literature on Employer Learning and Statistical Discrimination (EL-SD). First, we adapt the public learning setup of Altonji and Pierret (2001) for education and race to the case of marriage. The main idea is that as the experience in the labor market of workers increase, the returns on easy-to-observe variables vis-à-vis the returns on hard-to-observe variables are informative of the existence of EL-SD. We also consider the asymmetric learning (or private learning) framework of Schönberg (2007) and Pinkston (2009) which is an extension of the setup of Altonji and Pierret (2001). This setting allows current employers to accrue superior (and private) information about workers that the rest of the market does not.[25] We adapt our NLSY79 sample in order to be able to extend these empirical strategies to the case of marriage.[26]

The main mechanism in the model of Altonji and Pierret (2001) can be described as follows. Employers value the productivity of workers. Some of the determinants of productivity are easily observable by employers while others are not. With out loss of generality, consider one easy-to-observe variable such as marital status (which might or might not affect productivity) and

[23] See, for example, Ginther and Zavodny (2001) and Antonovics and Town (2004).

[24] We acknowledge the fact that, in the US, it is illegal to formally discriminate in favor of married candidates/workers and that job applicants cannot be forced to disclose their martial status. However, the implicit assumption is that it comes at a low cost for employers to have a good approximation of the marital status of a job applicant or recently hired worker. Consider the content of casual conversations in the workplace, the fact that many individuals display their marital status through elements of clothing (such as wedding rings), or that if there exists positive discrimination towards married individuals it is optimal for these individuals to reveal this information to their (potential) employer.

[25] The conclusions we draw from the asymmetric learning framework are virtually identical to those from the public learning. We describe the asymmetric learning setup and its results in Section D.2 of Appendix D.

[26] Section D.1 of Appendix D details how we construct the sample.

a hard-to-observe variable such as cognitive ability/intelligence which determines productivity. Altonji and Pierret (2001) show that if these two variables are positively correlated, their returns in a wage equation indicate if there is employer learning and statistical discrimination. First, if the returns to the hard-to-observe variable increase with experience that is indicative of employer learning. The rationale is that, as the worker accumulates experience in the labor market, employers are better able to discern workers' true endowment of the hard-to-observe variable. Second, in the presence of statistical discrimination, that is, when the easy-to-observe variable is used to proxy the hard-to-observe variable, the returns on the easy-to-observe variable decrease with experience. The informational content of the easy-to-observe variable decreases and, hence, its returns diminish.

5.1 Empirical Specification

We consider the following specification:

$$y_i = \alpha_0 M_i + \alpha_1 (M_i \times x_i) + \beta_0 A_i + \beta_1 (A_i \times x_i) + C_i^{'} \gamma + \epsilon_i. \tag{19}$$

M_i is an indicator for being married. A_i is the the hard-to-observe determinant of productivity. As it is common in the EL-SD literature, we use the normalized and age-adjusted Armed Forces Qualification Test score (AFQT) from the NLSY as a measure of cognitive ability/intelligence. $(M_i \times x_i)$ and $(A_i \times x_i)$ are interactions between labor market experience (x_i) and, respectively, marriage (M_i) and the AFQT (A_i). The vector C_i contains a series of control variables. We follow the literature to include: highest educational attainment dummies, interactions of educational attainment dummies with time (to absorb changing returns to education), polynomials up to order three in time and in experience (to not conflate changes that occur over time with the experience interaction of focus), an urban residence indicator, children dummies, and tenure.

A common issue in this framework is the fact that cognitive ability might determine actual experience and bias the estimates which are interacted with this variable. We follow the common approach in the literature and instrument experience with potential experience.[27]

5.2 Results

Table 7 and Table 8 present the results for men and women, respectively. We report the coefficients from three different specifications, incrementally adding regressors, to show the impact of their inclusion on the estimated coefficients associated with marriage, the easy-to-observe variable, and AFQT, the hard-to-observe variable. In both tables, columns (1) to (3) display the estimated coefficients when we use potential experience as x_i in Equation 19. In columns (4) to (6) we report the estimates for the case in which we use actual experience instrumented with potential experience as x_i.[28] The presence of employer learning implies that the returns to the hard-to-observe variable increase with experience. That is, the interaction between experience and the AFQT has to be positive. Positive employer discrimination implies that the returns to the easy-to-observe variable are positive when the worker has no experience and decrease while the worker accumulates experience. Hence, the married coefficient needs to be positive while

[27]In the IV regression all instances of actual experience (polynomials and interaction terms) are instrumented.
[28]To be consistent with the literature, we divide years of experience by 10.

the interaction between marriage and experience has to be negative.

Table 7: Testing SD-EL Model of Public Learning, Men - NLSY79
Dependent Variable: Log(Hourly Wage) in 2006 Dollars

	OLS Potential Experience			IV Actual Experience		
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Married	0.209***	0.146***	0.147***	0.200***	0.110***	0.113***
	(0.026)	(0.037)	(0.037)	(0.027)	(0.037)	(0.036)
Married*Experience/10		0.049**	0.049**		0.089***	0.090***
		(0.025)	(0.024)		(0.032)	(0.032)
AFQT	0.077***	0.077***	0.038*	0.073***	0.074***	0.024
	(0.013)	(0.013)	(0.021)	(0.013)	(0.013)	(0.021)
AFQT*Experience/10			0.030**			0.047***
			(0.013)			(0.017)
Adjusted R^2	0.328	0.329	0.330	0.323	0.325	0.326
Observations	8,271	8,271	8,271	8,271	8,271	8,271

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. Standard errors are reported in parentheses, where these are clustered by individual. The dependent variable in all columns is the natural log of wages. In addition to a dummy for married and the normalized AFQT score, the following additional control variables are included in all specifications: dummies for highest level of educational attainment, the education dummies interacted with a linear time trend, tenure, dummies for number of children, a dummy for urban residence and polynomials up to order 3 in both time and experience. Columns 2, 3, 5 and 6 include an interaction between the married dummy and experience/10. Columns 3 and 6 include an interaction between normalized AFQT and experience/10. Results from a pooled OLS model with experience captured by potential experience are presented in Columns 1-3. Results from an IV model where all experience terms are actual experience instrumented by potential experience are presented in Columns 4-6.

For men, the patterns of the OLS and IV estimates are almost identical. In columns (1) and (4), when we regress wages only on marriage, AFQT, and controls (without the experience interactions), the marriage and the AFQT coefficients are positive and statistically significant. Marriage is associated with a premium of about 20% (0.209 in column (1) and 0.200 in column (2)) while an increase of the AFQT of one standard deviation from the mean is associated with around a 7% higher wage (0.077 in the column (1) and 0.073 in column (4)). The inclusion of the interaction between marriage and experience in the regression, columns (2) and (5), reveals that the marriage premium of men increases over the working life. According to the IV estimates in column (5), the marriage premium of men with no experience is of around 11% (about 15% in the OLS estimates) while each additional decade of experience is associated with an increase of the premium of around 9 percentage points (about 5 for the OLS case). The coefficients from columns (3) and (6) indicate that the returns to a cognitive ability also evolve over the working life. In particular, both the OLS and IV coefficients suggest that, for men with no experience, a higher AFQT is associated with higher wages albeit the estimates are not precisely estimated. The p-value in the OLS case is 0.075 (the point estimate is 0.038) while it is 0.244 for the IV (the coefficient is 0.024). It is clearer in both cases that the returns to cognitive ability increase with experience. The OLS point estimate in column (3) is 0.030 with a p-value of 0.024 while its IV analog is 0.047 with a p-value of 0.006. The fact that the returns to cognitive ability, the hard-to-observe variable, increase with experience are consistent with the presence of employer learning. However, there is no evidence of employer discrimination as the returns to marriage also increase with experience instead of decreasing.

Table 8: Testing SD-EL Model of Public Learning, Women - NLSY79
Dependent Variable: Log(Hourly Wage) in 2006 Dollars

	OLS Potential Experience			**IV Actual Experience**		
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Married	0.027	-0.091**	-0.092**	0.005	-0.082**	-0.082**
	(0.030)	(0.037)	(0.037)	(0.032)	(0.035)	(0.035)
Married*Experience/10		0.104***	0.105***		0.095***	0.097***
		(0.030)	(0.030)		(0.036)	(0.036)
AFQT	0.093***	0.094***	0.077***	0.082***	0.082***	0.072***
	(0.013)	(0.013)	(0.020)	(0.013)	(0.013)	(0.018)
AFQT*Experience/10			0.015			0.012
			(0.017)			(0.020)
Adjusted R^2	0.260	0.263	0.263	0.295	0.296	0.297
Observations	6,899	6,899	6,899	6,899	6,899	6,899

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. Standard errors are reported in parentheses, where these are clustered by individual. The dependent variable in all columns is the natural log of wages. In addition to a dummy for married and the normalized AFQT score, the following additional control variables are included in all specifications: dummies for highest level of educational attainment, the education dummies interacted with a linear time trend, tenure, dummies for number of children, a dummy for urban residence and polynomials up to order 3 in both time and experience. Columns 2, 3, 5 and 6 include an interaction between the married dummy and experience/10. Columns 3 and 6 include an interaction between normalized AFQT and experience/10. Results from a pooled OLS model with experience captured by potential experience are presented in Columns 1-3. Results from an IV model where all experience terms are actual experience instrumented by potential experience are presented in Columns 4-6.

In the case of women, the interpretation of the results is also equivalent between the OLS and IV estimates. In columns (1) and (4), the AFQT coefficient is positive and significant while the married coefficient is not precisely estimated and *low*. The explanation for these low and imprecise estimates of the marriage premium for women is found in columns (2) and (5). When we include the interaction between marriage and experience, we see that the *average* premium from columns (1) and (4) is, in fact, a composite of a penalty for married women with no labor market experience which evolves into a premium when experience increases. In particular, married women without experience earn around 8-9% less than their single counterparts while an extra decade of experience increases their wages by around 10 percentage points. Differently from men, the returns to cognitive ability for women do not change significantly over the working life. The intercept of the returns to cognitive ability is positive and significant (0.077 in the OLS and 0.072 in the IV) while the interaction between the AFQT score and experience has a coefficient that is not statistically different from zero. In the OLS, the point estimate is 0.015 with a p-value of 0.391 while, in the IV case, the point estimate is 0.012 with a p-value of 0.564.

The fact that there is not a precise positive estimate for the interaction between experience and the AFQT, the hard-to-observe variable, implies that the evidence towards the presence of employer learning for women is rather weak. Moreover, the coefficient associated with the interaction between marriage and experience is significant and positive which is at odds with the existence of any type of positive employer discrimination that rationalizes a wage premium for married women. Nevertheless, it is relevant that the marriage wage premium of women is the reflection of an initial penalty that turns into a premium. In particular, this pattern can be consistent with the presence of statistical discrimination based on traditional gender roles within (married) households. The idea is that, when employers observe a married female worker with no experience, they use marriage to proxy unobservables such us attachment to the labor force or willingness to work long hours that might be negatively related with the stereotypical role of a married woman. As the labor market experience of married women increases, the true values of those characteristics become less difficult to observe and the penalty disappears. We see this mechanism as speculative, especially because it reflects a prior that should not survive, but indicative of how a marriage wage premium for women can coexist with wage penalties based on traditional gender roles.

