On the Marriage Wage Premium*

Brendon McConnell†
University of Southampton

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

July 10, 2020

Abstract

In this paper, we use a novel instrument based on local social norms towards marriage to present a new fact: marriage has a positive causal effect on the wages of both men and women. Despite the striking changes in the labor market and the composition of families that occurred over the last decades, the substantial positive effect of marriage on the wages of men has remained largely unchanged. Conversely, while marriage decreased the wages of women until the 1980s, we document the emergence of a sizable marriage wage premium from the late 2000s onward. The fact that marriage increases the wages of women displaces the main hypotheses that the literature discussed to explain the positive relationship between marriage and the wages of men. Namely, the idea that married men are able to devote more resources to their careers than their single counterparts because their wives specialize in home work. Further, we highlight the fact that the effect of marriage on wages is heterogeneous both between and within genders. In particular, the marriage wage premium is larger for women above the median of the wage distribution, whereas for men we find the opposite.

*We thank Effrosyni Adamopoulou, Alex Armand, Nezih Guner, Ezgi Kaya, Michael Knaus, Michael Lechner, Attila Lindner, Joan Llull, Shelly Lundberg, Jaime Milln-Quijano, Richard Murphy, Imran Rasul, Anna Raute and seminar participants at the Berlin Applied Micro Seminar, the Swiss Macro Workshop, the University of Bristol, the University of Nottingham, the University of St. Gallen, the University of York, and the Online Discrimination and Disparities Seminar for valuable comments. All errors are ours.

†b.i.mcconnell@soton.ac.uk.

‡arnau.valladares@gmail.com.

1 Introduction

Over the last decades, the United States, along with many other countries, has experienced a remarkable 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 mating have all dramatically changed.[1] Placing the lens on gender, the labor market outcomes of women have also 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 document 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 use Current Population Survey (CPS) data from 1977 to 2018 to show that, while until the mid 1980s marriage caused a penalty on the wages of women, since the mid 2000s there is a sizable premium. Interestingly, the causal effect of marriage on the wages of men has remained relatively stable over the same period. The MWP for women is crucial to understand the secular changes that have taken place over the last decades. 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 the Second World War 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 in order to understand the social transformation of the last decades. Moreover, the theories that have been proposed to explain the MWP of men often rely on intra-household arguments. In particular, the literature 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 in recent years is at odds with this hypothesis.[6]

Establishing a causal effect of marriage on wages presents three main challenges.[7] First, for women, there is a sizable part of the population that does not participate in employment. The underlying economic decision that generates this outcome implies that the sample of observed

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

[2] There is a vast literature studying the evolution of the labor market outcomes of women. See 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, 2020) or Guner, Kulikova, and Llull (2018).

[4] See Hill (1979), Korenman and Neumark (1992), Loughran and Zissimopoulos (2009), Ginther and Sundström (2010), Juhn and McCue (2016, 2017), and Pilossoph and Wee (2019).

[5] See Korenman and Neumark (1991), Loh (1996), Cornwell and Rupert (1997), Ginther and Zavodny (2001), Stratton (2002), Antonovics and Town (2004), Krashinsky (2004), Ahituv and Lerman (2007), Bardasi and Taylor (2008), and Killewald and Gough (2013).

[6] Another hypothesis discussed in the literature 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. In Appendix E, we adapt the approaches of Altonji and Pierret (2001) and Pinkston (2009) to the case of marriage and show that statistical discrimination is not a relevant mechanism behind the marriage wage premium of women and men.

[7] Some authors analyze other aspects related to the relationship between marriage and wages. See Gray (1997) and Maasoumi, Millimet, and Sarkar (2009).

wages for women is not a random representation of the population. Hence, the estimated coefficients in the wage equation might suffer sample-selection bias. Second, there may be unobservable variables that affect both the propensity to be married and wages. That is, the estimated coefficients on marriage in the wage equation may suffer from omitted-variable bias. Thirdly, there may be an issue of reverse causality if wages also affect the probability of being married. In our main specification, we tackle all three issues combining a Heckman (1979) sample-selection correction with a novel instrument for marriage based on local social norms.[8] 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 individuals in our analysis to proxy for the relevant local social norms that affect the marriage decision of that individual.[9] For women, we correct for sample-selection bias into employment using the age of the youngest child in the household, which enables us to control for the presence of children in the wage equation. Our results indicate that, nowadays, marriage increases the wages of women by about 9 percentage points, while the effect is of around 20 percentage points for men.[10]

We present evidence that the effect of marriage on wages is notably different across the distribution of wages. For men, the MWP presents a decreasing trend along the wage distribution. That is, for men at the lower end of the wage distribution the effect of marriage on their wages is larger than for men at the top. In the late 1970s and early 1980s this pattern is specially pronounced while it has flattened over time. In recent years, the MWP of men is fairly similar along the wage distribution. The MWP of women has experienced a similar evolution albeit with an opposite starting point. These patterns suggest that the within-household-specialization channel, along with the degree of assortative mating, might have been relevant to understand why women at the bottom of the wage distribution experienced a penalty until the early 2000s while men at the bottom enjoyed a larger premium than men at the top. However, neither the emergence of a MWP for women at the top of the wage distribution in the early 1990s nor the current patterns of the MWP for both genders are consistent with within-household specialization being the main driver of the effect of marriage on wages.

We make three key contributions. First, our analysis is the first to present causal evidence that, nowadays, marriage generates a positive effect on the wages of both men and women. We document that, while marriage has increased the wages of men for several decades, the premium for women emerged in the mid 2000s. The marriage wage premium is important to understand both the relationship between the family and outcomes in the labor market, and the social transformation of the last decades. Second, we provide evidence that the effect of marriage on wages is heterogeneous. This suggest that there might be several mechanisms that bring about the positive effect of marriage on wages and that different people might be impacted by different mechanisms. Thirdly, we establish that there is no unifying theory that explains the existence

[8]There is a growing literature that studies how social norms affect family outcomes. See Drewianka (2003), Fernández and Fogli (2009), Adamopoulou (2012), Mourifié and Siow (2017), Adamopoulou and Kaya (2018), and Vickery and Anderberg (2019).

[9]We provide comprehensive evidence to support each of the identifying assumptions underling the instrumental variable approach. Notably, we use the plausibly exogenous method of Conley, Hansen, and Rossi (2012) and the imperfect instrumental variable approach of Nevo and Rosen (2012) to show that our results are robust to violations of the exclusion restriction.

[10]Importantly, the change in the selection pattern of women into employment of the last decades plays a negligible role in the determination of the marriage wage premium for women.

of the marriage premium and its evolution over the last decades. Our paper establishes the key elements that a successful theory about the effect of marriage on wages needs to reproduce.

The rest of the paper is organized as follows. In Section 2, we describe the data we use and the sample restrictions we impose, as well as, presenting descriptive evidence on the relationship between marriage and wages over the last decades. In Section 3, we describe our instrument and present the causal patterns of the effect of marriage on wages. In Section 4, we analyze the causal effect of marriage along the wage distribution. Section 5 discusses how to reconcile our results with the different theories on the marriage wage premium. Finally, Section 6 concludes.

2 Descriptives

We use data from the March Supplement of the CPS from 1977 to 2018.[11] Our 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 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.[12]

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.[13] We use the Annual Social and Economic Supplement weights in all the analysis.

Our final sample contains 1,531,669 observations, 731,632 men and 800,037 women. Tables 1 and 2 present key descriptive statistics for men and women respectively. We divide the sample in seven-years periods in order to study the evolution of the marriage wage premium over time. The observed patterns, both for men and women, are consistent with well-documented trends 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.

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

[12] 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 A, we reproduce our main analysis using a sample in which the non-married group includes never-married, divorced, and separated individuals. The coefficients estimated with this alternative definition of the non-married group are in line with our main results.

[13] 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.

Table 1: Descriptive Statistics - Men
Means, Standard Deviations in Parentheses

	1977-1983	**1984-1990**	**1991-1997**	**1998-2004**	**2005-2011**	**2012-2018**
Sample Size	121,172	120,080	112,404	129,390	137,634	110,952
Married	0.841	0.791	0.760	0.745	0.709	0.662
Hourly Wage (1999 Dollars)	19.81	19.29	18.40	19.63	19.81	19.81
	(9.31)	(9.94)	(9.88)	(10.71)	(11.33)	(11.82)
Age	37.55	37.13	38.05	39.27	39.53	39.10
	(8.83)	(8.38)	(8.24)	(8.40)	(8.78)	(8.93)
Highest Level of Education:						
HS Dropout	0.168	0.116	0.085	0.070	0.062	0.051
HS Graduate	0.368	0.378	0.335	0.310	0.310	0.277
Some College	0.187	0.201	0.262	0.277	0.274	0.275
College Graduate	0.150	0.171	0.209	0.238	0.243	0.269
Advanced Graduate	0.126	0.134	0.110	0.105	0.111	0.129
Number Children, 0-4	0.297	0.296	0.268	0.248	0.250	0.238
	(0.592)	(0.599)	(0.570)	(0.557)	(0.562)	(0.549)
Number Children, 5-17	1.12	0.898	0.842	0.828	0.788	0.753
	(1.29)	(1.12)	(1.08)	(1.08)	(1.07)	(1.08)
Children, 18 and over	0.160	0.141	0.129	0.127	0.122	0.114

Notes: Data used: CPS, 1977-2018.

Table 2: Descriptive Statistics - Women
Means, Standard Deviations in Parentheses

	1977-1983	**1984-1990**	**1991-1997**	**1998-2004**	**2005-2011**	**2012-2018**
Sample Size	132,308	130,154	121,247	142,474	152,620	121,234
Married	0.901	0.864	0.846	0.829	0.803	0.751
Employed	0.619	0.713	0.763	0.774	0.759	0.751
Hourly Wage (1999 Dollars)	11.91	12.86	13.65	15.18	15.79	16.26
	(6.25)	(7.24)	(8.04)	(9.08)	(9.55)	(10.14)
Age	37.80	37.28	38.19	39.44	39.80	39.31
	(8.83)	(8.39)	(8.22)	(8.34)	(8.73)	(8.92)
Highest Level of Education:						
HS Dropout	0.162	0.104	0.073	0.053	0.045	0.037
HS Graduate	0.475	0.447	0.362	0.305	0.262	0.219
Some College	0.173	0.205	0.275	0.295	0.295	0.283
College Graduate	0.121	0.153	0.203	0.244	0.271	0.300
Advanced Graduate	0.068	0.091	0.087	0.103	0.126	0.161
Number Children, 0-4	0.270	0.288	0.272	0.252	0.261	0.255
	(0.569)	(0.589)	(0.570)	(0.559)	(0.570)	(0.564)
Number Children, 5-17	1.25	1.01	0.958	0.948	0.926	0.903
	(1.30)	(1.13)	(1.10)	(1.10)	(1.11)	(1.12)
Children, 18 and over	0.203	0.178	0.161	0.154	0.156	0.150

Notes: Data used: CPS, 1977-2018.

2.1 The Correlation between Marriage and Wages

We start by measuring the conditional correlation between being married and hourly wages over time separately for men and women. Using OLS we estimate the following linear regression model:

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

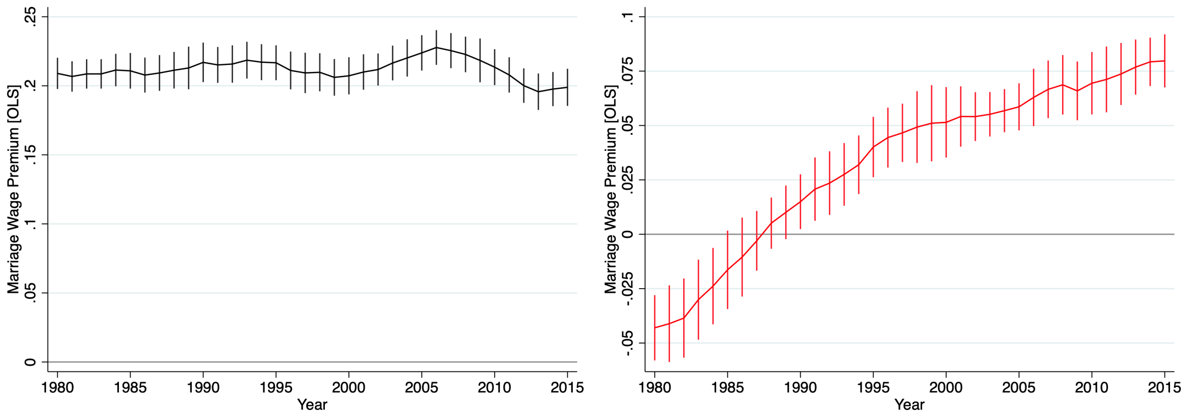
where y_i is the natural logarithm of hourly wages of person i, M_i is a dummy variable that equals 1 when an individual reports to be married and living with their spouse, X_i is a set of demographic controls which consists of education-category dummies, the number of children below the age of 5, the number of children aged 5-17, a dummy for a child over the age of 18, and dummies for years of potential experience.[14] θ_s and ϕ_t are state and year fixed effects respectively. We cluster standard errors at the state level.

Figure 1 describes how the relationship between marriage and wages ($\hat{\alpha}$) has evolved over time for men and women. We estimate Equation 1 for all years in our sample, from $t = 1977$ to $t = 2018$, with a +/-3 year window around each year, and plot the resulting set of $\hat{\alpha}$ coefficients over time. The conditional correlation between marriage and hourly wages has changed markedly for women (Figure 1b). At the beginning of the period, marriage was associated with a wage penalty of 4.3%. This penalty linearly reduced over time until the mid 1980s. In the early 1990s, a positive correlation emerged and it continued to increase until the end of the sample period. In the last period, marriage is associated with a premium of 8.0%. This change in the correlation between being married and the wages of women is especially important in the context of the evolution of female labor force participation and the decline of marriage. For men, despite the remarkable changes in family structure that are documented in the literature, such as the decrease in the marriage rate, the increase of divorce, the rise of assortative mating, and the marked 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.[15]

[14] As it is standard in the literature, potential experience is computed as age minus years of education minus seven.

[15] A significant part of the literature on the MWP uses within-individual variation from panel data to identify the effect of marriage. For example, Killewald and Lundberg (2017) and Ludwig and Brüderl (2018) find no support for a positive effect of marriage on the wages of man using the National Longitudinal Survey of the Youth 1979 (NLSY79). In Appendix C, we provide a comprehensive comparison between our CPS-based results and their counterparts in the NLSY79. One of the main advantages of using the CPS is that it enables us to analyze the evolution of the effect of marriage on wages over more than four decades.

Figure 1: OLS-estimated Marriage Wage Premium over Time



(a) Men.

(b) Women.

Notes: The figures plot the OLS estimate $\hat{\alpha}$ from Equation 1, and 95% confidence intervals based on state-clustered standard errors as the vertical spikes. Each point centered on year t is estimated using observations from year $t-3$ to $t+3$. Control variables are as described in Section 2.1 above. Data used: CPS, 1977-2018.