6 Conclusions

We draw three main conclusions in this paper. First, there is evidence that marriage has a positive causal effect on wages for both men and women. We address the three main challenges to measure the returns to marriage (sample-selection bias, omitted-variable bias, and simultaneity bias) from different angles. Both the Oster (2019) bounds on cross-sectional data and the fixed-effects model on panel data indicate that although omitted variables can account for a sizable part of the observed relationship between marriage and wages, a significant positive association remains. We use a new exclusion restriction to tackle sample-selection bias and show that our results are robust to this correction. We present a new variable to instrument marriage in the wage equation based on local social norms. When we use this approach, we find *higher* returns to marriage than the OLS-estimated positive effects for both men and women. We argue that the higher coefficients are a consequence of marriage having heterogeneous returns among the population and the instrument impacting a set of compliers at the higher end of the treatment effect distribution. Without exogenous variation in marriage, we cannot address the concern of simultaneity bias for the whole population of individuals we consider. However, we show that the fraction of compliers in our setup is likely to be sizable. We also show that the compliers are likely to be younger and slightly less educated than the whole population. Our interpretation of the results is that the effect of marriage on wages is higher for individuals with lower earnings potential. We provide further support to this interpretation. We estimate, bound, and correct for selection the effect of marriage on wages along the wage distribution and show that the marriage premium is much higher at the lower end of the distribution.

Second, we find that the two main hypotheses that can be found in the literature to reconcile the existence of a marriage wage premium for men have little support in the data. We show that for both men and women, the returns to marriage increase with labor market experience which is inconsistent with the idea that the marriage wage premium is due to positive discrimination by employers. The household-specialization hypothesis is at odds with the sizable wage premium for married women and the fact that the premium for women mainly emerges has labor market experience increases.

Thirdly, we highlight that the relationship between wages and marriage is a relevant dimension of the social transformation that has occurred in the last decades both in terms of the family and the role of women in the labor market. The fact that, nowadays, married people have higher wages that their single counterparts is relevant not only to understand secular trends but also to think about policies. From taxation to welfare programs, a wide range of policies depend explicitly or implicitly on marital status. We need to better understand the relationship between marriage and wages, especially if this relationship is heterogeneous among individuals, to design more effective policies.

References

Avner Ahituv and Robert I. Lerman. How do marital status, work effort, and wage rates interact? *Demography*, 44(3):623–647, August 2007.

Stefania Albanesi and Claudia Olivetti. Home production, market production and the gender wage gap: Incentives and expectations. *Review of Economic Dynamics*, 12(1):80–107, January 2009.

Joseph G. Altonji and Charles R. Pierret. Employer learning and statistical discrimination. *The Quarterly Journal of Economics*, 116(1):313–350, February 2001.

Joseph G. Altonji, Todd E. Elder, and Christopher R. Taber. Selection on observed and unobserved variables: Assessing the effectiveness of catholic schools. *Journal of Political Economy*, 113(1):151–184, February 2005.

Nikolay Angelov, Per Johansson, and Erica Lindahl. Parenthood and the gender gap in pay. *Journal of Labor Economics*, 34(3):545–579, July 2016.

Kate Antonovics and Robert Town. Are all the good men married? uncovering the sources of the marital wage premium. *American Economic Review*, 94(2):317–321, May 2004. Papers and Proceedings of the One Hundred Sixteenth Annual Meeting of the American Economic Association.

Peter Arcidiacono, Patrick Bayer, and Aurel Hizmo. Beyond signaling and human capital: Education and the revelation of ability. *American Economic Journal: Applied Economics*, 2 (4):76–104, October 2010.

Orazio Attanasio, Hamish Low, and Virginia Sánchez-Marcos. Explaining changes in female labor supply in a life-cycle model. *American Economic Review*, 98(4):1517–1552, September 2008.

Anthony Bald, Eric Chyn, Justine S. Hastings, and Margarita Machelett. The causal impact of removing children from abusive and neglectful homes. Working Paper 25419, National Bureau of Economic Research, January 2019.

Manudeep Bhuller, Gordon Dahl, Katrine Løken, and Magne Mogstad. Incarceration, recidivism and employment. *Journal of Political Economy*, Forthcoming.

Francine D. Blau and Lawrence M. Kahn. Changes in the Labor Supply Behavior of Married Women: 1980-2000. *Journal of Labor Economics*, 25(3):393–438, July 2007.

Francine D. Blau and Lawrence M. Kahn. The gender wage gap: Extent, trends, and explanations. *Journal of Economic Literature*, 55(3):789–865, September 2017.

Christian N. Brinch, Magne Mogstad, and Matthew Wiswall. Beyond late with a discrete instrument. *Journal of Political Economy*, 125(4):985–1039, August 2017.

Sekyu Choi and Arnau Valladares-Esteban. The marriage unemployment gap. *The B.E. Journal of Macroeconomics*, 18(1), January 2018.

YoonKyung Chung, Barbara Downs, Danielle H. Sandler, and Robert Sienkiewicz. The parental gender earnings gap in the united states. Working Papers 17-68, Center for Economic Studies, U.S. Census Bureau, November 2017.

Christopher Cornwell and Peter Rupert. Unobservable individual effects, marriage, and the earnings of young men. *Economic Inquiry*, 35(2):285–294, April 1997.

Gordon B. Dahl, Andreas Ravndal Kostøl, and Magne Mogstad. Family welfare cultures. *The Quarterly Journal of Economics*, 129(4):1711–1752, November 2014.

Christian Dustmann and María Engracia Rochina-Barrachina. Selection correction in panel data models: An application to the estimation of females' wage equations. *The Econometrics Journal*, 10(2):263–293, June 2007.

Raquel Fernández. Cultural change as learning: The evolution of female labor force participation over a century. *American Economic Review*, 103(1):472–500, February 2013.

Sergio Firpo, Nicole M. Fortin, and Thomas Lemieux. Unconditional quantile regressions. *Econometrica*, 77(3):953–973, May 2009.

Sara Flood, Miriam King, Steven Ruggles, and J. Robert Warren. Integrated public use microdata series, current population survey: Version 4.0. [dataset]. *Minneapolis: University of Minnesota*, http://doi.org/10.18128/D030.V4.0., 2015.

Donna K. Ginther and Madeline Zavodny. Is the male marriage premium due to selection? the effect of shotgun weddings on the return to marriage. *Journal of Population Economics*, 14 (2):313–328, June 2001.

Claudia Goldin. A grand gender convergence: Its last chapter. *American Economic Review*, 104 (4):1091–1119, April 2014.

Jeffrey S. Gray. The Fall in Men's Return to Marriage: Declining Productivity Effects or Changing Selection? *The Journal of Human Resources*, 32(3):481–504, June 1997.

Jeremy Greenwood, Nezih Guner, and Guillaume Vandenbroucke. Family economics writ large. *Journal of Economic Literature*, 55(4):1346–1434, December 2017.

Nezih Guner, Yuliya Kulikova, and Joan Llull. Marriage and health: Selection, protection, and assortative mating. *European Economic Review*, 104:138–166, May 2018.

Jens Hainmueller. Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies. *Political Analysis*, 20(1):25–46, December 2012.

James Heckman. Sample selection bias as a specification error. *Econometrica*, 47(1):153–161, January 1979.

James J. Heckman and Edward J. Vytlacil. Econometric Evaluation of Social Programs, Part II: Using the Marginal Treatment Effect to Organize Alternative Econometric Estimators to Evaluate Social Programs, and to Forecast their Effects in New Environments. In J.J. Heckman and E.E. Leamer, editors, *Handbook of Econometrics*, volume 6, chapter 71. Elsevier, January 2007.

Martha S. Hill. The wage effects of marital status and children. *The Journal of Human Resources*, 14(4):579–594, October 1979.

Guido W. Imbens and Joshua D. Angrist. Identification and estimation of local average treatment effects. *Econometrica*, 62(2):467–475, March 1994.

Chinhui Juhn and Kristin McCue. Evolution of the marriage earnings gap for women. *American Economic Review*, 106(5):252–56, May 2016.

Alexandra Killewald. A reconsideration of the fatherhood premium: Marriage, coresidence, biology, and fathers' wages. *American Sociological Review*, 78(1):96–116, February 2013.

Alexandra Killewald and Margaret Gough. Does specialization explain marriage penalties and premiums? *American Sociological Review*, 78(3):477–502, April 2013.

Alexandra Killewald and Ian Lundberg. New evidence against a causal marriage wage premium. *Demography*, 54(3):1007–1028, June 2017.

Henrik Kleven. The EITC and the Extensive Margin: A Reappraisal. Mimeo, September 2019.

Henrik Kleven, Camille Landais, and Jakob Egholt Sgaard. Children and Gender Inequality: Evidence from Denmark. *American Economic Journal: Applied Economics*, 11(4):181–209, October 2019.

Sanders Korenman and David Neumark. Does marriage really make men more productive? *The Journal of Human Resources*, 26(2):282–307, April 1991.

Sanders Korenman and David Neumark. Marriage, motherhood, and wages. *The Journal of Human Resources*, 27(2):233–255, April 1992.

Harry A. Krashinsky. Do marital status and computer usage really change the wage structure? *The Journal of Human Resources*, 39(3):774–791, July 2004.

Ilyana Kuziemko, Jessica Pan, Jenny Shen, and Ebonya Washington. The mommy effect: Do women anticipate the employment effects of motherhood? Working Paper 24740, National Bureau of Economic Research, June 2018.

Eng Seng Loh. Productivity differences and the marriage wage premium for white males. *The Journal of Human Resources*, 31(3):566–589, July 1996.

Volker Ludwig and Josef Brüderl. Is there a male marital wage premium? new evidence from the united states. *American Sociological Review*, 83(4):744–770, August 2018.

Shelly Lundberg and Robert A. Pollak. The american family and family economics. *Journal of Economic Perspectives*, 21(2):3–26, June 2007.

Esfandiar Maasoumi, Daniel L. Millimet, and Dipanwita Sarkar. Who benefits from marriage? *Oxford Bulletin of Economics and Statistics*, 71(1):1–33, January 2009.

Casey B. Mulligan and Yona Rubinstein. Selection, investment, and women's relative wages over time. *The Quarterly Journal of Economics*, 123(3):1061–1110, August 2008.

Claudia Olivetti. Changes in women's aggregate hours of work: The role of returns to experience. *Review of Economic Dynamics*, 9(4):557–587, October 2006.

Emily Oster. Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business & Economic Statistics*, 37(2):187–204, 2019.

Laura Pilossoph and Shu Lin Wee. Household search and the marital wage premium. Mimeo, May 2019.

Joshua C. Pinkston. A model of asymmetric employer learning with testable implications. *The Review of Economic Studies*, 76(1):367–394, January 2009.

A. D. Roy. Some thoughts on the distribution of earnings. *Oxford Economic Papers*, 3(2): 135–146, June 1951.

Uta Schönberg. Testing for asymmetric employer learning. *Journal of Labor Economics*, 25(4): 651–691, October 2007.

Betsey Stevenson and Justin Wolfers. Marriage and divorce: Changes and their driving forces. *Journal of Economic Perspectives*, 21(2):27–52, June 2007.

Leslie S. Stratton. Examining the wage differential for married and cohabiting men. *Economic Inquiry*, 40(2):199–212, April 2002.

Appendix

A The Relationship between Marriage and Wages along the Wage Distribution

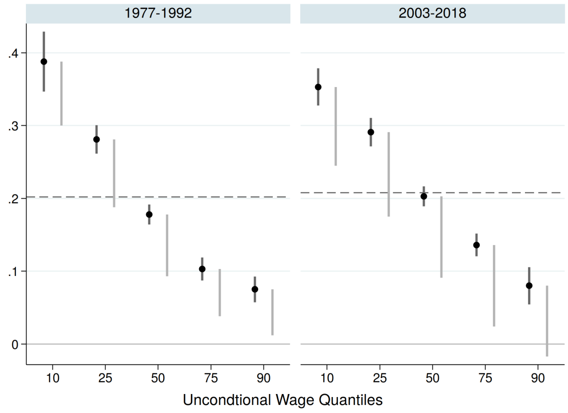
We use the Recentered Influence Function (RIF) approach of Firpo, Fortin, and Lemieux (2009) to study the relationship between marriage and wages across the (unconditional) wage distribution. Let us denote q_τ the τth quantile of the marginal (unconditional) distribution of the logarithm of hourly wages $F_y(y)$. The RIF can be written as:

$$\mathrm{RIF}(y; q_\tau, F_y) = q_\tau + \frac{(\tau - \mathbb{1}\{y \leq q_\tau\})}{f_y(q_\tau)}. \tag{20}$$

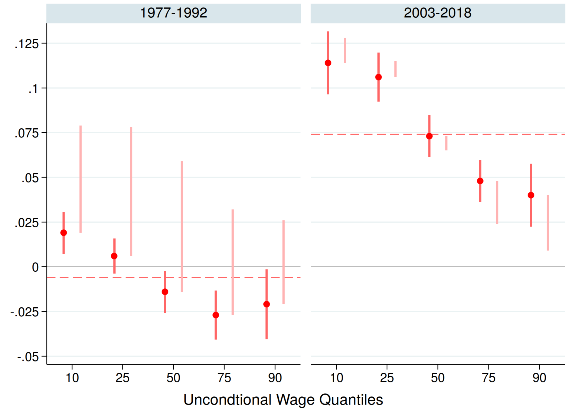
In order to compute the RIF, we estimate the relevant sample quantile (q_τ), then estimate the density $f_Y(q_\tau)$ at q_τ, and finally, construct the indicator dummy $\mathbb{1}\{y \leq q_\tau\}$. With this approach, we can estimate, using OLS, the following unconditional quantile regression (UQR) equation for various values of τ:

$$\mathrm{RIF}(y_i, q_\tau) = \alpha_\tau M_i + X_i^{'}\beta_\tau + \theta_{\tau s} + \phi_{\tau t} + \epsilon_{\tau i}. \tag{21}$$

In Figure A1 we present the coefficients estimated from the UQR in Equation 21 at the 10th, 25th, 50th, 75th, and 90th quintiles. Figures A1a and A1b present the point estimates, the associated 95% confidence intervals, and the Oster bounds, based on the UQR point estimates, computed as described in Section 4.1. Figure A1c combines the Heckman two-step selection correction described in Section 4.2 with the unconditional quantile regression specification. Table A1 presents all the the UQR estimates shown in Figure A1.



(a) Men



(b) Women

(c) RIF-Heckman's Two-Step

Figure A1: Uncondtional Quantile Regressions MWP Estimates

Notes: The figures plot $\hat{\alpha}_\tau$ from Equation 21 for $\tau = 10, 25, 50, 75$ and 90. The dependent variable in all cases is the natural log of wages. Year and state fixed effects are included in all regressions. The following additional controls are included: dummies for highest level of educational attainment, potential experience, potential experience squared, dummies for number of children below the age of 5, dummies for number of children aged 5-17. For each unconditional quantile, two lines are plotted. The left-hand line is the 95% confidence intervals (based on state-clustered standard errors), centered around $\hat{\alpha}_\tau$. The right-hand line represents the bounds on the UQR estimate accounting for selection on unobservables using the Oster (2019) method: the bounds are set assuming the coefficient of proportionality is zero or one. The dashed horizontal line is the OLS estimate, in order to provide a reference point for the UQR estimates.

Table A1: Mean and Unconditional Quantile Regressions, CPS

Dependent Variable: Log(Hourly Wage) in 1999 Dollars

	1977-1992						**2003-2018**					
	Mean	$\tau = 10$	$\tau = 25$	$\tau = 50$	$\tau = 75$	$\tau = 90$	**Mean**	$\tau = 10$	$\tau = 25$	$\tau = 50$	$\tau = 75$	$\tau = 90$
Men												
Married	0.202***	0.388***	0.281***	0.178***	0.103***	0.075***	0.208***	0.353***	0.291***	0.203***	0.136***	0.080***
	(0.005)	(0.021)	(0.010)	(0.007)	(0.008)	(0.009)	(0.005)	(0.013)	(0.010)	(0.007)	(0.008)	(0.013)
Oster Bounds	[0.124, 0.202]	[0.300, 0.388]	[0.188, 0.281]	[0.093, 0.178]	[0.038, 0.103]	[0.012, 0.075]	[0.102, 0.208]	[0.245, 0.353]	[0.175, 0.291]	[0.091, 0.203]	[0.024, 0.136]	[-0.017, 0.080]
δ Required for Coefficient of 0	1.869	1.985	1.925	1.678	1.418	1.150	1.573	1.704	1.686	1.512	1.165	0.854
Adjusted R^2	0.229	0.072	0.129	0.165	0.154	0.114	0.265	0.076	0.146	0.194	0.181	0.113
Observations	261,737	261,737	261,737	261,737	261,737	261,737	267,259	267,259	267,259	267,259	267,259	267,259
Women												
Married	-0.006	0.019***	0.006	-0.014**	-0.027***	-0.021**	0.074***	0.114***	0.106***	0.073***	0.048***	0.040***
	(0.005)	(0.006)	(0.005)	(0.006)	(0.007)	(0.010)	(0.005)	(0.009)	(0.007)	(0.006)	(0.006)	(0.009)
Oster Bounds	[-0.006, 0.056]	[0.019, 0.079]	[0.006, 0.078]	[-0.014, 0.059]	[-0.027, 0.032]	[-0.021, 0.026]	[0.067, 0.074]	[0.114, 0.128]	[0.106, 0.115]	[0.065, 0.073]	[0.024, 0.048]	[0.009, 0.040]
δ Required for Coefficient of 0	0.107	-0.348	-0.095	0.204	0.482	0.481	4.854	-163.655	28.353	4.715	1.836	1.255
Adjusted R^2	0.188	0.040	0.091	0.146	0.149	0.101	0.226	0.057	0.121	0.183	0.158	0.091
Observations	204,261	204,261	204,261	204,261	204,261	204,261	245,038	245,038	245,038	245,038	245,038	245,038
Women, Heckman's Two-Step Estimator												
Married	-0.003	0.029***	0.015***	-0.009	-0.029***	-0.027***	0.073***	0.115***	0.107***	0.070***	0.044***	0.037***
	(0.005)	(0.006)	(0.006)	(0.006)	(0.008)	(0.009)	(0.005)	(0.008)	(0.007)	(0.007)	(0.007)	(0.009)
Inverse Mills Ratio	-0.070***	-0.217***	-0.201***	-0.093***	0.049**	0.123***	0.072***	-0.071	0.012	0.126***	0.156***	0.145***
	(0.019)	(0.034)	(0.031)	(0.021)	(0.020)	(0.025)	(0.025)	(0.056)	(0.038)	(0.028)	(0.030)	(0.035)
Adjusted R^2	0.189	0.040	0.092	0.146	0.149	0.101	0.227	0.057	0.121	0.183	0.158	0.091
Observations	204,261	204,261	204,261	204,261	204,261	204,261	245,038	245,038	245,038	245,038	245,038	245,038

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. Standard errors are reported in parentheses, where these are clustered by state. The dependent variable in all columns is the natural log of wages. Year and state fixed effects are included in all regressions. The following additional controls are included: dummies for highest level of educational attainment, potential experience, potential experience squared, dummies for number of children below the age of 5, dummies for number of children aged 5-17. For both genders, estimates are presented in the following order: mean, 10%, 25%, 50%, 75%, 90% percentiles of the unconditional distribution of log wages. The unconditional quantile regression estimates are based on the RIF method of Firpo, Fortin and Lemieux (2009). For the Heckman's two-step method results, the exclusion restrictions are a series of dummies for age of youngest child in the household, where age less that 1 is the base category, and a dummy for 18 and over and no children are also included.

B Alternative Definition of the Non-married Group

As discussed in Section 2.1, the literature defines the MWP as the difference in wages between married individuals and those who are never married. The choice of the non-treated group as the never married raises the question of whether looking at the separated and the divorced, people who used to be treated, can shed any light on the MWP. Given that two of the main concerns when estimating the MWP are omitted-variable bias and simultaneity bias, if the termination of marriage is exogenous, the separated and divorced are the perfect control group to estimate the effect of marriage on wages. However, that is an assumption at odds with the large literature on the determinants of marriage dissolution.[29] Despite separation/divorce being endogenous, it is informative to estimate the returns to marriage with a different control group as a robustness check.

In this section, we present a series of estimates that are analogous to the main analysis we perform in Section 4 which use as the non-married group the separated and divorced. The only omission is the instrumental variable approach of Section 4.4. When we define the non-married group as the separated and divorced, the instrument is insufficiently correlated with marriage, that is, it is a weak instrument. This result is relevant because it is informative of the compliant sub-population in the main analysis and the feasibility of using our instrument in different contexts.

Table B1: OLS with Oster Bounds, CPS

Dependent Variable: Log(Hourly Wage) in 1999 Dollars

	Men		**Women**	
	1977-1992	**2003-2018**	**1977-1992**	**2003-2018**
Married	0.126***	0.141***	-0.016***	0.040***
	(0.005)	(0.006)	(0.004)	(0.003)
Oster Bounds	[0.099, 0.126]	[0.102, 0.141]	[-0.016, -0.014]	[0.017, 0.040]
δ Required for Coefficient of 0	3.685	3.128	10.191	1.721
Unadjusted R^2	0.209	0.233	0.182	0.224
$R_{\max}$	0.271	0.304	0.236	0.291
Adjusted R^2	0.209	0.233	0.182	0.224
Observations	246,595	243,354	213,483	246,769

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. Standard errors are reported in parentheses, where these are clustered by state. The dependent variable in all columns is the natural log of wages. Year and state fixed effects are included in all regressions. The following additional controls are included: dummies for highest level of educational attainment, potential experience, potential experience squared, dummies for number of children below the age of 5, dummies for number of children aged 5-17. In brackets we report bounds on the OLS estimate accounting for selection on unobservables using the Oster (2019) method: the bounds are set assuming the coefficient of proportionality is zero or one. Below the bounds we report the coefficient of proportionality (δ) that is required for the implied point estimate to be zero.

In Table B1, we reproduce the bounding exercise of Section 4.1. The results obtained with the alternative definition of the non-married group are qualitatively the same as in the main analysis. Namely, married men have a higher wage than their separated/divorced counterparts. This difference does not seem to be driven by unobservable variables as the estimated δ is

[29] See Stevenson and Wolfers (2007).

large. That is, assuming that the pattern through which unobservable variables affect the married coefficient is symmetric to that of observable variables, the role of unobservables in the wage equation needs to be more than three times larger than that of observables in order to drive the correlation between marriage and wages to zero. For women, we observe a significant penalty in the 1977-1992 period which evolves into a premium in the 2003-2018 as in the main exercise. Analogously to the case of men, unobservable variables need to have a bigger role than observables in the wage equation for the coefficient of marriage to be zero.