3 The Causal Effect of Marriage on Wages

In this section, we present a novel instrument to estimate the causal effect of marriage on wages. We think of marriage as being not only a product of economic factors, marriage market conditions, preferences, and chance but also local social norms. At the same time, we assume that the social norms that affect marriage do not affect individual productivity and, thus, wages. To measure the prevalence of local social norms on the propensity of being married, 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. 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 affect the marriage decisions of the reference cohort persist to affect the marriage decisions of the individuals in our sample. We match on education and the presence of children because observable characteristics are a predictor of different social norms and also to proxy for homophilic social networks.

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

3.1 Empirical Specification

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

$$M_i = \pi_1 Z_{M,i} + X_i^{'}\pi_2 + \theta_{1s} + \phi_{1t} + \mu_i, \tag{2}$$
$$y_i = \alpha M_i + X_i^{'}\beta + \theta_{2s} + \phi_{2t} + \epsilon_i. \tag{3}$$

Equation 2 is the first stage which models marriage as dependent on the covariates used in Section 2.1 and the instrument $Z_{M,i}$. Equation 3 describes the second stage. It specifies how the logarithm of wages, y_i, depends on marriage and the same covariates as in previous specifications.

For women, we also run a selection-corrected version of the 2SLS specification described in Equation 2 and Equation 3:

$$E_i = \mathbb{1}\{\kappa_1 Z_{E,i} + \kappa_2 Z_{M,i} + X_i^{'}\kappa_3 + \theta_{1s} + \phi_{1t} + \xi_i > 0\} = \mathbb{1}\{Z_i^{'}\kappa + \xi_i > 0\}, \tag{4}$$
$$M_i = \pi_1 Z_{M,i} + X_i^{'}\pi_2 + \theta_{2s} + \phi_{2t} + \pi_5\lambda(Z_i^{'}\kappa) + \mu_i, \tag{5}$$
$$y_i = \alpha M_i + X_i^{'}\beta + \theta_{3s} + \phi_{3t} + \sigma_{13}\lambda(Z_i^{'}\kappa) + \epsilon_i. \tag{6}$$

Because we treat marriage as an endogenous variable, we use the instrument $Z_{M,i}$ instead of the dummy for marriage M_i in the employment equation (Equation 4). We start by estimating the employment decision (Equation 4) 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 2-3 and Equations 5-6 using 2SLS. We bootstrap the standard errors in the selection-corrected 2SLS procedure.[16]

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.[17] A particularly common exclusion restriction in the literature that studies the labor market outcomes of women is to use a dummy variable for the presence of own children in the household.[18] However, this option is not compatible with controlling for children in the wage equation. 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 (is more likely to affect the employment margin) than a teenager. Given this insight, we use the age of the youngest own child in the household as an exclusion restriction. To the best of our knowledge, our paper is the first to use this exclusion restriction. We note two relevant points. First, the dummies for the age of the youngest child are jointly significant in the probit employment equation. Secondly, the set of controls (X_i) in the wage equation includes the number of children below the age of 5, the number of children aged 5-17, and a dummy for a child over the age of 18. Hence, because we already control

[16]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.

[17]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).

[18]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 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.

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).[19] In the next section, we highlight the main features of the supporting evidence we provide in Appendix F 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 that social norms impact wages only through the marriage channel (Exclusion restriction).

3.2 Support for the Identifying Assumptions

First Stage. We provide evidence supporting the relevance of the instrument in three places. Figure F1 shows the first stage graphically, as well as presenting information on the distribution of the instrument. For both men and women, 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 3.1). In addition, the first column of Table F2 presents the first stage coefficient for each of the key sample specifications. Finally, we present first-stage F-statistics along with the results of the IV estimation in Tables 3, 4, and 5, which we present in Section 3.3. All pieces of evidence provide strong support for the relevance of the instrument.

Conditional Independence. Table F1 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 time period we analyze, hence, we do not 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. The estimates in Table F1 indicate that, indeed, 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 strong do 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 employs a setup of a binary endogenous regressor

[19] See Imbens and Angrist (1994).

and a continuous instrument as we do, notes that monotonicity implies we should see a non-negative first stage coefficient for any sub-sample. Table F2 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.

Exclusion restriction. Strictly speaking, the exclusion restriction is non-testable. In Appendix Section F.1.2, we use the imperfect instrumental variable approach of Nevo and Rosen (2012) and the plausibly exogenous method of Conley et al. (2012) to assess how moderate violations of the exclusion restriction affect our estimates. Both procedures indicate that allowing the instrument to violate the exclusion restriction does not substantially affect the coefficients estimated with our main specification. That is, the economic interpretation we derive from our results is robust to a moderate fail of the exclusion restriction.

3.3 Results

Table 3 presents the marriage coefficients estimated using IV ($\hat{\alpha}$ from Equation 3) for men in the six sub-periods of seven years that cover all data we have. Table 4 is its equivalent for women, while Table 5 contains the estimates from the specification in Equations 4, 5, and 6, which combines the selection correction with the IV. To ease comparison with the OLS estimates, the first row of each table contains the marriage coefficients estimated using OLS ($\hat{\alpha}$ from Equation 1). One of the differences between the OLS estimates and the IV coefficients is that the latter are estimated out of the complier population while the first are based on the whole population of interest. In order to understand how much of the difference between the OLS and IV estimates is due to the distinct populations from which they are estimated, the second row presents OLS estimates based on reweighting the main sample to reflect the observable characteristics of the complier population.[20] That is, the coefficients in the second and third rows are based on samples that reflect the same observable characteristics. Figure 2 presents the IV estimates using a 7-year rolling window over all years of our sample as in Figure 1.

The patterns observed from the IV estimates are in line with those observed out of the descriptive evidence of Section 2.1. Namely, the positive effect of marriage on the wages of men is sizable and has remained fairly stable over the last decades. As seen in Figure 2a, in the early 1980s marriage increases the wages of men by around 30%. The effect slightly decreases up until the mid 1990s, when it is of around 20%. Since the late 2010s, marriage rises the wages of men by around 24 percentage points. For women, marriage negatively affects wages during the 1980s. From the late 1980s until the mid 2000s, the effect is non-significantly different from zero with negative point estimates which are close to zero. Since the late 2000s, marriage increases the wages of women by around 9 percentage points.[21]

[20]In Appendix Section F.2, we provide the details of how we back out the observable characteristics of the complier population.

[21]In Appendix B, we show that these patterns are robust to including other controls. In particular to adding controls on industry and occupation. We do not include these controls in our main specification because we consider these characteristics to be endogenous to the treatment. That is, we think as marriage plausibly affecting the industry/occupation in which people work. Moreover, the employment-selection specification cannot include industry and occupation controls as these are undefined for non-employed individuals.

Table 3: IV - Men

A. OLS:	**(1)** **'77-'83**	**(2)** **'84-'90**	**(3)** **'91-'97**	**(4)** **'98-'04**	**(5)** **'05-'11**	**(6)** **'12-'18**
Baseline:						
Married	0.213***	0.208***	0.214***	0.209***	0.222***	0.200***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Complier Reweighted:						
Married	0.201***	0.195***	0.204***	0.198***	0.217***	0.197***
	(0.006)	(0.006)	(0.007)	(0.006)	(0.007)	(0.006)
B. IV:						
Married	0.306***	0.269***	0.246***	0.226***	0.266***	0.237***
	(0.013)	(0.015)	(0.013)	(0.016)	(0.017)	(0.021)
First-Stage F-Statistic	1461.3	2116.8	1304.2	1425.5	1426.4	1128.9
Adjusted R^2	0.198	0.238	0.263	0.249	0.268	0.265
Observations	104,970	104,545	96,210	112,807	120,606	94,896

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, dummies for year of potential experience, the number of children below the age of 5, the number of children aged 5-17, and a dummy for a child over the age of 18. The instrument in all specification is the proportion of individuals' reference group that are married. For the complier reweighted regressions, we first separate each sample into six mutually exclusive groups based on education and age, as outlined in Section F.2.2. We estimate the proportion of compliers in each sub-group, and then reweight our main estimation samples so that the complier proportion in each of the six sub-groups matches the proportion of the main sample for the selfsame sub-group. Data used: CPS, 1977-2018.

Table 4: IV - Women

A. OLS:	**(1)** **'77-'83**	**(2)** **'84-'90**	**(3)** **'91-'97**	**(4)** **'98-'04**	**(5)** **'05-'11**	**(6)** **'12-'18**
Baseline:						
Married	-0.053***	-0.001	0.031***	0.054***	0.070***	0.079***
	(0.008)	(0.008)	(0.007)	(0.007)	(0.007)	(0.006)
Complier Reweighted:						
Married	-0.032***	0.018**	0.049***	0.064***	0.081***	0.087***
	(0.008)	(0.009)	(0.007)	(0.007)	(0.007)	(0.007)
B. IV:						
Married	-0.159***	-0.061	-0.030	-0.034	0.078***	0.088***
	(0.025)	(0.045)	(0.029)	(0.044)	(0.027)	(0.032)
First-Stage F-Statistic	361.9	478.3	291.4	425.6	378.0	283.9
Adjusted R^2	0.148	0.200	0.237	0.227	0.226	0.227
Observations	70,872	83,482	83,791	101,175	109,624	85,427

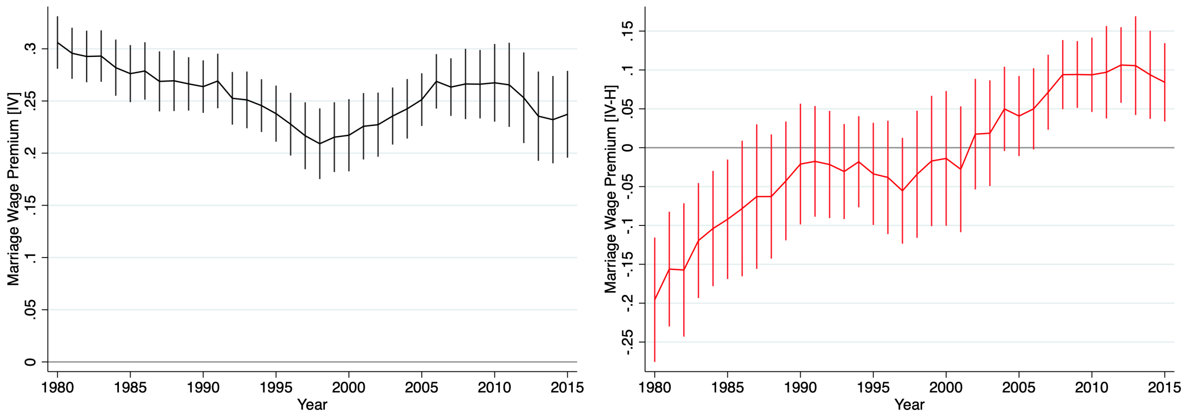
Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. Standard errors are reported in parentheses, where these are clustered by state. The dependent variable in all columns is the natural log of wages. Year and state fixed effects are included in all regressions. The following additional controls are included: dummies for highest level of educational attainment, dummies for year of potential experience, the number of children below the age of 5, the number of children aged 5-17, and a dummy for a child over the age of 18. The instrument in all specification is the proportion of individuals' reference group that are married. For the complier reweighted regressions, we first separate each sample into six mutually exclusive groups based on education and age, as outlined in Section F.2.2. We estimate the proportion of compliers in each sub-group, and then reweight our main estimation samples so that the complier proportion in each of the six sub-groups matches the proportion of the main sample for the selfsame sub-group. Data used: CPS, 1977-2018.

Table 5: IV-Heckman - Women

A. Heckman:	**(1)** **'77-'83**	**(2)** **'84-'90**	**(3)** **'91-'97**	**(4)** **'98-'04**	**(5)** **'05-'11**	**(6)** **'12-'18**
Baseline:						
Married	-0.044***	0.000	0.033***	0.052***	0.068***	0.080***
	(0.008)	(0.008)	(0.007)	(0.008)	(0.007)	(0.006)
Complier Reweighted:						
Married	-0.025***	0.019**	0.051***	0.061***	0.079***	0.087***
	(0.008)	(0.008)	(0.007)	(0.008)	(0.007)	(0.007)
B. IV-Heckman:						
Married	-0.149***	-0.058	-0.018	-0.033	0.104***	0.086***
	(0.022)	(0.037)	(0.028)	(0.040)	(0.031)	(0.032)
First-Stage F-Statistic	352.6	485.8	265.2	417.4	311.0	236.8
Adjusted R^2	0.149	0.200	0.237	0.227	0.225	0.227
Observations	70,872	83,482	83,791	101,175	109,624	85,427

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, dummies for year of potential experience, the number of children below the age of 5, the number of children aged 5-17, and a dummy for a child over the age of 18. The instrument in all specification is the proportion of individuals' reference group that are married. The exclusion restrictions for the employment equation are a series of dummies for age of youngest child in the household from 1-18, where age less that 1 is the base category, a dummy for ages 19-24, a dummy for 25 and over, and no children are also included. In this case we bootstrap standard errors, allowing for clustering at the state level, and using 500 iterations. For the complier reweighted regressions, we first separate each sample into six mutually exclusive groups based on education and age, as outlined in Section F.2.2. We estimate the proportion of compliers in each sub-group, and then reweight our main estimation samples so that the complier proportion in each of the six sub-groups matches the proportion of the main sample for the selfsame sub-group. Data used: CPS, 1977-2018.

Figure 2: Marriage Wage Premium over Time - 3 year window



(a) IV - Men. (b) IV-Heckman - Women.

Notes: The figures plot the IV and IV-Heckman estimates of α from Equations 3 and 6 for men and women respectively, and 95% confidence intervals based on state-clustered standard errors as the vertical spikes. Each point centered on year t is estimated using observations from year $t-3$ to $t+3$. Control variables are as outlined in Section 3.1 above. Data used: CPS, 1977-2018.