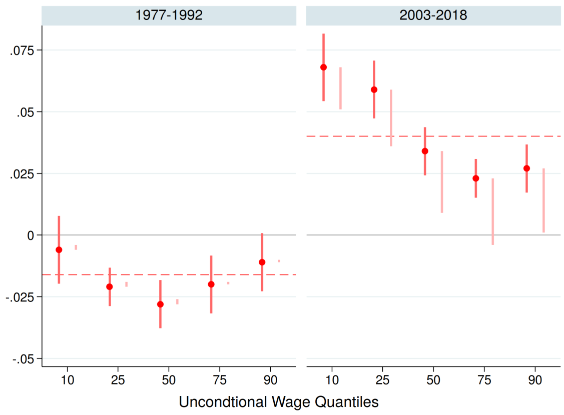
Table B2: Heckman Two-Step and Maximum Likelihood Models, CPS
Dependent Variable: Log(Hourly Wage) in 1999 Dollars

	Heckman Two-Step		**Full Maximum Likelihood**	
	1977-1992	**2003-2018**	**1977-1992**	**2003-2018**
Married	0.003	0.036***	-0.003	0.038***
	(0.006)	(0.004)	(0.006)	(0.003)
Inverse Mills Ratio	-0.101***	0.055**	-0.070***	0.022**
	(0.018)	(0.027)	(0.018)	(0.009)
Adjusted R^2	0.182	0.224		
Observations	213,483	246,769	213,483	246,769

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. Standard errors are reported in parentheses, where these are clustered by state. The dependent variable in all columns is the natural log of wages. Year and state fixed effects are included in all regressions. The following additional controls are included: dummies for highest level of educational attainment, potential experience, potential experience squared, dummies for number of children below the age of 5, dummies for number of children aged 5-17. In columns 1 and 2, we implement Heckman's two-step method. In this case we bootstrap standard errors, allowing for clustering at the state level, and using 200 iterations. In columns 3 and 4 we estimate both stages jointly via Maximum Likelihood. The exclusion restrictions are a series of dummies for age of youngest child in the household, where age less that 1 is the base category, and a dummy for 18 and over and no children are also included.

Table B2 presents the selection-correction of Section 4.2 using the alternative definition of the non-married group. Both the sign of the coefficients and the patters of selection captured by the inverse Mills ratio are analogous to those presented in the main text.

(a) Men



(b) Women

Figure B1: Uncondtional Quantile Regressions MWP Estimates

Notes: The figures plot $\hat{\alpha}_\tau$ from Equation 21 for $\tau = 10, 25, 50, 75$ and 90. The dependent variable in all cases is the natural log of wages. Year and state fixed effects are included in all regressions. The following additional controls are included: dummies for highest level of educational attainment, potential experience, potential experience squared, dummies for number of children below the age of 5, dummies for number of children aged 5-17. For each unconditional quantile, two lines are plotted. The left-hand line is the 95% confidence intervals (based on state-clustered standard errors), centered around $\hat{\alpha}_\tau$. The right-hand line represents the bounds on the UQR estimate accounting for selection on unobservables using the Oster (2019) method: the bounds are set assuming the coefficient of proportionality is zero or one. The dashed horizontal line is the OLS estimate, in order to provide a reference point for the UQR estimates.

In Figure B1, we replicate the analysis of the correlation between wages and marriage along the wage distribution of Appendix A. The observed patters are broadly similar to those observed in the main sample. That is, the MWP decreases as we move up the wage distribution for both sexes. The only difference with respect to Figure A1 is that for women in the 1977-1992 period, the higher two quintiles are slightly above the median.

Overall, the qualitative patterns of the main analysis are also present when we use the separated/divorced as the non-married group. Quantitatively, the married coefficients estimated with this alternative definition of the non-married are slightly smaller. This is consistent with the idea that part of the MWP is due to unobservables driving the decision to marry (which also affect productivity) and that the separated and the divorced might also affected by this mechanism. We interpret the results of this section as evidence that the inference derived in the main analysis is robust to the inclusion of the separated and the divorced in the non-married group. At the same time, we believe that excluding these individuals from the main analysis renders a cleaner exercise as the endogeneity concerns are more complex if the non-married group contains individuals that are affected by the marriage termination.

C Pseudo-NLSY CPS Samples

C.1 Creation of the Samples

In this section we describe the different matching approaches we use to construct samples that are based on the CPS data but are similar to the our NLSY79 sample in terms of descriptive statistics. We focus on three main dimensions for each gender: the survey years covered, the age distribution, and the average of several of the key covariates we use in the main analysis. Notice that, by matching on surveyed years and age distribution we mechanically tackle the issue that the NLSY data focuses on a particular birth cohort while the CPS contains many.

We construct four different pseudo-NLSY samples. In the first sample, which we name *Simple*, we focus only on selecting observations from the CPS that have the same characteristics as our NLSY79 sample in terms of surveyed year and the ages of the respondents. We start by computing the 10th and 90th percentiles of the distribution of surveyed years from the NLSY79 sample and find all the observations in the CPS data that fall within this time range. This is not a trivial exercise because the NLSY is not balanced across years. Then, for each NLSY survey year (annual from 1979 to 1993, biannual thereafter) we restrict the observations from the CPS to match the age range of the NLSY79 sample.

The second and third samples, build on the Simple pseudo-NLSY and add matching on covariates. That is, these samples also include the restrictions on survey years and age ranges. Both samples use propensity score matching to select CPS observations that match the NLSY79 sample using the following variables: a married dummy, dummies for highest educational attainment, number of children, and age. We use nearest neighbor match, allowing for ties (given he discrete natures of the matching variables), to construct the second pseudo-NLSY sample, which we call *NN(1).* For the third pseudo-NLSY sample, which we label *Kernel* we use kernel-matching methods (with the Epanechnikov kernel). For both samples, we store the matching weights and use these in the subsequent analysis instead of the CPS weights.

Lastly, we name the fourth pseudo-NLSY sample *Entropy Balance.* We take a similar approach to the second and third sample but we use the entropy-balancing method of Hainmueller (2012). The aim is to balance the first moments of the matching variables. We use the weights generated from this procedure in the subsequent analysis.

Table C1: Descriptive Statistics,Pseudo-NLSY79 based on CPS
Means, Standard Deviations in Parentheses

	Simple		**NN(1)**		**Kernel**		**Entropy Balance**	
	Men	Women	Men	Women	Men	Women	Men	Women
Sample Size	77,457	66,754	91,192	74,369	114,744	103,200	116,095	105,013
Weighted Sample Size			14,668	10,362	14,668	10,358	14,736	10,375
Married	0.650	0.711	0.465	0.462	0.482	0.484	0.459	0.460
Hourly Wage (1999 Dollars)	16.35	12.63	15.71	13.08	15.66	12.97	15.61	13.03
	(8.86)	(7.35)	(8.98)	(7.66)	(8.89)	(7.54)	(8.90)	(7.60)
Age	31.08	30.06	30.81	30.17	30.86	30.20	30.80	30.18
	(5.47)	(5.11)	(6.63)	(6.58)	(6.70)	(6.65)	(6.62)	(6.59)
Highest Level of Education:								
HS Dropout	0.082	0.053	0.075	0.022	0.078	0.024	0.079	0.022
HS Graduate	0.381	0.372	0.434	0.374	0.421	0.359	0.434	0.375
Some College	0.245	0.270	0.187	0.233	0.205	0.252	0.188	0.233
College Graduate	0.204	0.225	0.206	0.258	0.207	0.265	0.205	0.259
Advanced Graduate	0.088	0.080	0.097	0.113	0.090	0.100	0.094	0.111
Number Children, 0-4	0.391	0.356	0.217	0.164	0.236	0.175	0.226	0.164
	(0.665)	(0.619)	(0.497)	(0.428)	(0.521)	(0.441)	(0.511)	(0.427)
Number Children, 5-17	0.534	0.577	0.300	0.279	0.328	0.318	0.306	0.288
	(0.923)	(0.931)	(0.700)	(0.673)	(0.735)	(0.722)	(0.714)	(0.691)
Number Children, 0-17	0.925	0.933	0.517	0.443	0.564	0.493	0.532	0.452
	(1.14)	(1.09)	(0.883)	(0.806)	(0.920)	(0.852)	(0.901)	(0.823)

Table C1 presents summary statistics for the four pseudo-NLSY samples. Three differences stand out when comparing the descriptive statistics for our CPS sample (Table 1) with those of our NLSY79 sample (Table 2). In the CPS sample, individuals are older, the rate of married people is larger, and the education level is higher than in the NLSY79 sample. The descriptive statistics of Table C1 show that, except for our Simple pseudo-NLSY sample, these three differences are tackled with the matching all the procedures. In the following section, we reproduce our main analysis using the four pseudo-NLSY samples.

C.2 Key Results for the Pseudo-NLSY Samples

Table C2: OLS with Oster Bounds, CPS

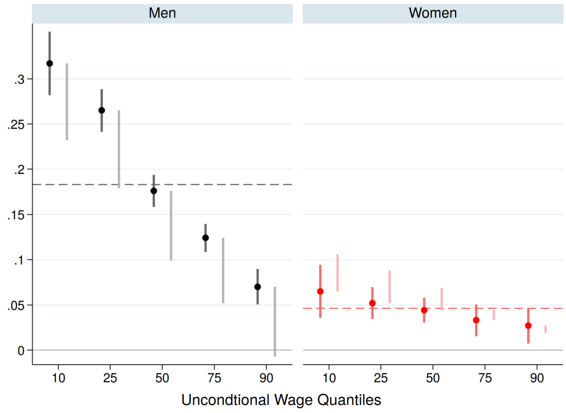
Dependent Variable: Log(Hourly Wage) in 1999 Dollars

	Simple		**NN(1)**		**Kernel**		**Entropy Balance**	
	Men	Women	Men	Women	Men	Women	Men	Women
Married	0.183***	0.046***	0.183***	0.056***	0.186***	0.052***	0.183***	0.049***
	(0.007)	(0.007)	(0.008)	(0.009)	(0.006)	(0.007)	(0.007)	(0.007)
Oster Bounds	[0.105, 0.183]	[0.046, 0.068]	[0.066, 0.183]	[0.017, 0.056]	[0.103, 0.186]	[0.043, 0.052]	[0.105, 0.183]	[0.044, 0.049]
δ Required for Coefficient of 0	1.675	-3.448	1.330	1.351	1.697	3.813	1.753	5.122
Unadjusted R^2	0.251	0.240	0.263	0.270	0.258	0.256	0.257	0.264
$R_{\max}$	0.326	0.312	0.343	0.350	0.335	0.333	0.334	0.343
Adjusted R^2	0.250	0.239	0.263	0.269	0.257	0.256	0.256	0.263
Observations	77,457	66,754	91,192	74,369	114,744	103,200	116,095	105,013

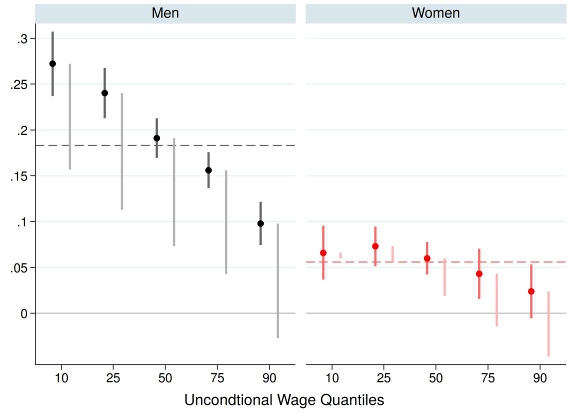
Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. Standard errors are reported in parentheses, where these are clustered by state. The dependent variable in all columns is the natural log of wages. The analysis uses four different datasets, the construction of which is detailed in section C.1. Year and state fixed effects are included in all regressions. The following additional controls are included: dummies for highest level of educational attainment, potential experience, potential experience squared, dummies for number of children below the age of 5, dummies for number of children aged 5-17. The following additional controls are included: dummies for highest level of educational attainment, potential experience, potential experience squared, dummies for number of children below the age of 5, dummies for number of children aged 5-17. In brackets we report bounds on the OLS estimate accounting for selection on unobservables using the Oster (2019) method: the bounds are set assuming the coefficient of proportionality is zero or one. Below the bounds we report the coefficient of proportionality (δ) that is required for the implied point estimate to be zero.

In Table C2 we replicate the bounding exercise in Table 3 of Section 4.1. Note that in these pseudo-NLSY samples, the survey years correspond to those in the NLSY79 sample, hence, we are not looking to analyze the two periods as we do in the CPS sample. The conceptually-comparable results are those of the FE model in Table 5 of Section 4.3. The results from Table C2 show that, for both genders, the married coefficient is similar across pseudo-NLSY samples. For men marriage is associated with around an 18% premium while married women earn roughly 5% more than their single counterparts. The magnitude of the marriage premiums is roughly similar to that obtained with Pooled OLS in Table 5 of Section 4.3. In the case of men, the Pooled OLS coefficient is 0.123, around 33% smaller than what we obtain in the pseudo-NLSY samples. For women, the married coefficient is 0.072 which is approximately 44% larger than the estimates from the pseudo-NLSY samples. The direction of the omitted-variable bias suggested by the bounds is in line with the FE model results. For men, in all four pseudo-NLSY samples, the bounds show that the married coefficient is lower when the effect of unobservables is symmetric to that of observables. That is the married coefficient from the OLS estimation is upward biased due to omitted-variable bias. In the case of women, the results are equivalent to those of men except for the Simple pseudo-NLSY sample, in which the direction of the bias is the opposite.

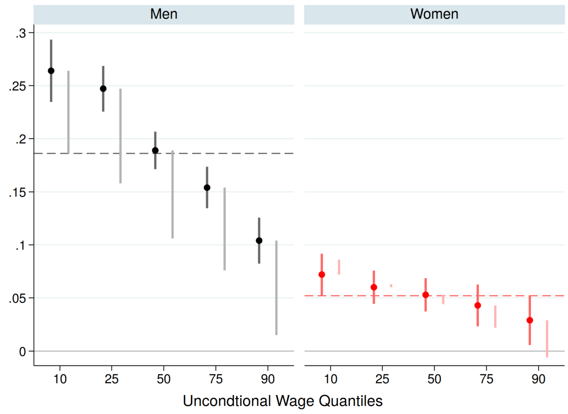
All in all, we interpret the results from Table C2 as confirming the inference we derive from the main analysis. That is, a sizable part of the observed relationship between marriage and wages can be accounted for by omitted-variable bias while there is an important portion of the relationship that survives the correction. Moreover, the coefficients in Table C2 show that part of the difference in the magnitude of the marriage wage premiums estimated by OLS from the CPS sample and the NLSY79 sample are due to the distinct observable characteristics between the individuals in each sample. In particular, the married coefficients we obtain from the CPS sample (Table 3) are somewhat larger than those from the NLYS79 sample (Table 5). The difference between the coefficients between from pseudo-NLSY samples and the NLSY79 sample are significantly smaller.



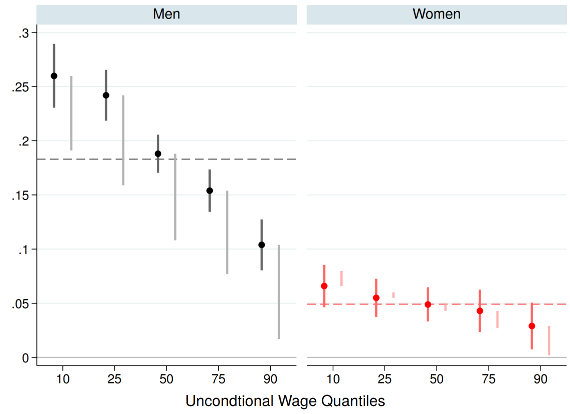
(a) Simple Restrictions



(b) 1 Nearest Neighbour



(c) Kernel



(d) Entropy Balance

Figure C1: Uncondtional Quantile Regressions MWP Estimates

Notes: The figures plot $\hat{\alpha}_\tau$ from Equation 21 for $\tau = 10, 25, 50, 75$ and 90. The dependent variable in all cases is the natural log of wages. Year and state fixed effects are included in all regressions. The following additional controls are included: dummies for highest level of educational attainment, potential experience, potential experience squared, dummies for number of children below the age of 5, dummies for number of children aged 5-17. For each unconditional quantile, two lines are plotted. The left-hand line is the 95% confidence intervals (based on state-clustered standard errors), centered around $\hat{\alpha}_\tau$. The right-hand line represents the bounds on the UQR estimate accounting for selection on unobservables using the Oster (2019) method: the bounds are set assuming the coefficient of proportionality is zero or one. The dashed horizontal line is the OLS estimate, in order to provide a reference point for the UQR estimates.

In Figure C1, we present the exercise from Appendix A applied to the four pseudo-NLSY samples. Broadly speaking, we also observe that the starkness of the relationship between marriage and wages declines along the wage distribution, as it is the case for the main CPS sample.

D More on Testing the Statistical Discrimination Hypothesis

D.1 Data

In order to test implications of the EL-SD models we consider, we modify our baseline NLSY79 sample. We select the sample restrictions to balance two objectives. First, we follow the EL-SD literature as much as possible so that our results are comparable to those in the literature. Secondly, we restrict the sample to account for the fact that marital status can change over time. Hence, we require that marital status is fixed within job-spell.

Broadly speaking, we follow the criteria laid out in Altonji and Pierret (2001), Pinkston (2009), and Arcidiacono, Bayer, and Hizmo (2010). Because we do not focus on education as the easy-to-observe variable, we do not impose further restrictions based on educational attainment. As in Pinkston (2009), we drop observations where the measure of actual experience exceeds potential experience by a year or more. For ever-married individuals, we consider marital status in each of their job-spells, and restrict the sample to job-spells where the ever-married enter the job married. As our focus is on employer learning and statistical discrimination based on marital status, spells that occur before marriage are not informative of the mechanism we test. We also aim to rule out cases where there is employee learning about the statistical discrimination process, if it exists, whereby individuals make marital decisions based on perceived employer-based statistical discrimination.

Table D1: Descriptive Statistics, NLSY79 EL-SD Sub-Sample
Means, Standard Deviations in Parentheses

	Men	**Women**
Sample Size	8,271	6,899
Number of Individuals	1,369	1,390
Married	0.639	0.732
Ever Observed Married in Panel	0.694	0.796
Hourly Wage (2006 Dollars)	17.38	12.96
	(9.57)	(6.45)
Job Tenure	3.44	3.28
	(3.69)	(3.75)
Experience	10.62	9.01
	(6.18)	(5.95)
Potential Experience	13.30	11.78
	(6.47)	(6.19)
Age	31.13	29.88
	(6.63)	(6.35)
Highest Level of Education:		
HS Dropout	0.122	0.050
HS Graduate	0.523	0.534
Some College	0.183	0.234
College Graduate	0.128	0.134
Advanced Graduate	0.045	0.049
	(1.05)	(1.04)
Normalized AFQT	-0.000	0.000
Urban Residence	0.719	0.724
Number Children, 0-17	0.889	0.915
	(1.00)	(1.00)

Table D1 presents the summary statistics of the sample we use to test the to EL-SD models we consider. The extra sample restrictions with respect to the baseline NLSY79 (Table 2) sample imply a considerable decrease in the number of observations. Notably, the sample does not contains less observations from single individuals are reflected by the higher marriage rate compared to that in Table 2. Otherwise, most statistics in Table D1 are broadly in line with their counterparts in Table 2.

D.2 Asymmetric Employer Learning

The model of asymmetric (or private) employer learning is an extension of the public learning framework. In particular, the models considers that there can exist two channels for the employer to improve their perception on the employee hard-to-observe characteristics. One, it the public channel already present in the base model. The second is that some information might be reveals over a tenure spell, which might only be available to the current employer. The main idea is that both the public learning mechanism, which is associated to experience in the labor market, and the private learning mechanism, which is related to tenure with a particular employer, may be relevant to understand EL-SD.

D.2.1 Empirical Specification

Formally, we extend the specification from Equation 19:

$$\begin{aligned} y_i = &\alpha_0 M_i + \alpha_1(M_i \times x_i) + \alpha_2(M_i \times t_i) \\ &+ \beta_0 A_i + \beta_1(A_i \times x_i) + \beta_2(A_i \times t_i) + C_i^{'}\gamma + \epsilon_i. \end{aligned} \tag{22}$$

We now include two interactions terms in tenure (t_i) in addition to those in experience (x_i). The vector C_i is also augmented to include polynomials of tenure up to order three to mirror our controls for experience.

Analogously to the concerns about the potential relationship between experience and productivity, tenure might also be correlated with unobserved productivity, thus biasing the tenure interaction terms. We adapt the approach in Pinkston (2009) to instrument for tenure. Specifically, we regress tenure in period t on actual experience, full duration of current tenure spell, and career-average tenure spells. The career-average tenure spells is a measure that encapsulates individuals' propensity to stay in a job over their (observed) careers and their general ability to enter well-matched jobs. In addition, the full duration of current job spell should capture firm-worker match-specific elements that may be correlated with the residual in the wage equation.[30] To the extent that these variables capture the channels through which tenure is correlated with the residual in Equation 22, we can use the residual from this regression as an instrument for tenure.