The IV estimates reveal a moderately larger effect of marriage on wages than the OLS coefficients. According to the IV estimates, both the marriage wage premium (for men and women) and the marriage wage penalty that women experience in the 1980s are larger than what is estimated from OLS. This difference between IV and OLS estimates may appear through three different channels. First, if marriage has heterogeneous effects, it is possible that, in the complier population, marriage has a larger effect than in the whole population. The OLS estimates reweighted to reflect the observable characteristics of the complier population (second row of Tables 3, 4, and 5) indicate that, for women, this channel may be part of the explanation. As seen in Tables 4 and 5, the complier-reweighted estimates are systematically higher than the OLS for the full sample. In the case of men (Table 3), we see an opposite pattern. Second, if there are unobservable factors that increase (decrease) wages while decreasing (increasing) the probability of being married, the OLS estimates are downward biased. For women, there is evidence that certain traits and behavior that are positively associated with career success may reduce the likelihood of marriage.[22] On the other hand, for men, the literature on the MWP suggests that the effect of unobservable variables downward biases the OLS estimates. Thirdly, if an exogenous increase (decrease) in wages reduces (increases) the probability of being married, the OLS coefficients are downward biased due to reverse causality. Theoretically, an exogenous change in wages affects the material resources that a single person brings to marriage changing their option value of marriage and how appealing they might be to potential spouses. At the same time, it also changes their option value to remain single, which implies that the direction of the simultaneity bias is ambiguous. Regalia, Ríos-Rull, and Short (2011), Salcedo, Schoellman, and Tertilt (2012), and Greenwood, Guner, Kocharkov, and Santos (2016) present quantitative models in which the rate of marriage is driven down by the improvement of the option value of singlehood. Hence, it is plausible that the OLS coefficients are downward biased due to reverse causality.[23]

It is worth highlighting that the uncertainty about the discrepancy between the OLS and IV estimates is irrelevant for the key economic interpretation of our results. As discussed above, in Appendix Section F.1.2, we present evidence that the patterns described by our IV estimates are robust to violations of the exclusion restriction. That is, the presence of a sizable positive causal effect of marriage on the wages of men and the emergence of an analogous effect on the wages of women exists even when the exclusion restriction is relaxed.

4 The Effect of Marriage along the Wage Distribution

One of the most relevant economic changes of the last decades has been the increase in income inequality.[24] At the same time, the divergence in marriage rates along important determinants of income has also increased. For example, Lundberg, Pollak, and Stearns (2016) document that, while the marriage rate was virtually the same across education groups up until the mid 1980s, nowadays collage graduates are significantly more likely to be married than people with a high school diploma. These two trends naturally lead to investigate how heterogeneous is the effect of marriage along the wage distribution and its evolution over the last decades. We do so in this section.

[22]Bursztyn, Fujiwara, and Pallais (2017) study the behavior of MBA students and find that single females express less willingness to conform with the demands of high paying jobs in front of single male peers. Taylor, Hart, Smith, Whalley, Hole, Wilson, and Deary (2005) show that IQ at age 11 is negatively associated with the probability of being married during adulthood for women.

[23]Folke and Rickne (2020) show that, in Sweden, women that are elected for public office, which under certain conditions can be understood as an exogenous positive shock to earnings, are more likely to divorce.

[24]See Autor (2014) and Guvenen, Kaplan, Song, and Weidner (2017).

4.1 Empirical Specification

We estimate the causal effect of marriage along the unconditional wage distribution, the unconditional IV quantile tratment effects (IVQTEs), using the Generalized Quantile Regression approach of Powell (Forthcoming).[25] We follow Powell (Forthcoming) and specify a rank variable $U_i^* \sim U(0,1)$, an unknown, unspecified function of observed (X_i) and unobserved factors (U_i) such that $U_i^* = f(X_i, U_i)$. We can write our outcome variable as a function of our treatment variable as $Y_i = M_i{}'\alpha(U_i^*), \quad U_i^* \sim U(0,1)$. Our aim is to estimate the Structural Quantile Function (SQF):

$$S_Y(\tau|m) = m'\alpha(\tau). \tag{7}$$

The SQF defines the τ^{th} quantile of the outcome distribution given the treatment if each individual in the sample had $M = m$. The quantile function is written as $M'\alpha(\tau)$. The GQR estimator uses the following moment conditions:

$$E\left\{Z_i\left[\mathbb{1}(Y_i \leq M_i'\alpha(\tau)) - \tau_X\right]\right\} = 0, \tag{8}$$

$$E\left[\mathbb{1}(Y_i \leq M_i'\alpha(\tau)) - \tau\right] = 0, \tag{9}$$

where τ_X represents $P(Y_i \leq M_i'\alpha(\tau) \mid X_i)$ and is estimated by:

$$\tau_X = F(X_i\delta(\tau)). \tag{10}$$

To be clear, covariates do not enter in Equation 7, the quantile functions are unconditional. Rather, covariates are used to determine the probability that log wages are below the quantile function given the covariates, as seen in Equation 8.[26]

4.2 Results

Figure 3 presents the IVQTEs for men and women divided in the six periods that cover all our sample years. For each gender-period, we compute the causal effect of marriage on wages at the 10th, 25th, 50th, 75th, and 90th percentiles of the unconditional wage distribution. For men at the beginning of the sample period (panel 1977-1983 in Figure 3a), the effect of marriage on wages monotonically decreases as we move up the unconditional wage distribution. While for men below the median marriage increases wages by around 40%, the effect is halved at the top of the distribution. This pattern is fairly consistent over time although the magnitude of the difference decreases. In the last period, 2012-2018 (Figure 3b) there are virtually no differences in the effect of marriage on wages at different points of the wage distribution. In the case of women, we observe a similar evolution. However, the starting point is the opposite to the case of men. In the early periods (Figure 3c) women below the median experience a bigger marriage penalty that women above. For women at the 90th percentile, the wage premium emerges in

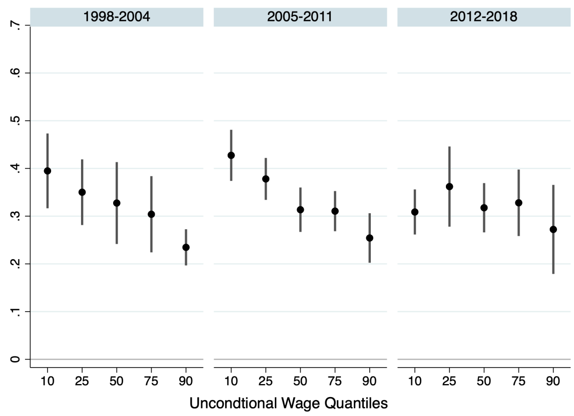
[25]This approach is a useful synthesis of (conditional) IVQR method of Chernozhukov and Hansen (2006) and the Recentered Influence Function (RIF) approach of Firpo, Fortin, and Lemieux (2009), which estimates unconditional QTEs but does not permit the use of IV methods. The utility of using unconditional quantiles is notable in our setting given that the wage distribution has changed remarkably over the last decades.

[26]Covariates are additionally used in our setting for the purposes of identification. Given that our instrument is constructed conditional on individual realizations of education, age and children, it is likely not unconditionally exogenous, but is conditionally exogenous.

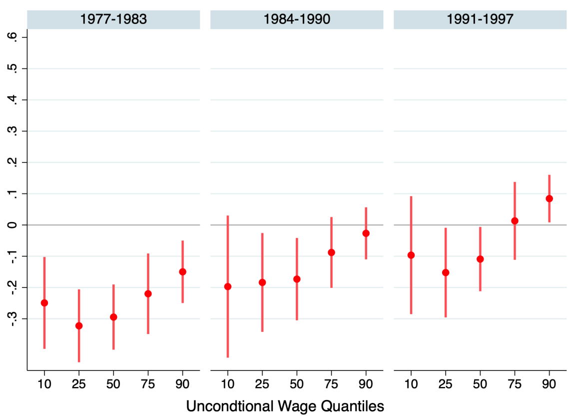
the mid 1990s, while the average effect in this period is still non-significant on average (Table 5). Conversely, for women at the 10th and 25th percentiles there is no conclusive evidence of the emergence of a positive effect of marriage on wages. Although the point estimates in the last period are positive (panel 2012-2018 in Figure 3d), their 95% confidence intervals include 0.

Figure 3: IVQTE

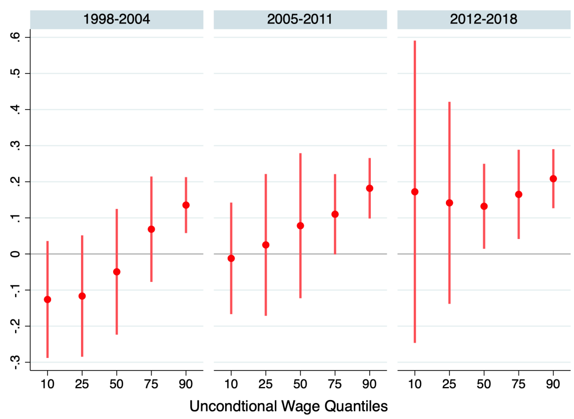
(a) Men - IV - Periods 1-3.



(b) Men - IV - Periods 4-6.



(c) Women - IV - Periods 1-3.



(d) Women - IV - Periods 4-6.

Notes: The figures present IVQTE estimates, and 95% confidence intervals based on state-clustered standard errors. These IVQTE estimates were produced using the Generalized Quantile Regression approach of Powell (Forthcoming). The dependent variable in all columns is the natural log of wages. The proneness variables - variables used to predict the probability that the log wage is below the quantile function - include: year and state fixed effects, dummies for highest level of educational attainment, dummies for year of potential experience, the number of children below the age of 5, the number of children aged 5-17, and a dummy for a child over the age of 18. The instrument in all specification is the proportion of individuals' reference group that are married interacted with a full set of state dummies. For the numerical optimization we used an adaptive MCMC optimization procedure, with a Metropolis-within-Gibbs sampler, 20,000 draws, a burn rate of .30, an acceptance rate of .50. The retained draws were jumbled to reduce autocorrelation between draws. Data used: CPS, 1977-2018.

5 Discussion

Our results can be summarized as follows. Nowadays, marriage causally increases the wage of the average men and the average women. While for the average men this considerable positive effect has barely reduced over the last decades, the causal impact of marriage on the wages of women has evolved from a penalty to the emergence of a sizable premium in the mid 2000s. These patterns are not uniform across the wage distribution. For men below the median wage, the marriage premium has reduced significantly even though it remains positive and sizable. Conversely, for men above the median wage, the effect has stayed fairly constant. In the case of women, the emergence of a premium happened earlier for women above the median wage than for the average. For women below the median, the effect of marriage on wages is unlikely to be negative. However, our estimates are too noisy to conclude that a premium exists.

These facts naturally lead to assess if within household specialization, the main mechanism discussed in the literature to account for the marriage wage premium of men, is consistent with the data. The idea behind this mechanism is that married men are able to put more effort into their job/career than their single counterparts because their wives specialize in non-market work. It follows that the opposite is true for women. Hence, we should observe a wage penalty for married women due to household specialization. Our results regarding the average effect of marriage on the wages of men and women (Figure 2) indicate that, since the appearance of the marriage wage premium for women, the within-household-specialization mechanism cannot be the main driver of the effect of marriage on wages. One possibility is that this mechanism was responsible for the male premium when there was a penalty for women. However, the MWP of men seems to be unaffected by the emergence of a premium for women. Hence, it seems implausible for the within-household-specialization mechanism to be the principal factor behind the male premium.

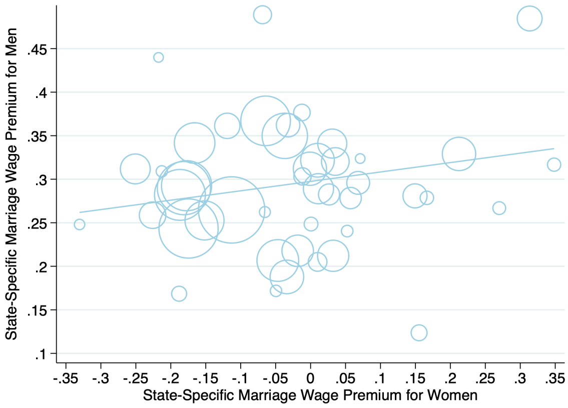
The patterns of the effect of marriage on wages across the quintiles of the wage distribution (Figure 3) provide the best case scenario for the relevance of the within-household-specialization mechanism. Taking into account the patterns of assortative mating, it is reasonable to assume that men below the median wage are more likely to be married to women below the median wage than to women above. For below-median-wage individuals, we observe that the reduction of the female penalty happens as the male premium decreases. If within household specialization plays a role in the determination of wages for below-median-wage individuals, this pattern indicates that the reduction in the female penalty and the male premium could have occurred due to a decrease in the extent to which these individuals engage in within household specialization. However, in the mid 2000s, the penalty for women below the median wage vanishes while a sizable premium remains for men below the median wage. That is, at best, the within-household-specialization mechanism is one of different channels that partially explain the marriage premium of men below the median wage.

In order to further test the relevance of the within-household-specialization mechanism, we use our IV estimates to compute state-specific coefficients across time for both men and women.[27] We then look at the correlation between the causal effect of marriage for men and women across states and time. If the within-household-specialization mechanism is a primary factor behind the returns of marriage, we expect a negative relationship between the effect of marriage on the wages of men and women across states. Figure 4 presents these correlations grouping together the years available in our sample in three sub-periods of fourteen years each. For all sub-periods,

[27]Specifically we re-run our IV and IV-Heckman specifications for men and women respectively, allowing the coefficient on marriage to differ by state. This amounts to replacing (i) $\pi_1 Z_{M,i}$ with $\sum_s \pi_{1,s} Z_{M,i}$ in Equations 2 and 5, and (ii) αM_i with $\sum_s \alpha_s M_i$ in Equations 3 and 6. Naturally, s denotes state and we sum over all states.

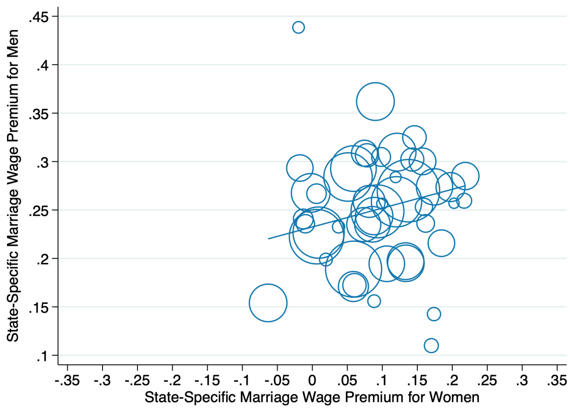
the correlation between the returns of marriage of men and women is positive. That is, in states in which the causal effect of marriage on the wages of women is bigger, the premium for men tends to be bigger too. Analogously to the discussion above, these patterns are at odds with the hypothesis that the within-household-specialization mechanism is the main driver of the returns to marriage.

Figure 4: The Relationship between the Causal MWP of Men and Women



(a) IV, 1977-1990.

(b) IV, 1991-2004.



(c) IV, 2005-2018.

Notes: The figures plot the coefficient on a dummy for marriage interacted with a full set of state dummies from an IV and IV-Heckman regression for men and women respectively. Each state is weighted by the combined (CPS-weighted) population. 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, dummies for year of potential experience, the number of children below the age of 5, the number of children aged 5-17, and a dummy for a child over the age of 18. The instrument in all specification is the proportion of individuals' reference group that are married interacted with a full set of state dummies. The exclusion restrictions for the employment equation are a series of dummies for age of youngest child in the household from 1-18, where age less that 1 is the base category, a dummy for ages 19-24, a dummy for 25 and over, and no children are also included. We estimate 51 state-specific treatment effects (50 states plus DC), which requires 51 first stage regressions. The IV was weak for some state-time period combinations. In addition, certain small states gave very erratic MWP estimates - something that did not occur when we estimated the wage equations by OLS. In order to present meaningful results, we plot state-specific MWP estimates subject to i.) a CPS-weighted population cut-off of 100,000 and ii.) a cut-off for Shea's Adjusted Partial R^2 of 0.07. The patterns presented are robust to the precise cut-offs implemented. Data used: CPS, 1977-2018.