[30] The regression output is summarized as follows:

$$\begin{aligned} \text{Men: } t_i = &\underset{(0.080)}{-0.995} + \underset{(0.018)}{0.095}\ \bar{t}_i^{career} + \underset{(0.006)}{0.183}\ dur_i + \underset{(0.005)}{0.222}\ x_i + \hat{e}_i\ , \quad \text{Adjusted } R^2: 0.446 \\ \text{Women: } t_i = &\underset{(0.078)}{-0.878} + \underset{(0.019)}{0.069}\ \bar{t}_i^{career} + \underset{(0.007)}{0.193}\ dur_i + \underset{(0.006)}{0.244}\ x_i + \hat{e}_i\ , \quad \text{Adjusted } R^2: 0.484 \end{aligned}$$

D.2.2 Results

Table D2: Testing SD-EL Model of Public and Private Learning, Men - NLSY79
Dependent Variable: Log(Hourly Wage) in 2006 Dollars

	OLS Potential Experience, Actual Tenure			IV Actual Experience, Actual Tenure		
	(1)	(2)	(3)	(4)	(5)	(6)
Married	0.201***	0.150***	0.151***	0.197***	0.117***	0.120***
	(0.026)	(0.037)	(0.036)	(0.027)	(0.037)	(0.037)
Married*Experience/10		0.046	0.046*		0.099***	0.100***
		(0.028)	(0.028)		(0.037)	(0.037)
Married*Tenure		-0.003	-0.003		-0.007	-0.007
		(0.006)	(0.006)		(0.006)	(0.006)
AFQT	0.074***	0.074***	0.035*	0.075***	0.076***	0.030
	(0.013)	(0.013)	(0.021)	(0.014)	(0.014)	(0.021)
AFQT*Experience/10			0.034**			0.045**
			(0.015)			(0.019)
AFQT*Tenure			-0.001			-0.000
			(0.003)			(0.003)
Adjusted R^2	0.337	0.337	0.338	0.313	0.316	0.316
Observations	8,271	8,271	8,271	8,270	8,270	8,270

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. Standard errors are reported in parentheses, where these are clustered by individual. The dependent variable in all columns is the natural log of wages. In addition to a dummy for married and the normalized AFQT score, the following additional control variables are included in all specifications: dummies for highest level of educational attainment, the education dummies interacted with a linear time trend, dummies for number of children, a dummy for urban residence and polynomials up to order 3 in time, tenure and experience. Columns 2, 3, 5 and 6 include an two interaction terms between the married dummy and i. experience/10 and ii. tenure. Columns 3 and 6 include two interaction terms between normalized AFQT and i. experience/10 and ii. tenure. Results from a pooled OLS model with tenure captured by actual tenure and experience captured by potential experience are presented in Columns 1-3. Results from an IV model where all tenure terms are instrumented using the approach outlined in section D.2.1, and experience terms are actual experience instrumented by potential experience are presented in Columns 4-6.

Table D2 presents the results for men. The conclusions regarding public learning from Table 7 are robust to the inclusion of tenure. That is, the interaction between AFQT and experience in columns (3) and (6) remains positive and significant, indicating there exist public learning. The interaction between marriage and experience is also positive and significant, which indicates that there is no statistical discrimination based on marital status. Both the interaction between AFQT and tenure and the interaction between being married and tenure are statistically very close to 0. We interpret these coefficients as evidence that the private mechanism of employer learning is not relevant to explain the returns marriage.

Table D3: Testing SD-EL Model of Public and Private Learning, Women - NLSY79
Dependent Variable: Log(Hourly Wage) in 2006 Dollars

	OLS Potential Experience, Actual Tenure			**IV Actual Experience, Actual Tenure**		
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Married	0.022	-0.072**	-0.072**	-0.002	-0.101***	-0.100***
	(0.030)	(0.036)	(0.036)	(0.032)	(0.038)	(0.038)
Married*Experience/10		0.047	0.044		0.059	0.056
		(0.035)	(0.035)		(0.041)	(0.041)
Married*Tenure		0.013*	0.014**		0.014	0.015*
		(0.007)	(0.007)		(0.009)	(0.009)
AFQT	0.090***	0.091***	0.072***	0.082***	0.082***	0.084***
	(0.013)	(0.013)	(0.019)	(0.013)	(0.013)	(0.019)
AFQT*Experience/10			-0.005			-0.017
			(0.018)			(0.023)
AFQT*Tenure			0.007**			0.004
			(0.003)			(0.004)
Adjusted R^2	0.279	0.282	0.285	0.260	0.264	0.266
Observations	6,899	6,899	6,899	6,898	6,898	6,898

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. Standard errors are reported in parentheses, where these are clustered by individual. The dependent variable in all columns is the natural log of wages. In addition to a dummy for married and the normalized AFQT score, the following additional control variables are included in all specifications: dummies for highest level of educational attainment, the education dummies interacted with a linear time trend, dummies for number of children, a dummy for urban residence and polynomials up to order 3 in time, tenure and experience. Columns 2, 3, 5 and 6 include an two interaction terms between the married dummy and i. experience/10 and ii. tenure. Columns 3 and 6 include two interaction terms between normalized AFQT and i. experience/10 and ii. tenure. Results from a pooled OLS model with tenure captured by actual tenure and experience captured by potential experience are presented in Columns 1-3. Results from an IV model where all tenure terms are instrumented using the approach outlined in section D.2.1, and experience terms are actual experience instrumented by potential experience are presented in Columns 4-6.

Table D3 presents the estimates for women. The inclusion of tenure enriches the inference that we draw form Table 8 but do not modify the main conclusion. The interaction between AFQT and experience indicate that there is no evidence of public learning. The coefficients are slightly negative but not significant. In the OLS results, the estimate associated to the interaction between tenure and AFQT points towards the existence of private learning. However, this result is not robust to instrumenting tenure in column (6). The interaction between being married and experience and the interaction between being married and tenure confirm the composition of the MWP for women described in Section 5. Married women start their careers experiencing a wage penalty with respect to their single counterparts. As their career progresses, this penalty becomes a premium. This effects works both through the public and private learning channels albeit the coefficients associated to public learning (experience) are too noisy to be significant.

E More on the Instrumental Variables Approach

E.1 Assumptions

We set the assumptions within the potential outcomes framework. We denote $M_i(k)$ the potential marriage outcome if the value of the instrument $Z_{M,i}$ is equal to k, with $M_i = 1$ if the individual is married, and 0 otherwise. $y_i(M_i, Z_{M,i})$ is the potential wage outcome given M_i and $Z_{M,i}$. Formally, we specify the four IV assumptions as follows.

- First stage:
$$E[M_i(k) - M_i(k-1)] \neq 0 \quad \forall k. \tag{23}$$

- Independence:
$$[y_i(0), y_i(1), \{M_i(k); \forall k\}] \perp Z_{M,i}. \tag{24}$$

- Exclusion:
$$y_i(M_i, Z_{M,i}) = y_i(M_i) \text{ for } M_i = 0, 1. \tag{25}$$

- Monotonicity:
$$M_i(k) \geq M_i(k-1) \quad \forall k. \tag{26}$$

Table E1: 2SLS First Stage for Various Sub-Samples, CPS

		Predicted Married		Region				Age		College	
	Full Sample	**Below Median**	**Above Median**	**North-East**	**Mid-West**	**South**	**West**	**25-39**	**40-54**	**No**	**Yes**
Men, 1977-1992											
Proportion of	0.800***	0.084***	0.032***	0.792***	0.794***	0.752***	0.856***	0.796***	0.607***	0.841***	0.756***
RG Married	(0.017)	(0.019)	(0.010)	(0.018)	(0.011)	(0.030)	(0.044)	(0.016)	(0.064)	(0.017)	(0.028)
Observations	240,204	115,532	124,672	61,114	67,054	61,850	50,186	152,486	87,718	124,077	116,127
Men, 2003-2018											
Proportion of	0.652***	0.026	0.194***	0.645***	0.635***	0.646***	0.678***	0.644***	0.623***	0.562***	0.686***
RG Married	(0.020)	(0.025)	(0.020)	(0.041)	(0.022)	(0.022)	(0.077)	(0.021)	(0.049)	(0.018)	(0.024)
Observations	254,562	100,723	153,839	57,023	71,508	71,214	54,817	127,810	126,752	80,610	173,952
Women, 1977-1992											
Proportion of	0.573***	0.038	0.025***	0.592***	0.539***	0.499***	0.587***	0.579***	0.401***	0.597***	0.486***
RG Married	(0.028)	(0.023)	(0.007)	(0.057)	(0.021)	(0.042)	(0.109)	(0.033)	(0.060)	(0.035)	(0.033)
Observations	181,094	87,059	94,035	46,538	50,605	46,758	37,193	116,059	65,035	102,502	78,592
Women, 2003-2018											
Proportion of	0.429***	0.102***	0.125***	0.472***	0.416***	0.416***	0.390***	0.462***	0.311***	0.235***	0.418***
RG Married	(0.021)	(0.026)	(0.015)	(0.016)	(0.030)	(0.031)	(0.074)	(0.023)	(0.037)	(0.022)	(0.024)
Observations	230,181	96,617	133,564	53,784	66,707	63,776	45,914	116,265	113,916	55,585	174,596
Women, 1977-1992 Heckman's Two-Step											
Proportion of	0.585***	0.056**	0.027***	0.608***	0.550***	0.508***	0.601***	0.587***	0.407***	0.605***	0.503***
RG Married	(0.029)	(0.023)	(0.008)	(0.059)	(0.022)	(0.043)	(0.109)	(0.033)	(0.060)	(0.035)	(0.032)
Observations	181,094	87,071	94,023	46,538	50,605	46,758	37,193	116,059	65,035	102,502	78,592
Women, 2003-2018 Heckman's Two-Step											
Proportion of	0.422***	0.122***	0.112***	0.457***	0.405***	0.415***	0.381***	0.419***	0.310***	0.245***	0.409***
RG Married	(0.020)	(0.026)	(0.014)	(0.015)	(0.030)	(0.031)	(0.072)	(0.024)	(0.037)	(0.023)	(0.024)
Observations	230,181	95,856	134,325	53,784	66,707	63,776	45,914	116,265	113,916	55,585	174,596

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. Standard errors are reported in parentheses, where these are clustered by state. The dependent variable in all columns is a dummy for married. This regression represents the first-stage of the 2SLS procedure. Year and state fixed effects are included in all regressions. The following additional controls are included: dummies for highest level of educational attainment, potential experience, potential experience squared, dummies for number of children below the age of 5, dummies for number of children aged 5-17. The instrument in all specification in the proportion of individuals' reference group that are married. The first four row blocks present IV estimates, the final two present selection-corrected IV estimates. Columns 2 and 3 present results based on whether individuals where above or below the median based on predicted marriage. This involved running a linear probability model of the married dummy on all key covariates, but not the instrument.

Table E2: Alternative 2SLS First Stage Specification, CPS

	Men				**Women**							
Inverse Mills Ratio	**No**				**No**				**Yes**			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Proportion of RG Married	0.652***	0.653***	0.637***	0.637***	0.429***	0.430***	0.428***	0.428***	0.422***	0.420***	0.414***	0.414***
	(0.020)	(0.020)	(0.023)	(0.023)	(0.021)	(0.021)	(0.025)	(0.025)	(0.020)	(0.020)	(0.025)	(0.025)
U.S. Born		-0.042***	-0.043***	-0.043***		-0.062***	-0.068***	-0.067***		-0.103***	-0.115***	-0.113***
		(0.005)	(0.006)	(0.006)		(0.007)	(0.010)	(0.010)		(0.008)	(0.011)	(0.011)
Health Status:												
Excellent		0.042***	0.038***	0.037***		0.046***	0.044***	0.043***		0.034***	0.030***	0.032***
		(0.003)	(0.004)	(0.004)		(0.004)	(0.005)	(0.005)		(0.004)	(0.006)	(0.006)
Very Good		0.028***	0.026***	0.026***		0.028***	0.028***	0.027***		0.009**	0.008*	0.010**
		(0.002)	(0.003)	(0.003)		(0.003)	(0.004)	(0.004)		(0.003)	(0.004)	(0.004)
Fair		-0.014***	-0.019***	-0.017***		-0.021***	-0.018***	-0.013**		0.061***	0.076***	0.060***
		(0.005)	(0.006)	(0.006)		(0.006)	(0.006)	(0.006)		(0.007)	(0.009)	(0.008)
Poor		-0.019	-0.030	-0.022		-0.019	-0.025	-0.013		0.189***	0.202***	0.161***
		(0.015)	(0.019)	(0.019)		(0.018)	(0.024)	(0.023)		(0.023)	(0.032)	(0.029)
Difficulty:												
Hearing				-0.009				-0.053**				-0.054**
				(0.011)				(0.023)				(0.022)
Vision				-0.071***				-0.031				0.021
				(0.016)				(0.022)				(0.021)
Physical				-0.033***				-0.067***				0.070***
				(0.012)				(0.017)				(0.020)
Years Available	2003-2018	2003-2018	2009-2018	2009-2018	2003-2018	2003-2018	2009-2018	2009-2018	2003-2018	2003-2018	2009-2018	2009-2018
p-value: Extra Covariates = 0		0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Adjusted R^2	0.411	0.413	0.415	0.415	0.276	0.279	0.289	0.290	0.277	0.280	0.290	0.291
Observations	254,562	254,562	145,664	145,664	230,181	230,181	131,470	131,470	230,181	230,181	131,470	131,470

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. Standard errors are reported in parentheses, where these are clustered by state. The dependent variable in all columns is a dummy for married. This regression represents the first-stage of the 2SLS procedure. Year and state fixed effects are included in all regressions. The following additional controls are included: dummies for highest level of educational attainment, potential experience, potential experience squared, dummies for number of children below the age of 5, dummies for number of children aged 5-17. The instrument in all specification in the proportion of individuals' reference group that are married. The first two column blocks present IV estimates, the final one presents selection-corrected IV estimates. The additional covariates presented in the tables are not available for the earlier 1977-1992 period. Those on physical difficulties are only available from 2009 onwards. Hence, columns 3, 7 and 11 re-estimate the columns 2, 6 and 10 specification, but for a restricted time period. This time period is noted at the bottom of the table. The p-value is from a joint test of statistical significance of all additional covariates.