Recent work by Pilossoph and Wee (2019) hypothesizes that the marriage wage premium for men and women is a byproduct of joint search behavior within the household. Due to the income pooling that takes place in married households, married individuals have a higher reservation wage than their single counterparts. This higher reservation wage leads to matching with jobs that pay higher wages. Moreover, married people are more willing to climb the job ladder as their success reinforces the higher reservation wage of their spouse. These mechanisms are consistent with the causal patterns we find in the later years of our sample. In particular for individuals above the median wage. However, they raise the question of why married people have significantly lower unemployment rates than their single counterparts.[28] Specifically, the non-employment rate of married men is lower than that of singles, while single and married women present similar non-employment rates.

Another hypothesis discussed in the literature 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. In Appendix E, we adapt the approaches of Altonji and Pierret (2001) and Pinkston (2009) to the case of marriage. Using the NLSY79, we find no evidence of marriage being used to positively discriminate men nor women. That is, employer statistical discrimination is not a relevant factor to explain the positive relationship between marriage and wages. However, we do find evidence that the marriage wage premium of women is composed of an initial penalty, when married women enter the labor market, which evolves into a premium as married women gain experience, at least for the cohort in the NLSY79.

Overall, there exist no unifying theory or mechanism that is able to account for the causal effect of marriage on wages. Some of the mechanisms discussed in the literature, within household specialization and the implications of joint household search, provide partial explanations. The limited scope of the existing theories, along with our results, indicate that the effect of marriage on wages is likely to be driven by multiple mechanisms, many of which have not been uncovered yet.

6 Conclusions

In this paper, we establish that marriage causes a higher wage for men and women. While the male premium has existed for years, the causal impact of marriage on the wages of women has evolved from a penalty to a sizable positive effect in the last decades. Our results indicate that the effect of marriage on wages is heterogeneous, in particular along the wage distribution. We highlight that there is no unifying theory that explains the existence of the marriage premium and its evolution over the last decades. Our paper provides the main facts that such a theory needs to account for. Understanding the different mechanisms that generate the marriage wage premium is crucial to understand the social and economic changes of the last decades. Moreover, it is relevant for the design of policies that rely on household composition such as taxation or welfare subsidies.

[28]See Choi and Valladares-Esteban (2018).

References

Effrosyni Adamopoulou. Peer Effects in Young Adults Marital Decisions. Working Paper 12-28, Universidad Carlos III de Madrid - Departamento de Economía, October 2012.

Effrosyni Adamopoulou and Ezgi Kaya. Young Adults Living with their Parents and the Influence of Peers. *Oxford Bulletin of Economics and Statistics*, 80(3):689–713, June 2018.

Avner Ahituv and Robert I. Lerman. How Do Marital Status, Work Effort, and Wage Rates Interact? *Demography*, 44(3):623–647, August 2007.

Joseph G. Altonji and Charles R. Pierret. Employer Learning and Statistical Discrimination. *Quarterly Journal of Economics*, 116(1):313–350, February 2001.

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 Papers and Proceedings*, 94(2): 317–321, May 2004.

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.

David H. Autor. Skills, education, and the rise of earnings inequality among the "other 99 percent". *Science*, 344(6186):843–851, May 2014.

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.

Elena Bardasi and Mark Taylor. Marriage and Wages: A Test of the Specialization Hypothesis. *Economica*, 75(299):569–591, July 2008.

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.

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

Victor Chernozhukov and Christian Hansen. Instrumental Quantile Regression Inference for Structural and Treatment Effect Models. *Journal of Econometrics*, 132(2):491 – 525, June 2006.

Sekyu Choi and Arnau Valladares-Esteban. The Marriage Unemployment Gap. *B.E. Journal of Macroeconomics*, 18(1), January 2018.

Sekyu Choi and Arnau Valladares-Esteban. On Households and Unemployment Insurance. *Quantiative Economics*, 11(1):437–469, January 2020.

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.

Timothy G. Conley, Christian B. Hansen, and Peter E. Rossi. Plausibly Exogenous. *Review of Economics and Statistics*, 94(1):260–272, February 2012.

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. *Quarterly Journal of Economics*, 129(4):1711–1752, November 2014.

Scott Drewianka. Estimating Social Effects in Matching Markets: Externalities in Spousal Search. *Review of Economics and Statistics*, 85(2):409–423, May 2003.

Christian Dustmann and María Engracia Rochina-Barrachina. Selection Correction in Panel Data Models: An Application to the Estimation of Females' Wage equations. *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.

Raquel Fernández and Alessandra Fogli. Culture: An Empirical Investigation of Beliefs, Work, and Fertility. *American Economic Journal: Macroeconomics*, 1(1):146–77, January 2009.

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.

Olle Folke and Johanna Rickne. All the Single Ladies: Job Promotions and the Durability of Marriage. *American Economic Journal: Applied Economics*, 12(1):260–87, January 2020.

Donna K. Ginther and Marianne Sundström. Does Marriage Lead to Specialization? An Evaluation of Swedish Trends in Adult Earnings Before and After Marriage. Mimeo, July 2010.

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? *Journal of Human Resources*, 32(3):481–504, June 1997.

Jeremy Greenwood, Nezih Guner, Georgi Kocharkov, and Cezar Santos. Technology and the Changing Family: A Unified Model of Marriage, Divorce, Educational Attainment, and Married Female Labor-Force Participation. *American Economic Journal: Macroeconomics*, 8(1): 1–41, January 2016.

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.

Fatih Guvenen, Greg Kaplan, Jae Song, and Justin Weidner. Lifetime Incomes in the United States over Six Decades. Working Paper 23371, National Bureau of Economic Research, April 2017.

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.

Martha S. Hill. The Wage Effects of Marital Status and Children. *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.

Guido W. Imbens and Donald B. Rubin. *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction.* Cambridge University Press, 2015.

Chinhui Juhn and Kristin McCue. Evolution of the Marriage Earnings Gap for Women. *American Economic Review*, 106(5):252–56, May 2016.

Chinhui Juhn and Kristin McCue. Specialization Then and Now: Marriage, Children, and the Gender Earnings Gap across Cohorts. *Journal of Economic Perspectives*, 31(1):183–204, February 2017.

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, 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? *Journal of Human Resources*, 26(2):282–307, April 1991.

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

Harry A. Krashinsky. Do Marital Status and Computer Usage Really Change the Wage Structure? *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. *Journal of Human Resources*, 31(3):566–589, July 1996.

David S. Loughran and Julie M. Zissimopoulos. Why Wait? The Effect of Marriage and Childbearing on the Wages of Men and Women. *Journal of Human Resources*, 44(2):326–349, Spring 2009.

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.

Shelly Lundberg, Robert A. Pollak, and Jenna Stearns. Family Inequality: Diverging Patterns in Marriage, Cohabitation, and Childbearing. *Journal of Economic Perspectives*, 30(2):79–102, May 2016.

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

Ismael Mourifié and Aloysius Siow. The Cobb Douglas marriage matching function: Marriage matching with peer and scale effects. Unpublished Manuscript, November 2017.

Casey B. Mulligan and Yona Rubinstein. Selection, Investment, and Women's Relative Wages Over Time. *Quarterly Journal of Economics*, 123(3):1061–1110, August 2008.

Aviv Nevo and Adam M. Rosen. Identification With Imperfect Instruments. *Review of Economics and Statistics*, 94(3):659–671, August 2012.

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.

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. *Review of Economic Studies*, 76(1):367–394, January 2009.

David Powell. Quantile Treatment Effects in the Presence of Covariates. *Review of Economics and Statistics*, Forthcoming.

Ferdinando Regalia, José-Víctor Ríos-Rull, and Jacob Short. What Accounts for the Increase in the Number of Single Households? Unpublished Manuscript, September 2011.

Alejandrina Salcedo, Todd Schoellman, and Michèle Tertilt. Families as roommates: Changes in U.S. household size from 1850 to 2000. *Quantitative Economics*, 3(1):133–175, March 2012.

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.

Michelle D. Taylor, Carole L. Hart, George Davey Smith, Lawrence J. Whalley, David J. Hole, Valerie Wilson, and Ian J. Deary. Childhood IQ and marriage by mid-life: the Scottish Mental Survey 1932 and the Midspan studies. *Personality and Individual Differences*, 38(7):1621 – 1630, May 2005.

Hans van Kippersluis and Cornelius A. Rietveld. Beyond Plausibly Exogenous. *Econometrics Journal*, 21(3):316–331, October 2018.

Alexander Vickery and Dan Anderberg. The Role of Own-Group Density and Local Social Norms for Ethnic Marital Sorting: Evidence from the UK. Unpublished Manuscript, May 2019.

Appendix

A Alternative Definition of the Non-married Group

As discussed in Section 2, the literature defines the MWP as the difference in wages between married individuals and those who are never married. That is, the divorced and separated are not included in the non-treated group. As the large literature on the determinants of marriage dissolution indicates, divorce and separation are endogenous.[29] Hence, an analysis aimed to understand the difference in wages between married and divorced/separated people needs to address not only the endogeneity of marriage formation but also that of marriage dissolution. Moreover, if marriage has persistence effects on wages, a comparison between married and divorced/separated is unable to identify these effects as the divorced/separated have also been exposed to the treatment. For these reasons, our main analysis follows the literature's definition of the non-married group as those that are never married. Nonetheless, in this section, we show that our main results are robust to the inclusion of divorced and separated individuals in the non-married group.

In Table A1 we present the OLS coefficients associated to marriage for men and women in all periods. Tables A2 and A3 replicate the IV estimates of Section 3.3. In general, the estimates display the same qualitative patterns as those that use only the never married in the non-married group. Namely, the MWP for men is sizable and virtually constant over time and the relationship between marriage and wages evolves from negative to significantly positive for women. The coefficients are also similar in magnitude.

Table A1: OLS

A. Men	**(1)** **'77-'83**	**(2)** **'84-'90**	**(3)** **'91-'97**	**(4)** **'98-'04**	**(5)** **'05-'11**	**(6)** **'12-'18**
Married	0.171***	0.174***	0.182***	0.182***	0.190***	0.173***
	(0.004)	(0.005)	(0.005)	(0.006)	(0.006)	(0.005)
Adjusted R^2	0.202	0.237	0.257	0.242	0.260	0.258
Observations	125,563	126,428	118,759	136,500	141,399	111,638
B. Women	**'77-'83**	**'84-'90**	**'91-'97**	**'98-'04**	**'05-'11**	**'12-'18**
Married	-0.034***	-0.009*	0.021***	0.040***	0.049***	0.066***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.004)	(0.004)
Adjusted R^2	0.157	0.200	0.231	0.228	0.228	0.233
Observations	97,810	112,059	112,605	133,807	139,766	107,669

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 in 1999 Dollars. Year and state fixed effects are included in all regressions. The following additional controls are included: dummies for highest level of educational attainment, dummies for year of potential experience, the number of children below the age of 5, the number of children aged 5-17, and a dummy for a child over the age of 18. Data used: CPS, 1977-2018.

[29] See Stevenson and Wolfers (2007).

Table A2: IV - Men

	'77-'83	**'84-'90**	**'91-'97**	**'98-'04**	**'05-'11**	**'12-'18**
Married	0.277***	0.265***	0.246***	0.229***	0.281***	0.231***
	(0.011)	(0.014)	(0.014)	(0.017)	(0.018)	(0.021)
First-Stage F-Statistic	2524.7	1989.5	1222.2	1341.4	1286.0	1125.2
Adjusted R^2	0.193	0.232	0.254	0.243	0.257	0.255
Observations	113,697	116,145	108,788	127,451	135,164	106,115

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, dummies for year of potential experience, the number of children below the age of 5, the number of children aged 5-17, and a dummy for a child over the age of 18. The instrument in all specification is the proportion of individuals' reference group that are married. Data used: CPS, 1977-2018.

Table A3: IV and IV-Heckman - Women

A. IV	**'77-'83**	**'84-'90**	**'91-'97**	**'98-'04**	**'05-'11**	**'12-'18**
Married	-0.211***	-0.077	-0.084*	-0.053	0.079**	0.112***
	(0.048)	(0.068)	(0.050)	(0.062)	(0.033)	(0.039)
First-Stage F-Statistic	438.6	455.2	179.6	232.9	242.1	239.6
Adjusted R^2	0.132	0.198	0.228	0.224	0.226	0.228
Observations	84,423	100,574	101,329	122,693	131,391	101,102
B. IV-Heckman	**'77-'83**	**'84-'90**	**'91-'97**	**'98-'04**	**'05-'11**	**'12-'18**
Married	-0.196***	-0.073	-0.064	-0.052	0.106***	0.104**
	(0.043)	(0.055)	(0.047)	(0.057)	(0.040)	(0.045)
First-Stage F-Statistic	477.8	469.6	157.7	209.2	208.0	191.3
Adjusted R^2	0.136	0.199	0.231	0.224	0.225	0.228
Observations	84,423	100,574	101,329	122,693	131,391	101,102

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, dummies for year of potential experience, the number of children below the age of 5, the number of children aged 5-17, and a dummy for a child over the age of 18. The instrument in all specification is the proportion of individuals' reference group that are married. In panel B, we present selection-corrected estimates. The exclusion restrictions for the employment equation are a series of dummies for age of youngest child in the household from 1-18, where age less that 1 is the base category, a dummy for ages 19-24, a dummy for 25 and over, and no children are also included. In this case we bootstrap standard errors, allowing for clustering at the state level, and using 500 iterations. Data used: CPS, 1977-2018.

B Robustness Checks on the Main IV Specifications

In this section, we provide a series of robustness checks on the main specification we use in Section 3 to estimate the causal effect of marriage on wages. In Table B1 we present the IV coefficients of Tables 3 (Men IV), 4 (Women IV), and 5 (Women IV-Heckman) along with relevant changes in the set of controls included in the wage regression. For the two specifications that do not include a selection-into-employment correction (Panels A and B of Table B1), we show that the quantitative and qualitative patterns described by our estimates in Section 3.3 are robust to including industry and occupation controls. We are unable to replicate these when we correct for selection-into-employment as industry and occupation are undefined for those who are not working. The estimates in Table B1 also establish that our results, in all three specifications, are robust to including state-year controls.

Table B1: Robustness Checks - Extended 2SLS Specifications

A. Men: IV	**'77-'83**	**'84-'90**	**'91-'97**	**'98-'04**	**'05-'11**	**'12-'18**
Baseline	0.306***	0.269***	0.246***	0.226***	0.266***	0.237***
	(0.013)	(0.015)	(0.013)	(0.016)	(0.017)	(0.021)
Baseline +	0.274***	0.227***	0.214***	0.196***	0.230***	0.199***
Industry Groups	(0.011)	(0.012)	(0.013)	(0.016)	(0.017)	(0.021)
Baseline +	0.284***	0.250***	0.227***	0.216***	0.257***	0.233***
Occupation Groups	(0.012)	(0.014)	(0.012)	(0.015)	(0.016)	(0.020)
Baseline +	0.251***	0.215***	0.201***	0.189***	0.224***	0.200***
Industry & Occupation	(0.011)	(0.012)	(0.012)	(0.015)	(0.016)	(0.020)
Baseline +	0.305***	0.271***	0.247***	0.224***	0.266***	0.238***
State $\times$ Year FEs	(0.013)	(0.015)	(0.012)	(0.016)	(0.017)	(0.021)
Observations	104,970	104,545	96,210	112,807	120,606	94,896
B. Women: IV						
Baseline	-0.159***	-0.061	-0.030	-0.034	0.078***	0.088***
	(0.025)	(0.045)	(0.029)	(0.044)	(0.027)	(0.032)
Baseline +	-0.138***	-0.063*	-0.038	-0.037	0.040	0.051*
Industry Groups	(0.026)	(0.037)	(0.028)	(0.040)	(0.028)	(0.029)
Baseline +	-0.166***	-0.070*	-0.044*	-0.046	0.058**	0.090***
Occupation Groups	(0.023)	(0.041)	(0.025)	(0.039)	(0.027)	(0.031)
Baseline +	-0.136***	-0.064*	-0.042*	-0.039	0.034	0.063**
Industry & Occupation	(0.023)	(0.035)	(0.024)	(0.037)	(0.028)	(0.029)
Baseline +	-0.158***	-0.066	-0.035	-0.040	0.079***	0.088***
State $\times$ Year FEs	(0.025)	(0.044)	(0.028)	(0.042)	(0.029)	(0.032)
Observations	70,872	83,482	83,791	101,175	109,624	85,427
C. Women: IV-H						
Baseline	-0.149***	-0.062	-0.021	-0.030	0.103***	0.082**
	(0.024)	(0.043)	(0.029)	(0.044)	(0.031)	(0.033)
Baseline +	-0.146***	-0.066	-0.024	-0.036	0.103***	0.081**
State $\times$ Year FEs	(0.024)	(0.043)	(0.029)	(0.042)	(0.032)	(0.035)
Observations	70,800	82,851	8.3e+04	100,448	109,456	8.5e+04

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, dummies for year of potential experience, the number of children below the age of 5, the number of children aged 5-17, and a dummy for a child over the age of 18. The instrument in all specification is the proportion of individuals' reference group that are married. For the IV-Heckman specifications, the exclusion restrictions for the employment equation are a series of dummies for age of youngest child in the household from 1-18, where age less that 1 is the base category, a dummy for ages 19-24, a dummy for 25 and over, and no children are also included. In this case we bootstrap standard errors, allowing for clustering at the state level, and using 500 iterations. Additional industry control variables are comprised of a set of 14 industry group dummies. Additional occupational controls are comprised of a set of 7 industry group dummies. Data used: CPS, 1977-2018.

C The Within-Individual Marriage Wage Premium

In this section, we provide a thorough comparison between our main results, based on the CPS, and their counterparts in the NLSY79. Given that the NLSY79 is widely used in the MWP literature, the results in this section provide a comprehensive account of the divergences in inference that might steam from inherent differences in the two data sets. In particular, we show how commonly used sample restrictions in the literature lead to samples with significant differences in observable characteristics between the two data sets. We compare different methods to address these differences and we assess how they affect our results. Finally, we estimate the effect of marriage on wages using a fixed effects framework on our NLSY79 sample. The fixed-effects marriage coefficients are positive and significant for both men and women albeit smaller than the IV coefficients we obtain from the CPS. We discuss the possible origin of this discrepancy and the fundamental differences between estimating the effect of marriage from the between-individual variation of the CPS and the within-individual variation of the NLSY79.

C.1 The NLSY79 Sample

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 years 1957-1964. As it is common in the literature, our estimates do not use the NLSY weights.

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. 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 NLSY79 sample consists of 1,446 men and 1,121 women which correspond to 14,736 and 10,375 observations respectively. Table C1 presents descriptive statistics related to marriage, wages, tenure, experience, age, education, and the number of children.

Table C1: Descriptive Statistics, NLSY79

Means, Standard Deviations in Parentheses

	(1) Men	(2) 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, Aged 0-4	0.292	0.235
	(0.593)	(0.528)
Number Children, Aged 5-17	0.232	0.206
	(0.614)	(0.565)
Children, 18 and over	0.005	0.008

Notes: Data used: NLSY79, 1979-2012.

Although we apply seemingly similar sample restrictions to construct our CPS and NLSY79 samples, a comparison between Tables 1 and 2 and Table C1 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 focuses on one cohort and some of covariates we use are differently measured between the CPS and the NLSY79. For example, in the NLSY79, we can compute actual experience while we relay on potential experience in the CPS. 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.

C.2 Bridging across the Two Samples

We balance two criteria when selecting variables and making sample restrictions. First, we are guided by the sample restrictions and variable definitions 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 and variable-selection 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 1977-2018 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. Thus, when comparing to the time periods into which we split the CPS, the NLSY79 median falls exactly between the second (1984-1990) and third (1991-1997) periods we consider. This is especially relevant for the results of women, given the changes in the female MWP that we document in 1b.

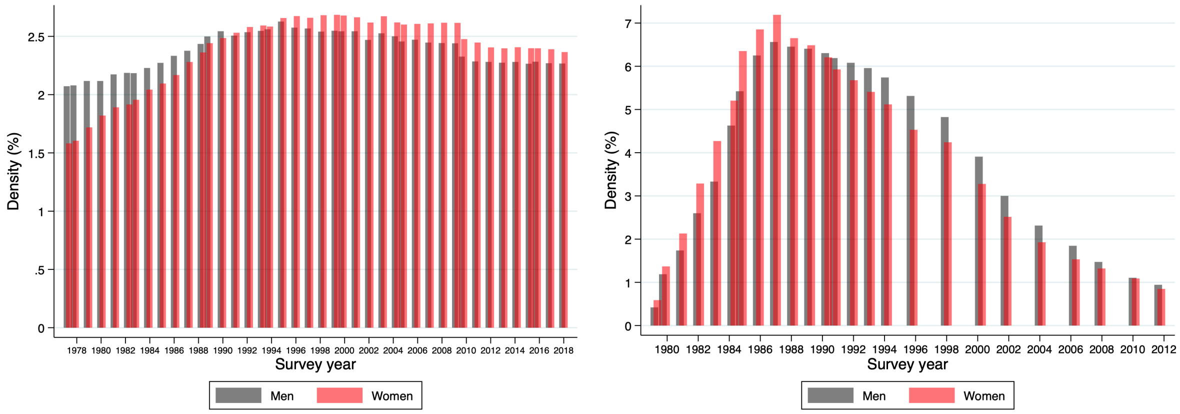
In order to bridge the differences between the two data sources, we perform two exercises. First, we use the data in the CPS to construct a set of samples which contain individuals with similar observable characteristics to the individuals in our NLSY79 sample. We employ different matching approaches 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. We use these pseudo-NLSY samples to run the same main analysis as on our CPS sample. 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.

Secondly, we apply the definitions of the covariates of the CPS to our NLSY79 sample. For example, in our main analysis, whenever we use the CPS sample we have to use potential experience as a proxy for actual experience while, in the NLSY79, we do observe actual experience. This exercise reveals that the distinct variable definitions do not alter substantially the direction of the results in the main analysis.

C.2.1 Creation of the pseudo-NLSY 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. Figure C1 provides an overview of the difference in survey years covered between the CPS and the NLSY.

Figure C1: Year Distribution - CPS and NLSY79 Comparison



(a) CPS.

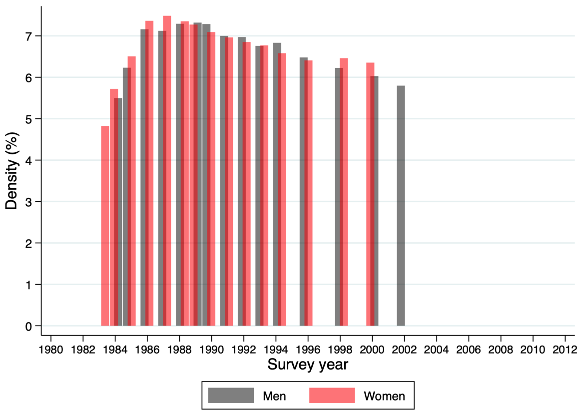
(b) NLSY.

Notes: Data used: CPS, 1977-2018 and NLSY79, 1979-2012.

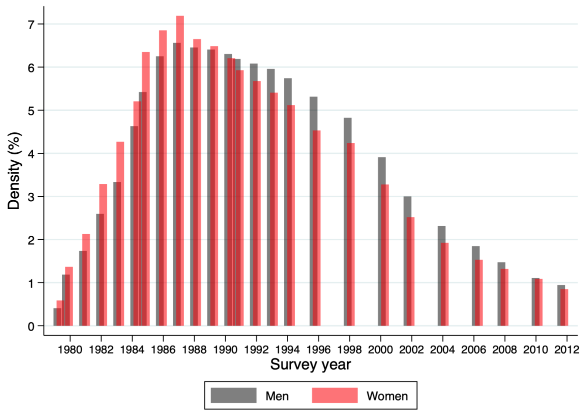
We construct four different pseudo-NLSY samples. In the first sample, which we name *Simple*, we focus only on selecting observations from the CPS that match our NLSY79 sample in terms of surveyed year and the ages of the respondents. In this sense we are using the CPS to build a quasi-cohort, which mimics the NLSY79. 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 the 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.

Figure C2 show how the different matching techniques manage to replicate the survey year composition of the NLSY (Figure C1b). Table C2 presents summary statistics for the four pseudo-NLSY samples. Three differences stand out when comparing the descriptive statistics for our CPS sample (Tables 1 and 2) with those of our NLSY79 sample (Table C1). 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 C2 show that, except for our Simple pseudo-NLSY sample, these three differences are tackled with all matching procedures.

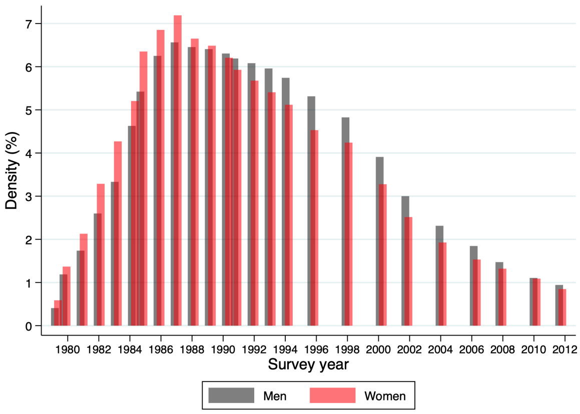
Figure C2: Year Distribution - Pseudo-NLSY samples



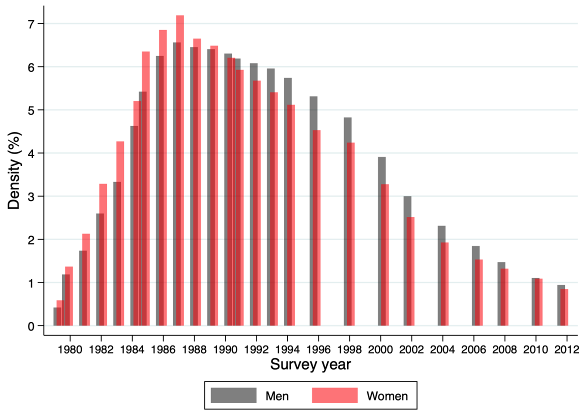
(a) Simple.



(b) NN(1).



(c) Kernel.



(d) Entropy Balance.

Notes: The details regarding the creation of the four pseudo-NLSY samples is outlined in Section C.2.1. Data used: CPS, 1979-2012 and NLSY79, 1979-2012.

Table C2: Descriptive Statistics, Pseudo-NLSY79

Means, Standard Deviations in Parentheses

	Simple		**NN(1)**		**Kernel**		**Entropy Balance**	
	(1) Men	(2) Women	(3) Men	(4) Women	(5) Men	(6) Women	(7) Men	(8) Women
Sample Size	77,232	66,514	78,818	62,099	115,026	104,004	115,810	104,721
Weighted Sample Size			14,734	10,375	14,734	10,375	14,736	10,375
Married	0.649	0.711	0.471	0.464	0.482	0.488	0.464	0.464
Hourly Wage (1999 Dollars)	16.36	12.63	15.81	13.23	15.72	13.12	15.67	13.17
	(8.87)	(7.35)	(9.06)	(7.72)	(8.94)	(7.64)	(8.96)	(7.71)
Age	31.09	30.06	30.80	30.17	30.85	30.19	30.80	30.17
	(5.47)	(5.12)	(6.62)	(6.59)	(6.70)	(6.65)	(6.62)	(6.59)
Highest Level of Education:								
HS Dropout	0.082	0.052	0.076	0.021	0.076	0.024	0.079	0.022
HS Graduate	0.380	0.372	0.429	0.372	0.416	0.355	0.434	0.375
Some College	0.245	0.270	0.188	0.234	0.206	0.251	0.189	0.233
College Graduate	0.205	0.225	0.209	0.260	0.211	0.269	0.204	0.259
Advanced Graduate	0.088	0.080	0.098	0.113	0.091	0.101	0.094	0.111
Number Children, 0-4	0.393	0.357	0.289	0.229	0.299	0.243	0.292	0.235
	(0.667)	(0.620)	(0.590)	(0.513)	(0.609)	(0.540)	(0.605)	(0.536)
Number Children, 5-17	0.549	0.595	0.226	0.203	0.255	0.247	0.235	0.207
	(0.932)	(0.944)	(0.613)	(0.558)	(0.653)	(0.639)	(0.615)	(0.569)
Children, 18 and over	0.020	0.023	0.002	0.003	0.010	0.015	0.005	0.008

Notes: Data used: CPS, 1979-2012.

C.2.2 Key Results for the Pseudo-NLSY Samples

In Table C3 we replicate the results of Tables 3 (Men IV), 4 (Women IV).[30] Note that in these pseudo-NLSY samples, the survey years correspond to those in the NLSY79 sample, hence, we are unable to perform any analysis across time. The results from Table C3 show that, for both genders, the effect of marriage on wages is similar across pseudo-NLSY samples and in line with our estimates in Section 3.3.