E.2 Complier Types

In this section, we follow the approach in Dahl et al. (2014) to calculate the fraction of compliers (those whose marriage decision was impacted by their value of $Z_{M,i}$), always takers (those who would marry irrespective of their value of $Z_{M,i}$), and never takers (those who would never marry irrespective of their value of $Z_{M,i}$). For compliers, we can write their proportion as:

$$\pi_c \equiv Pr(M_i = 1|Z_{M,i} = \overline{Z}) - Pr(M_i = 1|Z_{M,i} = \underline{Z}) = Pr(M_i(\overline{Z}) > M_i(\underline{Z})), \tag{27}$$

where $\underline{Z}$ and $\overline{Z}$ are the minimum and maximum values of the instrument, respectively. By conditional independence and monotonicity we can also write the proportion of always takers:

$$\pi_a \equiv Pr(M_i = 1|Z_{M,i} = \underline{Z}) = Pr(M_i(\overline{Z}) = M_i(\underline{Z}) = 1), \tag{28}$$

and the proportion of never takers:

$$\pi_n \equiv Pr(M_i = 1|Z_{M,i} = \overline{Z}) = Pr(M_i(\overline{Z}) = M_i(\underline{Z}) = 0). \tag{29}$$

Table E3 presents these proportions for each sample-specification combination, using both a local linear and a linear model, and using a variety of definitions for the values of $\underline{Z}$ and $\overline{Z}$. The local linear model is a flexible version of the first stage equation, Equation 14 for the non-selection corrected 2SLS approach and Equation 17 for the selection-corrected counterpart. After residualizing marriage with respect to our control variables, we run a local linear regression of resizualized marriage on our instrument. Figure E1 presents the local linear regression relation between residualized marriage and local social norms in marriage underlying these calculations.

We can also calculate these proportions using a linear model (i.e. Equations 14 and 17). In this case, we use the parameters from the first stage regression to calculate $\pi_c = \hat{\pi}_1(\overline{Z} - \underline{Z})$, $\pi_a = \hat{\pi}_{2,0} + \hat{\pi}_1\underline{Z}$ and $\pi_n = 1 - \hat{\pi}_{2,0} - \hat{\pi}_1\overline{Z}$, where $\hat{\pi}_{2,0}$ is the first stage constant and $\hat{\pi}_1$ the coefficient on the instrument. The proportion of compliers is typically larger when using the linear model. One can get a sense of why by reviewing Figure E1. The impact of the instrument is broadly linear, but does taper off towards the higher values. The local linear model captures this feature while the linear model does not.

It should be noted that the calculated proportion of compliers is large, a consequence of the fact that the instrument is age-dependent. The exercise implicit in the calculations imagines giving an individual the lowest and highest levels of the instrument, $\underline{Z}$ and $\overline{Z}$, and tracing out the impacts on marriage decisions. This exercise can never fully map to reality, as it involves changing the ages (and education levels) of individuals, in order that they are exposed to a different reference group.

Table E3: Sample Share by Compliance Category, CPS

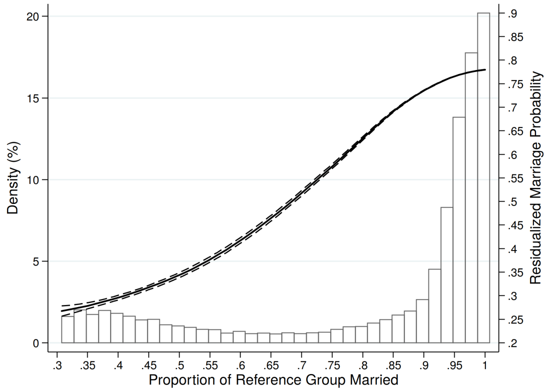
Model:	**Local Linear**				**Linear**			
%RG Married, top/bottom:	1%	1.5%	2%	5%	1%	1.5%	2%	5%
Men, 1977-1992								
Compliers	0.54	0.53	0.53	0.52	0.63	0.61	0.60	0.55
Never Takers	0.21	0.21	0.21	0.21	0.20	0.20	0.20	0.20
Always Takers	0.25	0.26	0.26	0.27	0.17	0.19	0.20	0.24
Men, 2003-2018								
Compliers	0.46	0.46	0.45	0.42	0.54	0.53	0.52	0.46
Never Takers	0.39	0.39	0.39	0.40	0.38	0.38	0.38	0.40
Always Takers	0.14	0.15	0.15	0.17	0.08	0.09	0.10	0.14
Women, 1977-1992								
Compliers	0.31	0.31	0.31	0.30	0.40	0.38	0.37	0.31
Never Takers	0.32	0.32	0.32	0.33	0.29	0.29	0.29	0.31
Always Takers	0.37	0.37	0.37	0.38	0.31	0.33	0.34	0.38
Women, 2003-2018								
Compliers	0.24	0.23	0.23	0.21	0.28	0.26	0.25	0.22
Never Takers	0.52	0.52	0.52	0.53	0.52	0.52	0.52	0.54
Always Takers	0.24	0.25	0.25	0.26	0.20	0.22	0.22	0.25
Women, 1977-1992 Heckman's Two-Step								
Compliers	0.31	0.31	0.31	0.29	0.40	0.38	0.37	0.31
Never Takers	0.26	0.26	0.26	0.27	0.23	0.23	0.23	0.25
Always Takers	0.43	0.43	0.43	0.44	0.37	0.39	0.40	0.44
Women, 2003-2018 Heckman's Two-Step								
Compliers	0.23	0.22	0.22	0.20	0.27	0.26	0.25	0.21
Never Takers	0.30	0.30	0.30	0.31	0.29	0.29	0.30	0.31
Always Takers	0.48	0.48	0.48	0.49	0.44	0.45	0.45	0.48

E.3 Characterizing Compliers

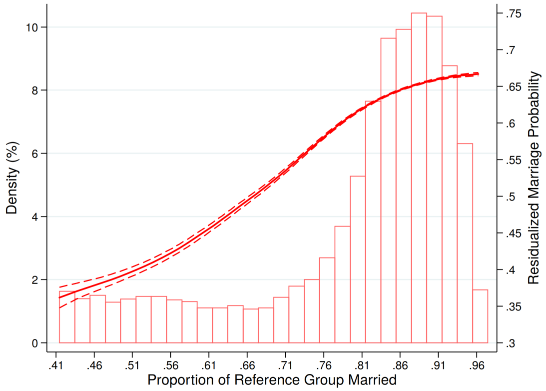
The statistic of interest to characterize the compliers is $\frac{P[X=x|complier]}{P[X=x]}$. In order to calculate the numerator, we calculate several ancillary statistics:

$$P[X = x|complier] = \frac{P[complier|X = x] \times P[X = x]}{P[complier]}, \tag{30}$$

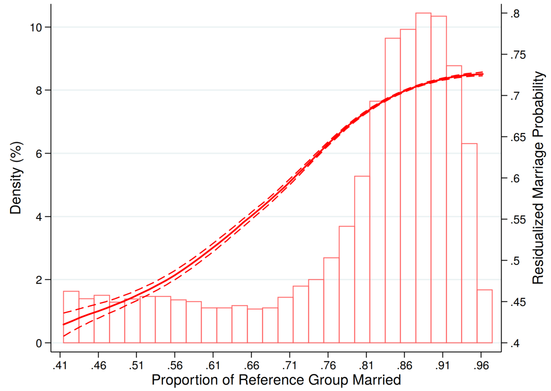
where $P[complier] = \hat{\pi}_1(\overline{Z} - \underline{Z})$ is calculated as described in the Section E.2 and $P[X = x]$ is the probability that $X = x$. $P[complier|X = x] = \hat{\pi}_{1,x}(\overline{Z} - \underline{Z})$, where $\pi_{1,x}$ is the first stage coefficient on the instrument based on the sub-sample $X = x$. Table E4 presents the results.



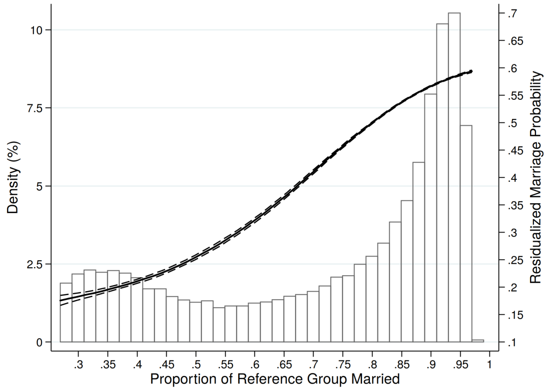
(a) Men, 1977-1992



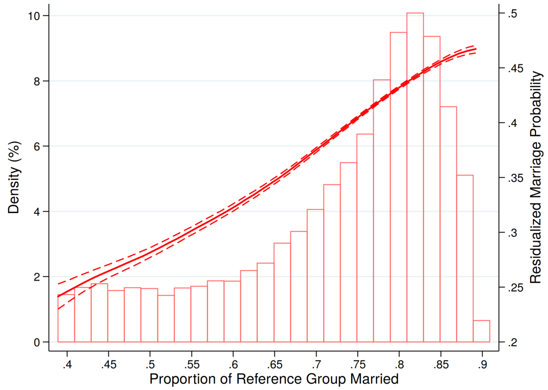
(b) Women, 1977-1992



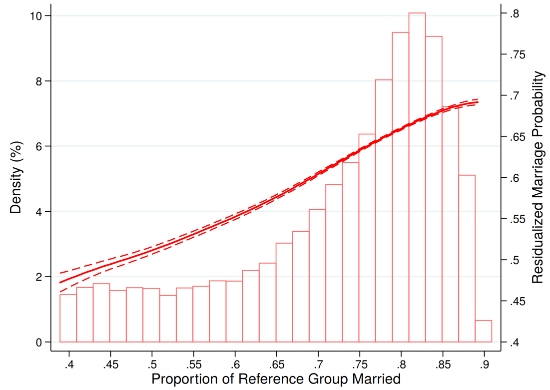
(c) Women - Selection Corrected, 1977-1992



(d) Men, 2003-2018



(e) Women, 2003-2018



(f) Women - Selection Corrected, 2003-2018

Figure E1: First Stage Relationship of Married on Proportion of Reference Group Married

Notes: The solid lines are local linear regression of residualized marriage on the instrumented, and is a flexible version of the first stage 2SLS equation. Marriage is residualized on year and state fixed effects,dummies for highest level of educational attainment, potential experience, potential experience squared, dummies for number of children below the age of 5, dummies for number of children aged 5-17. The instrument in all specification in the proportion of individuals' reference group that are married. Panels C and F include the Inverse Mills Ratio as an additional covariate. The dashed lines are 95% confidence intervals. The histogram of the instrument is shown in the background, with the top and bottom 1% excluded from the figure.