Table C3: 2SLS Results for Pseudo-NLSY Samples

	Simple		**NN(1)**		**Kernel**		**Entropy Balance**	
	(1) Men	(2) Women	(3) Men	(4) Women	(5) Men	(6) Women	(7) Men	(8) Women
Married	0.231***	-0.006	0.195***	-0.002	0.227***	-0.065*	0.227***	-0.060*
	(0.017)	(0.036)	(0.022)	(0.043)	(0.014)	(0.034)	(0.014)	(0.034)
First-Stage F-Statistic	2759.7	389.6	1095.4	227.5	2582.7	331.6	2595.5	330.3
Adjusted R^2	0.221	0.232	0.227	0.222	0.250	0.234	0.250	0.235
Observations	66,373	52,649	61,926	46,293	92,431	80,811	9.3e+04	81,021

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, dummies for year of potential experience, the number of children below the age of 5, the number of children aged 5-17, and a dummy for a child over the age of 18. The instrument in all specification is the proportion of individuals' reference group that are married. The construction of the four pseudo-NLSY samples on which these regressions are run is outlined in Section C.2.1. Data used: CPS, 1979-2012.

[30] Recall that one of the restrictions applied to the NLSY sample is that wages are observed in at least two periods. Hence, it is unfeasible to apply the selection-into-employment correction.

C.2.3 Equivalent Covariates Definitions

The differences in covariates between the NLSY79 and the CPS stem from the fact that in the NLSY79, given its panel structure, it is possible to measure actual work experience and tenure on the job. However, in the CPS, there is no measure of neither actual experience nor tenure which is consistently available for the base sample throughout the period we analyze. Hence, the wage equations we run on the CPS sample use potential experience as a covariate while the wage regressions we estimate using the NLSY79 contain controls for both actual experience and tenure. A difference we cannot bridge is the one related to the spatial component of wages. In the specifications we run on the CPS sample, we control for the state in which the individual lives. We do not observe state in the NLSY79. Instead, we include a dummy for individuals living in urban areas.

Table C4 presents a comparison of estimates obtained using the definition of covariates based on the information available in the CPS (columns named *Restricted*) and the coefficients computed using the additional information in the NLSY79 (columns named *Baseline*). The columns named *None* provide the reference point of the unconditional correlation between wages and marriage. For both genders, the estimates based on the variable definitions of the CPS are larger than those obtained based on the (richer) variable definitions of the NLSY79. In particular, for men the pooled OLS estimate is around 27% larger while that of women is 15%. This is not surprising given the fact that the in the CPS experience can only be proxied while it can be directly measured in the NLSY79. However, both the order of the magnitude and the direction of the estimates are the same in both specifications.

Table C4: Pooled OLS Models, NLSY

	Men			Women		
Covariates:	(1) None	(2) Restricted	(3) Baseline	(4) None	(5) Restricted	(6) Baseline
Married	0.309***	0.155***	0.124***	0.177***	0.070***	0.058***
	(0.016)	(0.018)	(0.018)	(0.018)	(0.020)	(0.018)
Adjusted R^2	0.092	0.292	0.337	0.032	0.251	0.318
Observations	14,557	14,557	14,557	10,330	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 in 2006 Dollars. Pooled OLS model estimates for men are presented in Columns 1-3, and for women in Columns 4-6. Columns 1 and 4 present uncondtional estimates of the MWP. In Columns 2 and 5 the following controls are included: dummies for highest level of educational attainment, dummies for year of potential experience, the number of children below the age of 5, the number of children aged 5-17, a dummy for a child over the age of 18, a dummy for urban residence and year dummies. In Columns 3 and 6 - the final, baseline specification - the following controls are included: dummies for highest level of educational attainment, dummies for deciles of both actual experience and tenure, the number of children below the age of 5, the number of children aged 5-17, a dummy for a child over the age of 18 and a dummy for urban residence. Data used: NLSY79, 1979-2012.

C.3 A Fixed Effects Framework

In this section, we use our NLSY sample to estimate the impact of marriage on wages using a fixed effects (FE) model. When a FE model is used to assess the effect of marriage on wages, one of the identifying assumptions is that the treatment of interest starts at the beginning of being marriage. That is, the effect of marriage is brought about by a within-individual comparison of wages before and after the time in which the individual becomes married. We discuss the implications of this assumption in Section C.3.2.

C.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 a 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.[31] 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.

C.3.2 Results

In Table C5, we present the coefficients 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. As discussed above part of the discrepancy between the CPS OLS coefficients and the NLSY Pooled OLS coefficient is rooted in structural

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

difference between the two data sets. In turn, the FE estimates are lower than the Pooled OLS. That is, a significant part of the relationship between marriage and wages can be accounted for unobservable individual fixed effects. Given the contrast between these estimates and the IV coefficients we obtain in Section 3.3, it is worth focusing on one of the identifying assumptions in the FE framework. The FE specification extracts the effect of marriage on wages from comparing the wages of people that marry before and after marriage. First, this implies that the effect of marriage is estimated out of only people that are observed married. Second, if marriage has anticipation effects, these are confounded in the non-treated observations. A relevant example is that of an individual that cohabits before actually becoming legally married.

Table C5: Panel Data Models

	Men		**Women**	
	(1) Pooled OLS	**(2) Fixed Effects**	**(3) Pooled OLS**	**(4) Fixed Effects**
Married	0.124***	0.051***	0.058***	0.023*
	(0.018)	(0.012)	(0.018)	(0.013)
R^2	0.337	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 in 2006 Dollars. 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, the number of children below the age of 5, the number of children aged 5-17, a dummy for a child over the age of 18 and a dummy for urban residence. Data used: NLSY79, 1979-2012.

D The Role of Selection into Employment

In this section, we build the intermediate step between the OLS and the IV-Heckman correction. We estimate the selection-into-employment correction in isolation. For women, the association between marriage and hourly wages 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.

D.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:

$$E_i = \mathbb{1}\{\gamma_1 Z_{E,i} + \gamma_2 M_i + X_i^{'}\gamma_3 + \theta_{1s} + \phi_{1t} + \epsilon_{1i} > 0\} = \mathbb{1}\{Z_i^{'}\gamma + \epsilon_{1i} > 0\}, \tag{14}$$

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

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:

$$\begin{pmatrix} \epsilon_{1i} \\ \epsilon_{2i} \end{pmatrix} \sim N \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & 1 \end{pmatrix}. \tag{16}$$

We specify Heckman's two-step approach as:

$$E_i = \mathbb{1}\{\gamma_1 Z_{E,i} + \gamma_2 M_i + X_i^{'}\gamma_3 + \theta_{1s} + \phi_{1t} + \xi_i > 0\} = \mathbb{1}\{Z_i^{'}\gamma + \xi_i > 0\}, \tag{17}$$

$$y_i = \alpha M_i + X_i^{'}\beta + \theta_{2s} + \phi_{2t} + \sigma_{12}\lambda(Z_i^{'}\gamma) + \epsilon_i, \tag{18}$$

We start by estimating Equation 17 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 18 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 14 and 17) but not in the wage equation (Equations 15 and 18). 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 child in the household, where age less that 1 is the base category. A dummy for 19-24, 25 and over, and no children are also included.

D.2 Results

Table D1 presents the estimated married coefficients (α in Equations 15 and 18) along with the inverse Mills ratio associated with the first stage (Equations 14 and 17) for the six periods we consider. The patterns implied by the two estimation procedures are qualitatively identical. Namely, the married coefficient evolves from being negative in the 1977-1983 period to a null coefficient in the 1984-1990 period, while it is increasingly positive in the following years. Interestingly, the descriptive patterns computed by uncorrected OLS in Figure 1b are robust to correcting by selection into employment. Qualitatively, the selection patterns in our data are broadly consistent with those described by Mulligan and Rubinstein (2008). Our inverse Mills ratios also indicate that the selection of women into employment is no longer negative. While Mulligan and Rubinstein (2008) study full-time and full-year workers, our definition of employment includes more work arrangements and we use a different exclusion restriction. Hence, it is reasonable to expect weaker selection patterns in our analysis. In the last period of our sample, which was not available at the time of the study of Mulligan and Rubinstein (2008), we do not find evidence of positive selection.

Table D1: Sample Selection Models

A. Heckman's Two-Step	**'77-'83**	**'84-'90**	**'91-'97**	**'98-'04**	**'05-'11**	**'12-'18**
Married	-0.032***	-0.001	0.033***	0.052***	0.068***	0.080***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.006)	(0.006)
Inverse Mills	-0.100***	-0.060	0.150***	0.095**	0.091***	-0.025
Ratio	(0.027)	(0.054)	(0.036)	(0.039)	(0.032)	(0.060)
Adjusted R^2	0.155	0.199	0.233	0.228	0.227	0.232
Observations	82,262	93,080	93,059	110,225	116,569	90,920
B. Maximum Likelihood	**'77-'83**	**'84-'90**	**'91-'97**	**'98-'04**	**'05-'11**	**'12-'18**
Married	-0.038***	-0.002	0.032***	0.054***	0.068***	0.080***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.006)
Inverse Mills	-0.041***	-0.017	0.030***	0.021**	0.024***	-0.009
Ratio	(0.013)	(0.017)	(0.006)	(0.009)	(0.009)	(0.024)
Observations	82,262	93,080	93,059	110,225	116,569	90,920
C. OLS Estimate	-0.043***	-0.003	0.032***	0.054***	0.069***	0.080***
	(0.008)	(0.007)	(0.007)	(0.007)	(0.007)	(0.006)

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, dummies for year of potential experience, the number of children below the age of 5, the number of children aged 5-17, and a dummy for a child over the age of 18. The instrument in all specification is the proportion of individuals' reference group that are married. The exclusion restrictions for the employment equation are a series of dummies for age of youngest child in the household from 1-18, where age less that 1 is the base category, a dummy for ages 19-24, a dummy for 25 and over, and no children are also included. In row panel A, we implement Heckman's two-step method. In this case we bootstrap standard errors, allowing for clustering at the state level, and using 500 iterations. In row panel B we estimate both stages jointly via Maximum Likelihood. Data used: CPS, 1977-2018.

E Testing the Statistical Discrimination Hypothesis

Several key papers in the literature on the MWP hypothesize that a potential mechanism behind the higher wages of married men is positive employer discrimination.[32] The idea is that marriage might be positively related to variables that are relevant for productivity which are hard to observe by employers while marital status is easier to observe.[33] 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 workers' experience in the labor market increases, 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). The asymmetric information setting allows for a distinction between the learning done by the current employer, which occurs over the tenure of a job, and the public learning that happens through the overall experience of a worker in the labor market. In order to test the EL-SD hypothesis for the MWP we slightly modify our NLSY79 sample in order to be able to check for the presence of this mechanism in the data.

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. Without loss of generality, consider one easy-to-observe variable such as marital status (which might or might not affect productivity) and 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, i.e., 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. That is, the informational content of the easy-to-observe variable decreases and, hence, its return diminish.

E.1 Data

In order to test implications of the EL-SD models we consider, we modify our NLSY79 sample. We select the sample restrictions to balance two objectives. First, we follow the EL-SD literature

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

[33] 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.

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 E1 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 C1) 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 C1. Otherwise, most statistics in Table E1 are broadly in line with their counterparts in Table C1.

Table E1: Descriptive Statistics, EL-SD Sub-Sample
Means, Standard Deviations in Parentheses

	(1) Men	**(2) 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
Normalized AFQT	-0.000	0.000
	(1.00)	(1.00)
Urban Residence	0.719	0.724
Number Children, 0-4	0.453	0.419
	(0.695)	(0.636)
Number Children, 5-17	0.415	0.482
	(0.771)	(0.817)
Children, 18 and over	0.008	0.008

Notes: Data used: NLSY79, 1979-2012.

E.2 Public Learning

E.2.1 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: interactions of both marriage and time, and AFQT and time (in order to account for possible secular changes in returns to marriage and ability over time), 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.[34]

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.[35]

E.2.2 Results

Table E2 and Table E3 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. 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 the interaction between marriage and experience has to be negative.

For men (Table E2), 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% (0.221 in column (1) and 0.212 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 0.074 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 15% while each additional year of experience is associated with an increase of the premium of around 4 percentage points (about 1.5 for the OLS case). The coefficients from columns (3) and (6) indicate that the returns to cognitive ability also evolve over the working life. The fact that the returns to cognitive ability, the hard-to-observe variable, increase with experience are consistent with the presence of employer learning.[36] However, there is no evidence of employer discrimination as the returns to marriage also increase with experience instead of decreasing.

[34] As Altonji and Pierret (2001) point out, the need to include a rich set of time-dependent controls to control for the effect that secular changes in the variables of interest may have on wages implies that we exploit the variation in experience across age cohorts in the NLSY79, which is limited given the nature of the data. Hence, the precision of the estimates is bound to be affected.

[35] In the IV regression all instances of actual experience (polynomials and interaction terms) are instrumented.

[36] The caveat of this interpretation is the preciseness of the estimates. The rich set of time-related controls and the little variation in age-cohorts of the NLSY79 leads to imprecise estimates. We obtain more precisely estimated coefficients, of almost equal magnitude, when we use a more parsimonious set of controls.

Table E2: Symmetric SD-EL Model, Men

	OLS Potential Experience			IV Actual Experience		
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Married	0.221***	0.153***	0.154***	0.212***	0.146***	0.151***
	(0.025)	(0.037)	(0.037)	(0.026)	(0.039)	(0.038)
Married*Experience		0.015*	0.015*		0.039**	0.040**
		(0.008)	(0.008)		(0.018)	(0.018)
AFQT	0.077***	0.077***	0.037*	0.074***	0.075***	0.031
	(0.013)	(0.013)	(0.021)	(0.013)	(0.013)	(0.023)
AFQT*Experience			0.003			0.008
			(0.004)			(0.008)
Adjusted R^2	0.325	0.327	0.327	0.319	0.314	0.313
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 2006 Dollars. 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, the number of children below the age of 5, the number of children aged 5-17, and a dummy for a child over the age of 18, a dummy for urban residence and cubic polynomials in both time and experience. Columns 2, 3, 5 and 6 include both an interaction between the married dummy and experience, and an interaction between the married dummy and time. Columns 3 and 6 further include both an interaction between normalized AFQT and experience, and an interaction between normalized AFQT and time. 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. Data used: NLSY79, 1979-2012.

In the case of women (Table E3), 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 4-11% less than their single counterparts while an extra year of experience increases their wages by around 2-5% percentage points. The intercept of the returns to cognitive ability is positive and significant (0.076 in the OLS and 0.079 in the IV) while the interaction between the AFQT score and experience has a coefficient that is not statistically different from zero. As it is the case for men, positive employer discrimination seems not to be a driver of the marriage wage premium for women. The coefficient associated with the interaction between marriage and experience is 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 is 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 in equilibrium, but indicative of how a marriage wage premium for women can coexist with wage penalties based on traditional gender roles.

Table E3: Symmetric SD-EL Model, Women

	OLS Potential Experience			**IV Actual Experience**		
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Married	0.017	-0.108***	-0.110***	-0.004	-0.048	-0.049
	(0.030)	(0.037)	(0.037)	(0.031)	(0.039)	(0.039)
Married*Experience		0.021*	0.021*		0.049**	0.049**
		(0.012)	(0.012)		(0.024)	(0.024)
AFQT	0.092***	0.093***	0.076***	0.081***	0.081***	0.071***
	(0.013)	(0.013)	(0.020)	(0.013)	(0.013)	(0.019)
AFQT*Experience			0.004			0.003
			(0.004)			(0.012)
Adjusted R^2	0.259	0.263	0.264	0.293	0.285	0.286
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 2006 Dollars. 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, the number of children below the age of 5, the number of children aged 5-17, and a dummy for a child over the age of 18, a dummy for urban residence and cubic polynomials in both time and experience. Columns 2, 3, 5 and 6 include both an interaction between the married dummy and experience, and an interaction between the married dummy and time. Columns 3 and 6 further include both an interaction between normalized AFQT and experience, and an interaction between normalized AFQT and time. 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. Data used: NLSY79, 1979-2012.

E.3 Asymmetric Employer Learning

E.3.1 Empirical Specification

Formally, we extend the specification from Equation 19:

$$\begin{aligned} y_i = {} & \alpha_0 M_i + \alpha_1 (M_i \times x_i) + \alpha_2 (M_i \times t_i) \\ & + \beta_0 A_i + \beta_1 (A_i \times x_i) + \beta_2 (A_i \times t_i) + C_i^{'} \gamma + \epsilon_i. \end{aligned} \tag{20}$$

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 firmworker match-specific elements that may be correlated with the residual in the wage equation.[37] To the extent that these variables capture the channels through which tenure is correlated with the residual in Equation 20, we can use the residual from this regression as an instrument for tenure.

E.3.2 Results

Table E4 presents the results for men. The conclusions regarding public learning from Table E2 are robust to the inclusion of tenure. That is, the interaction between AFQT and experience in columns (3) and (6) remains positive, indicating there exist public learning. The interaction between marriage and experience is also positive, which indicates that there is no statistical discrimination based on marital status.

Table E4: Asymmetric SD-EL Model, Men

	OLS Potential Experience, Actual Tenure			**IV Actual Experience, Actual Tenure**		
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Married	0.212***	0.159***	0.160***	0.210***	0.161***	0.166***
	(0.025)	(0.036)	(0.036)	(0.026)	(0.039)	(0.038)
Married*Experience		0.015*	0.015*		0.043**	0.045**
		(0.008)	(0.008)		(0.019)	(0.019)
Married*Tenure		-0.001	-0.001		-0.007	-0.007
		(0.006)	(0.006)		(0.006)	(0.006)
AFQT	0.074***	0.074***	0.035*	0.075***	0.078***	0.037
	(0.013)	(0.013)	(0.021)	(0.014)	(0.014)	(0.023)
AFQT*Experience			0.004			0.007
			(0.004)			(0.008)
AFQT*Tenure			-0.001			0.000
			(0.003)			(0.003)
Adjusted R^2	0.334	0.335	0.336	0.309	0.303	0.301
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 2006 Dollars. 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, the number of children below the age of 5, the number of children aged 5-17, and a dummy for a child over the age of 18, a dummy for urban residence and cubic polynomials in time, tenure and experience. Columns 2, 3, 5 and 6 include three interaction terms between the married dummy and i. experience, ii. tenure and iii. time. Columns 3 and 6 further include three interaction terms between normalized AFQT and i. experience, ii. tenure and iii. time. 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 E.3.1, and experience terms are actual experience instrumented by potential experience are presented in Columns 4-6. Data used: NLSY79, 1979-2012.

Table E5 presents the estimates for women. The inclusion of tenure enriches the inference that we draw form Table E3 but do not modify the main conclusion. In the OLS results, the estimated coefficient associated to the interaction between tenure and AFQT points towards

[37] The regression output is summarized as follows:

$$\text{Men: } t_i = \underset{(0.080)}{-0.995} + \underset{(0.018)}{0.095}\ \bar{t}_i^{career} + \underset{(0.006)}{0.183}\ dur_i + \underset{(0.005)}{0.222}\ x_i + \hat{e}_i\ , \quad \text{Adjusted } R^2 : 0.446$$

$$\text{Women: } t_i = \underset{(0.078)}{-0.878} + \underset{(0.019)}{0.069}\ \bar{t}_i^{career} + \underset{(0.007)}{0.193}\ dur_i + \underset{(0.006)}{0.244}\ x_i + \hat{e}_i\ , \quad \text{Adjusted } R^2 : 0.484$$

the existence of private learning. 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 E. Married women start their careers experiencing a wage penalty with respect to their single counterparts. As their career progresses, this penalty becomes a premium.

Table E5: Asymmetric SD-EL Model, Women

	OLS Potential Experience, Actual Tenure			**IV Actual Experience, Actual Tenure**		
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Married	0.010	-0.090**	-0.091**	-0.012	-0.062	-0.063
	(0.029)	(0.036)	(0.036)	(0.031)	(0.041)	(0.041)
Married*Experience		0.016	0.015		0.048**	0.046*
		(0.012)	(0.012)		(0.024)	(0.024)
Married*Tenure		0.014*	0.014**		0.012	0.012
		(0.007)	(0.007)		(0.008)	(0.008)
AFQT	0.089***	0.089***	0.073***	0.081***	0.081***	0.079***
	(0.013)	(0.013)	(0.019)	(0.013)	(0.013)	(0.020)
AFQT*Experience			0.003			-0.002
			(0.004)			(0.012)
AFQT*Tenure			0.008**			0.004
			(0.003)			(0.004)
Adjusted R^2	0.277	0.282	0.284	0.256	0.251	0.253
Observations	6,899	6,899	6,899	6,898	6,898	6,898

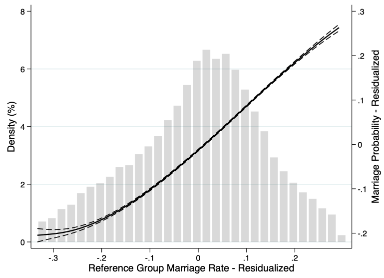
Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. Standard errors are reported in parentheses, where these are clustered by individual. The dependent variable in all columns is the natural log of wages in 2006 Dollars. The following additional control variables are included in all specifications: dummies for highest level of educational attainment, the education dummies interacted with a linear time trend, the number of children below the age of 5, the number of children aged 5-17, and a dummy for a child over the age of 18, a dummy for urban residence and cubic polynomials in time, tenure and experience. Columns 2, 3, 5 and 6 include three interaction terms between the married dummy and i. experience, ii. tenure and iii. time. Columns 3 and 6 further include three interaction terms between normalized AFQT and i. experience, ii. tenure and iii. time. 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 E.3.1, and experience terms are actual experience instrumented by potential experience are presented in Columns 4-6. Data used: NLSY79, 1979-2012.

F More on the Instrumental Variables Approach

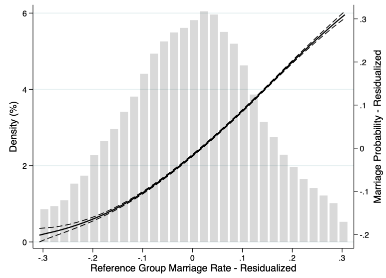
F.1 Support for the Identifying Assumptions

F.1.1 Supportive Evidence for the First Stage, Conditional Independence and Monotonicity Assumptions

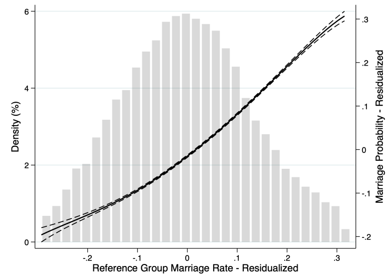
Figure F1: First Stage Relationship of Married on Proportion of Reference Group Married



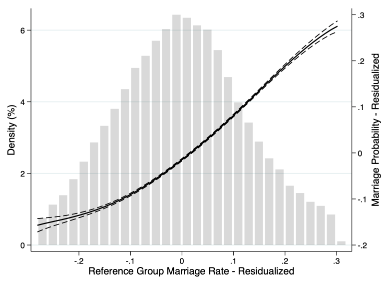
(a) Men, 1977-1983.



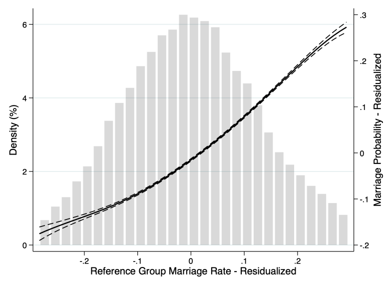
(b) Men, 1984-1990.



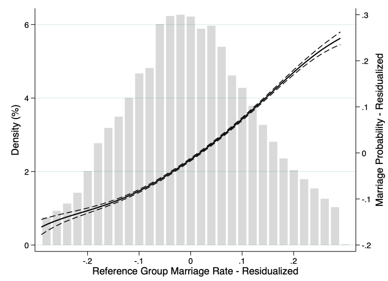
(c) Men, 1991-1997.



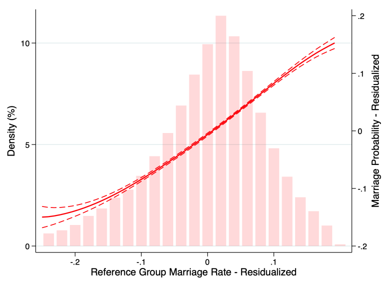
(d) Men, 1998-2004.



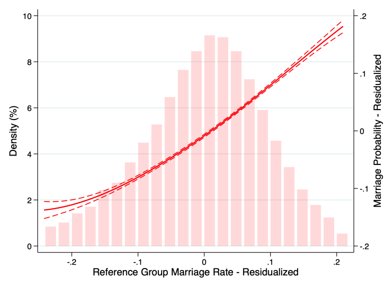
(e) Men, 2005-2011.



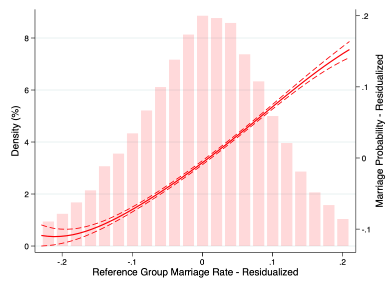
(f) Men, 2012-2018.



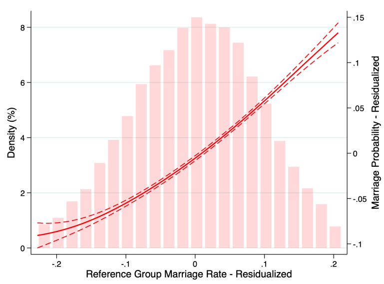
(g) Women, 1977-1983.



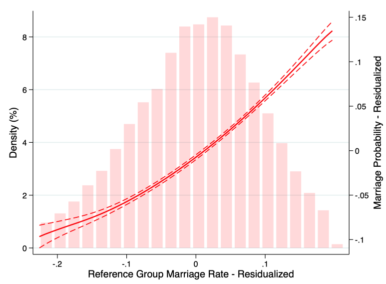
(h) Women, 1984-1990.



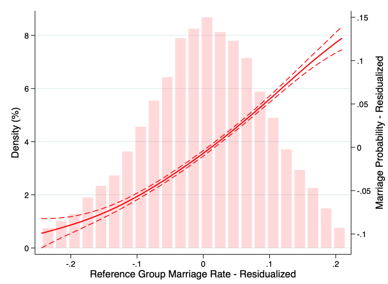
(i) Women, 1991-1997.



(j) Women, 1998-2004.



(k) Women, 2005-2011.



(l) Women, 2012-2018.

Notes: The solid lines are local linear regression of residualized marriage on the residualized instrument, and is a flexible version of the first stage 2SLS equation. Both marriage and the instrument are residualized on year and state fixed effects,dummies for highest level of educational attainment, dummies for years of potential experience, the number of children below the age of 5, the number of children aged 5-17, and a dummy for a child over the age of 18. The instrument in all specification in the proportion of individuals' reference group that are married. The specifications for women additionally include the Inverse Mills Ratio as a covariate. The dashed lines are 95% confidence intervals. The histogram of the (residualized) instrument is shown in the background, with the top and bottom 2% excluded from the figure. Data used: CPS, 1977-2018.

Table F1: Alternative 2SLS First Stage Specification

A. Men	**(1) 1998-2004**	**(2) 1998-2004**	**(3) 2005-2011**	**(4) 2005-2011**	**(5) 2012-2018**	**(6) 2012-2018**	**(7) 2012-2018**
RGMR	0.832***	0.830***	0.824***	0.822***	0.758***	0.758***	0.758***
	(0.022)	(0.022)	(0.022)	(0.022)	(0.023)	(0.023)	(0.023)
U.S. Born		-0.029***		-0.041***		-0.045***	-0.045***
		(0.006)		(0.007)		(0.007)	(0.007)
Health Status:							
Excellent		0.051***		0.048***		0.036***	0.035***
		(0.003)		(0.004)		(0.005)	(0.005)
Very Good		0.030***		0.033***		0.025***	0.025***
		(0.003)		(0.004)		(0.004)	(0.004)
Fair		-0.010		-0.014		-0.019**	-0.018**
		(0.007)		(0.010)		(0.008)	(0.008)
Poor		-0.038*		-0.013		-0.026	-0.020
		(0.020)		(0.020)		(0.024)	(0.025)
Difficulty:							
Hearing							-0.022
							(0.013)
Vision							-0.068***
							(0.020)
Physical							-0.023*
							(0.013)
p-value: Extra Covariates = 0	-	0.000	-	0.000	-	0.000	0.000
Observations	112,807	112,807	120,606	120,606	94,896	94,896	94,896
B. Women	**(7) 1998-2004**	**(8) 1998-2004**	**(9) 2005-2011**	**(10) 2005-2011**	**(11) 2012-2018**	**(12) 2012-2018**	**(13) 2012-2018**
RGMR	0.524***	0.526***	0.519***	0.528***	0.471***	0.471***	0.474***
	(0.026)	(0.026)	(0.029)	(0.029)	(0.031)	(0.031)	(0.031)
U.S. Born		-0.086***		-0.123***		-0.161***	-0.155***
		(0.008)		(0.011)		(0.012)	(0.013)
Health Status:							
Excellent		0.049***		0.032***		0.015**	0.020***
		(0.004)		(0.005)		(0.006)	(0.006)
Very Good		0.018***		-0.001		-0.010*	-0.004
		(0.005)		(0.005)		(0.006)	(0.006)
Fair		0.056***		0.103***		0.152***	0.119***
		(0.011)		(0.013)		(0.015)	(0.012)
Poor		0.215***		0.306***		0.382***	0.299***
		(0.025)		(0.029)		(0.049)	(0.044)
Difficulty:							
Hearing							-0.026
							(0.027)
Vision							0.056*
							(0.028)
Physical							0.166***
							(0.029)
p-value: Extra Covariates = 0	-	0.000	-	0.000	-	0.000	0.000
Observations	101,175	101,175	109,624	109,624	85,427	85,427	85,427

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, dummies for year of potential experience, the number of children below the age of 5, the number of children aged 5-17, and a dummy for a child over the age of 18. The instrument in all specification in the proportion of individuals' reference group that are married. In Panel B, the Inverse Mills Ratio is also included as a regressor. The additional covariates presented in the tables are not available for all periods of the data. US born is available from 1994 onwards, and health status from 1996 onwards. Information on physical difficulties are only available from 2009 onwards. Hence we present evidence only for periods where information on the additional covariates was available for the full period. The p-value is from a joint test of statistical significance of all additional covariates. Data used: CPS.