Table E4: Complier Characteristics, CPS

	1977-1992				2003-2018			
	First Stage	$P[X=x]$	$P[X=x\|complier]$	$\frac{P[X=x\|complier]}{P[X=x]}$	First Stage	$P[X=x]$	$P[X=x\|complier]$	$\frac{P[X=x\|complier]}{P[X=x]}$
Men								
Age:								
25-34	0.766	0.452	0.432	0.957	0.544	0.353	0.295	0.834
35-44	0.801	0.330	0.330	1.001	0.769	0.333	0.393	1.178
45-54	0.314	0.218	0.086	0.393	0.423	0.313	0.203	0.648
Education:								
HS Dropout	0.864	0.125	0.135	1.079	0.582	0.043	0.038	0.892
HS graduate	0.832	0.387	0.402	1.040	0.558	0.275	0.235	0.855
Some College	0.810	0.197	0.200	1.012	0.675	0.282	0.292	1.035
College Graduate	0.729	0.165	0.150	0.912	0.693	0.272	0.289	1.062
Advanced Graduate	0.705	0.126	0.111	0.881	0.680	0.128	0.133	1.042
Women								
Age:								
25-34	0.579	0.453	0.458	1.011	0.397	0.345	0.319	0.924
35-44	0.474	0.336	0.278	0.828	0.469	0.328	0.359	1.092
45-54	0.297	0.210	0.109	0.518	0.207	0.326	0.157	0.482
Education:								
HS Dropout	0.556	0.095	0.092	0.970	0.251	0.024	0.014	0.586
HS graduate	0.593	0.465	0.481	1.034	0.232	0.216	0.117	0.541
Some College	0.543	0.197	0.187	0.947	0.359	0.297	0.249	0.836
College Graduate	0.438	0.150	0.115	0.764	0.443	0.302	0.312	1.032
Advanced Graduate	0.456	0.092	0.073	0.796	0.483	0.160	0.180	1.125
Women, Heckman's Two-Step Estimator								
Age:								
25-34	0.588	0.453	0.456	1.005	0.360	0.345	0.295	0.853
35-44	0.486	0.336	0.279	0.830	0.452	0.328	0.352	1.071
45-54	0.305	0.210	0.110	0.521	0.206	0.326	0.160	0.489
Education:								
HS Dropout	0.561	0.095	0.091	0.959	0.275	0.024	0.016	0.651
HS graduate	0.598	0.465	0.476	1.022	0.240	0.216	0.123	0.568
Some College	0.558	0.197	0.188	0.953	0.359	0.297	0.253	0.850
College Graduate	0.452	0.150	0.116	0.773	0.430	0.302	0.308	1.019
Advanced Graduate	0.475	0.092	0.075	0.812	0.459	0.160	0.175	1.088

Notes:The dependent variable in Columns 1 and 5 is a dummy for married. This regression represents the first-stage of the 2SLS procedure. Year and state fixed effects are included in all regressions. The following additional controls are included: dummies for highest level of educational attainment, potential experience, potential experience squared, dummies for number of children below the age of 5, dummies for number of children aged 5-17. The instrument in all specification in the proportion of individuals' reference group that are married.

E.4 Marginal Treatment Effects

In order to consider heterogeneity of the treatment effect of marriage on wages, we expand our IV analysis and consider marginal treatment effects (MTEs). The underlying framework is a generalized Roy (1951) model, with Y_0 and Y_1 respectively denoting potential outcomes if never married and married. The marriage decision is written as:

$$M = \mathbb{1}\{\mu_M(X, Z_M) > V\}, \tag{31}$$

where μ_M is any function, X includes all control variables and fixed effects as above, Z_M is the instrument for marriage, and V is an unobserved, continuous variable. Given that V enters the latent index determining treatment, we think of it as unobserved resistance to treatment i.e. marriage. As long as V is indeed continuous, we can rewrite the marriage decision equation as:

$$M = \mathbb{1}\{P(X, Z_M) > U_M\}, \tag{32}$$

where $P(X, Z_M)$ is the propensity score, and U_M are quantiles of V.

MTEs trace the treatment effect along the (unobserved) resistance to treatment. The resistance to treatment is the driver of selection on unobserved gains to treatment/marriage. Those who choose to get married due to especially low resistance to marriage may have different gains than those with high resistance.

We estimate the MTEs using the separate approach as suggested by Heckman and Vytlacil (2007), and follow Brinch, Mogstad, and Wiswall (2017) in the implementation. Specifically we estimate the conditional expectation of Y separately for the married and never married with the regression

$$Y_j = X\beta_j + K_j(p) + \epsilon, \text{ for } j = 0, 1, \tag{33}$$

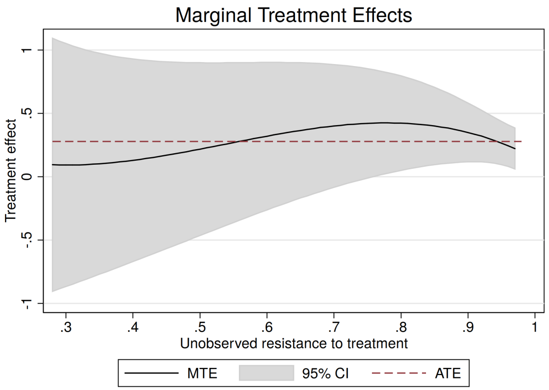
where the control function $K_j(p)$ is based on a cubic polynomial in p. With $K_j(p)$ in hand, we can estimate the MTE as

$$MTE(x, u) = \mathbb{E}(Y_1 | X = x, U_M = u) - \mathbb{E}(Y_0 | X = x, U_M = u) \tag{34}$$

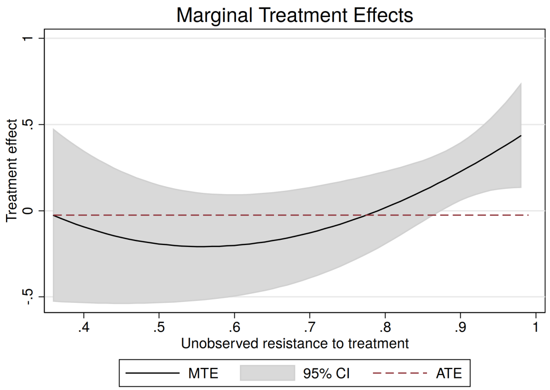
$$= x(\beta_1 - \beta_0) + k_1(u) - k_0(u) \ , \tag{35}$$

where $k_j(u) = \mathbb{E}(U_j | U_M = u)$. With the conditional independence assumption made in Section 4.4.2, as well as assuming separability between observed and unobserved heterogeneity in the treatment effects, the MTE is identified over the common support of the propensity score $P(X, Z_M)$.

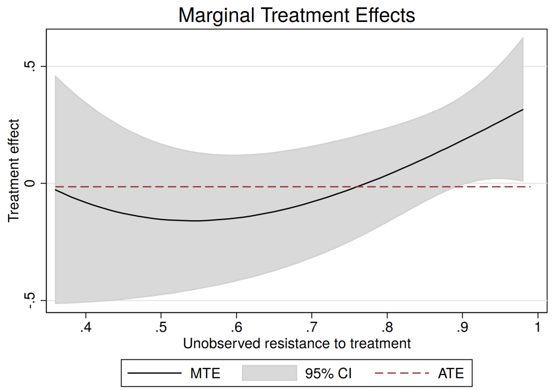
Using the separate approach and a cubic polynomial, we estimate MTE curves for our key sample-specification combinations, and present these in Figure E2.



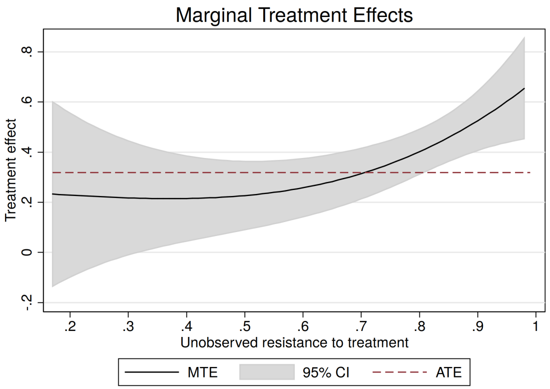
(a) Men, 1977-1992



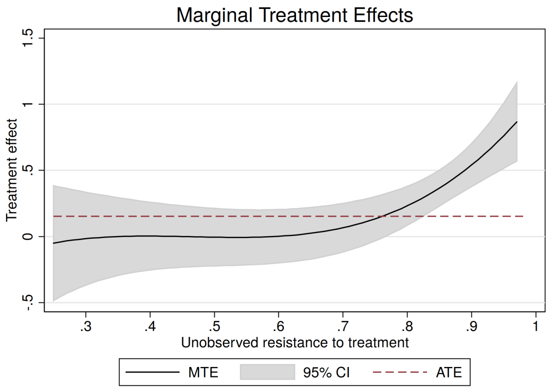
(b) Women, 1977-1992



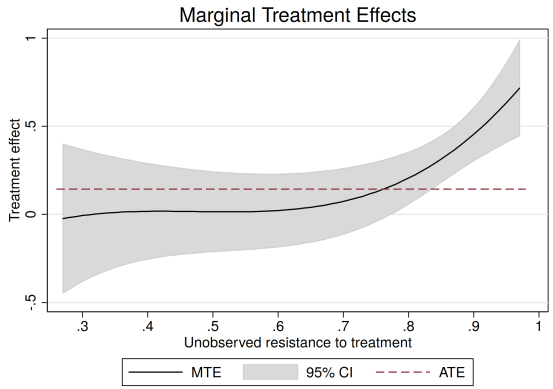
(c) Women - Selection Corrected, 1977-1992



(d) Men, 2003-2018



(e) Women, 2003-2018



(f) Women - Selection Corrected, 2003-2018

Figure E2: Marginal Treatment Effects

Notes: The figures plot the MTE curve over the support of the resistance to treatment, once the lowest and highest .5% of propensity scores have been trimmed. The MTEs are estimated based on the separate approach, with a cubic polynomial. 95% confidence interval bands are in gray. In all specification we include year and state fixed effects,as well as dummies for highest level of educational attainment, potential experience, potential experience squared, dummies for number of children below the age of 5, dummies for number of children aged 5-17. The instrument in all specification in the proportion of individuals' reference group that are married. Panels C and F include the Inverse Mills Ratio as an additional covariate.