Table F2: 2SLS First Stage for Various Sub-Samples

		Predicted Married		**Region**				**Age**		**College**	
A. Men	(1) Full Sample	(2) Below Median	(3) Above Median	(4) North-East	(5) Mid-West	(6) South	(7) West	(8) 25-39	(9) 40-54	(10) No	(11) Yes
1977-1983	0.839***	0.509***	0.163***	0.855***	0.813***	0.760***	0.939***	0.819***	1.115***	0.861***	0.815***
	(0.022)	(0.042)	(0.029)	(0.019)	(0.016)	(0.027)	(0.037)	(0.022)	(0.100)	(0.025)	(0.034)
1984-1990	0.906***	0.465***	0.230***	0.911***	0.909***	0.866***	0.964***	0.893***	0.840***	0.935***	0.875***
	(0.020)	(0.032)	(0.040)	(0.023)	(0.013)	(0.043)	(0.062)	(0.020)	(0.051)	(0.019)	(0.031)
1991-1997	0.867***	0.308***	0.247***	0.967***	0.872***	0.796***	0.839***	0.863***	0.875***	0.878***	0.856***
	(0.024)	(0.039)	(0.036)	(0.020)	(0.047)	(0.042)	(0.055)	(0.020)	(0.073)	(0.026)	(0.027)
1998-2004	0.832***	0.252***	0.277***	0.889***	0.812***	0.777***	0.885***	0.817***	0.944***	0.819***	0.832***
	(0.022)	(0.029)	(0.026)	(0.032)	(0.026)	(0.028)	(0.067)	(0.022)	(0.052)	(0.026)	(0.030)
2005-2011	0.824***	0.208***	0.382***	0.856***	0.812***	0.794***	0.844***	0.788***	0.979***	0.766***	0.851***
	(0.022)	(0.045)	(0.033)	(0.063)	(0.017)	(0.024)	(0.060)	(0.020)	(0.050)	(0.022)	(0.027)
2012-2018	0.758***	0.216***	0.529***	0.766***	0.729***	0.765***	0.791***	0.697***	0.903***	0.648***	0.802***
	(0.023)	(0.038)	(0.044)	(0.041)	(0.020)	(0.031)	(0.088)	(0.024)	(0.049)	(0.022)	(0.026)
B. Women											
1977-1983	0.746***	0.247***	0.038**	0.790***	0.713***	0.613***	0.875***	0.728***	0.759***	0.790***	0.653***
	(0.040)	(0.039)	(0.016)	(0.055)	(0.043)	(0.058)	(0.140)	(0.041)	(0.125)	(0.045)	(0.041)
1984-1990	0.739***	0.194***	0.059***	0.842***	0.737***	0.638***	0.724***	0.740***	0.574***	0.772***	0.692***
	(0.034)	(0.039)	(0.016)	(0.069)	(0.040)	(0.054)	(0.118)	(0.034)	(0.097)	(0.038)	(0.039)
1991-1997	0.661***	0.174***	0.066***	0.797***	0.631***	0.590***	0.600***	0.717***	0.453***	0.681***	0.636***
	(0.041)	(0.034)	(0.016)	(0.065)	(0.053)	(0.070)	(0.144)	(0.037)	(0.096)	(0.045)	(0.048)
1998-2004	0.524***	0.128***	0.093***	0.583***	0.462***	0.466***	0.558***	0.574***	0.418***	0.388***	0.541***
	(0.026)	(0.032)	(0.017)	(0.050)	(0.047)	(0.045)	(0.048)	(0.027)	(0.054)	(0.037)	(0.025)
2005-2011	0.519***	0.239***	0.088***	0.585***	0.514***	0.487***	0.456***	0.561***	0.410***	0.297***	0.553***
	(0.029)	(0.030)	(0.019)	(0.049)	(0.060)	(0.040)	(0.092)	(0.031)	(0.053)	(0.028)	(0.037)
2012-2018	0.471***	0.171***	0.193***	0.553***	0.401***	0.480***	0.444***	0.440***	0.422***	0.230***	0.479***
	(0.031)	(0.037)	(0.021)	(0.037)	(0.021)	(0.048)	(0.114)	(0.033)	(0.056)	(0.038)	(0.034)

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, dummies for year of potential experience, the number of children below the age of 5, the number of children aged 5-17, and a dummy for a child over the age of 18. In Panel B, the Inverse Mills Ratio is also included as a regressor. The instrument in all specification in the proportion of individuals' reference group that are married. Columns 2 and 3 present results based on whether individuals where above or below the median based on predicted marriage. This involves running a linear probability model of the married dummy on all key covariates, but not the instrument. Data used: CPS, 1977-2018.

F.1.2 The Exclusion Restriction

In this section we consider two different approaches that allow us to consider the impact of a violation of the exclusion restriction on our 2SLS estimates. These are the plausbibly exogenous approach of Conley et al. (2012) and the imperfect instrumental variable approach of Nevo and Rosen (2012). Both of these methods conisder violations of the exclusion restriction, albeit from slightly different persepectives. Simply put, the evidence we provide below shows that even with quite substantial violations of the exclusion restriction, our key results of both i.) a near constant causal effect of marriage on wages for men and ii.) the emergence of a causal MWP for women hold.

Plausibly Exogenous We follow the *plausibly exogenous* approach of Conley et al. (2012) to consider the impact of a (slight) violation of the exclusion restriction.

Abstracting from other control variables, Conley et al. (2012) present the 2SLS estimating equations, allowing for an instrument Z that is non-excludable, as:

$$Y = M\alpha + Z\gamma + \epsilon, \tag{21}$$

$$M = Z\Pi + V, \tag{22}$$

where $E[M\epsilon] \neq 0$, $E[Z\epsilon] = 0$ and the exclusion restriction can be thought of $\gamma \equiv 0$.

Conley et al. (2012) consider four different strategies to implement their technique, which handle deviations from $\gamma \equiv 0$. We consider the Local to Zero Approximation (LTZ) strategy and use the insights from van Kippersluis and Rietveld (2018) in order to specify a value for γ. The LTZ strategy assumes a prior on γ that follows a Normal distribution, with mean μ_γ and variance Ω_γ, where the uncertainty regarding γ reduces with sample size. In this setting we can write the plausibly exogenous estimator as:

$$\hat{\beta} \sim N(\beta_{2SLS} + A\mu_\gamma, W_{2SLS} + A\Omega_\gamma A'), \tag{23}$$

where $A = (X'Z(Z'Z)^{-1}Z'X)^{-1}(X'Z)$, β_{2SLS} is the 2SLS point estimate, and W_{2SLS} is the 2SLS variance-covariance matrix.

As van Kippersluis and Rietveld (2018) note, if there exists a sub-group for whom the first stage is zero (i.e. a group for whom the marriage decision is not impacted by our instrument of local social norms), then the reduced form for this group (i.e. the coefficient on local social norms in a wage regression) is informative of whether or not the exclusion restrictions holds. We refer to this sub-sample as the zero-first-stage sub-sample. To see how this approach is informative of the exclusion restriction, consider a reduced form equation for wages where we substitute Equation 22 into Equation 21:

$$Y = Z(\gamma + \alpha\Pi) + (\epsilon + \alpha V). \tag{24}$$

For the zero-first-stage sub-sample, $\Pi = 0$, the reduced form estimate of our instrument is given by γ. If we make the assumption of homogeneous direct effects of the instrument (i.e that the $\hat{\gamma}$ estimated from the zero-first-stage sub-sample is informative of γ for the full sample), then we can assess the impact of a violation of the exclusion restriction on our 2SLS estimates for the

Table F3: Zero-First-Stage Sample
First Stage and Reduced Form

	Men - IV			**Women - IV-Heckman**		
	1977-1990	**1991-2004**	**2005-2018**	**1977-1990**	**1991-2004**	**2005-2018**
First Stage	0.062**	0.021	0.042	0.000	-0.006	0.041
	(0.029)	(0.031)	(0.026)	(0.030)	(0.034)	(0.038)
Reduced Form	0.096***	0.026	0.010	0.034	-0.013	-0.026
	(0.027)	(0.033)	(0.037)	(0.029)	(0.023)	(0.025)
Observations	48,721	44,901	39,015	35,596	41,096	36,245

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 the First Stage section is a dummy for married, and in the Reduced Form panel 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, dummies for year of potential experience, the number of children below the age of 5, the number of children aged 5-17, and a dummy for a child over the age of 18. The instrument in all specification is the proportion of individuals' reference group that are married. In Columns 4-6, the Inverse Mills Ratio is also included as a regressor. The exclusion restrictions for the employment equation are a series of dummies for age of youngest child in the household from 1-18, where age less that 1 is the base category, a dummy for ages 19-24, a dummy for 25 and over, and no children are also included. The zero-first-stage sample is constructed by predicting marriage from a regression specification with all model controls listed, but not the instrument. Quartiles of the predicted married variable are created, and the first quartile form the zero-first-stage sample. Data used: CPS, 1977-2018.

full sample. Specifically this means we set $\mu_\gamma = \hat{\gamma}$. This approach is appealing as it provides an empirically grounded value for the direct effect of the instrument in our wage equations, rather than relying on arbitrary values for μ_γ.[38]

Table F3 displays first stage results for a variety of sub-samples of our data. We note that the first stage is weaker for those with predicted values of married below the median.[39] Hence, we consider a smaller sub-sample, those with predicted marriage in the lowest quartile, as a candidate sub-sample for the zero-first-stage sub-sample.

Table F3 highlights that, with the exception of the earliest period for men, we can consider this group as our zero-first-stage sub-sample because the instrument does not bite for this sub-sample.[40] Thus, we use the reduced form estimate on our instrument based on this sample as $\hat{\gamma}$, our estimate for μ_γ. The reduced form estimates are also presented in Table F3. In all cases but one the estimate of the direct effect of the instrument on wages is insignificantly different from zero. The single exception is for men in the 1977-1990 period, i.e. the only period for which we could not establish a meaningful zero-first-stage group.

Tables F4 and F5 present the results of the LTZ implementation of Conley et al. (2012)'s plausibly exogenous approach. In the first row, we report our IV estimates. The second and third rows present the estimates for those specifications in which we allow the exclusion restriction to be mildly violated according to our estimates for $\hat{\gamma}$ (reported in Table F3), both with and without uncertainty. The fourth row reports the value of γ required to cause the married coefficient to be

[38]The final ingredient in this approach is how to specify uncertainty regarding the direct effect of the instrument, Ω_γ. We take two approaches here, again following van Kippersluis and Rietveld (2018). The first is to set $\Omega_\gamma = 0$. The second is to follow a suggestion from Imbens and Rubin (2015) regarding normalized differences between covariates in a treatment and control group in a regression framework not exceeding 0.25. Applied to our setting, we specify $\Omega_\gamma = (.125\sqrt{S_0^2 + S_{-0}^2})^2$, where S_0 is the standard error on $\hat{\gamma}$ for the zero-first-stage sub-sample and S_{-0} is the equivalent for the remainder of the sample.

[39]We do not use the instrument when predicting married here.

[40]For completeness, we still present the results of this exercise for men in the 1977-1990 period below. To be clear, however, given the fact that we did not manage to construct a zero-first-stage sub-sample for men in this period, one should disregard the results for this group.

zero. We name this value γ^*.[41] The fifth row presents the ratio between γ^* and $\hat{\gamma}$ as a measure of how much larger the empirically-grounded $\hat{\gamma}$ would needs to be such that the violation of the exclusion restriction would render there to be no effect of marriage on wages. For men, $\hat{\gamma}$ would need to be an order of magnitude larger in order to drive the effect of marriage on wages to zero. For women in the later periods, the value of $\hat{\gamma}$ would need to change sign, and be substantially larger. Overall, we interpret the results of Tables F4 and F5 as providing strong supportive evidence for the robustness of our IV estimates to the violation of the exclusion restriction.

Table F4: Plausibly Exogenous - Men

	1977-1983	**1984-1990**	**1991-1997**	**1998-2004**	**2005-2011**	**2012-2018**
2SLS	0.306***	0.269***	0.246***	0.226***	0.266***	0.237***
	(0.013)	(0.015)	(0.013)	(0.016)	(0.017)	(0.021)
Plausibly Exogenous	0.187***	0.158***	0.214***	0.194***	0.254***	0.225***
	(0.013)	(0.015)	(0.013)	(0.016)	(0.017)	(0.021)
Plausibly Exogenous (With Uncertainty)	0.187***	0.158***	0.214***	0.194***	0.254***	0.225***
	(0.014)	(0.016)	(0.014)	(0.017)	(0.018)	(0.022)
γ^*	0.248	0.234	0.207	0.188	0.221	0.187
$\gamma^*/\hat{\gamma}_{ZFS}$	2.575	2.430	7.896	7.172	22.033	18.643
Observations	104,970	104,545	96,210	112,807	120,606	94,896

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. Year and state fixed effects are included in all regressions. The following additional controls are included: dummies for highest level of educational attainment, dummies for year of potential experience, the number of children below the age of 5, the number of children aged 5-17, and a dummy for a child over the age of 18. The instrument in all specification is the proportion of individuals' reference group that are married. We specify $\hat{\gamma}$ as the reduced form estimate from the zero-first-stage sample. For the plausibly exogenous with uncertainty results, we set $\Omega_\gamma = (.125\sqrt{S_0^2 + S_{-0}^2})^2$, where S_0 is the standard error on $\hat{\gamma}$ for the zero-first-stage sub-sample and S_{-0} is the equivalent for the remainder of the sample. γ^* is the value of γ that yields a coefficient on marriage equal to zero. Data used: CPS, 1977-2018.

[41] We compute γ^* by estimating the plausibly exogenous specification over a grid of values for γ. We find the value of γ that generates a zero married coefficient using bisection.

Table F5: Plausibly Exogenous - Women

	1977-1983	**1984-1990**	**1991-1997**	**1998-2004**	**2005-2011**	**2012-2018**
2SLS	-0.149***	-0.058	-0.018	-0.033	0.104***	0.086***
	(0.022)	(0.037)	(0.028)	(0.040)	(0.031)	(0.032)
Plausibly Exogenous	-0.197***	-0.106**	0.001	-0.010	0.153***	0.139***
	(0.025)	(0.044)	(0.030)	(0.045)	(0.031)	(0.034)
Plausibly Exogenous	-0.197***	-0.106**	0.001	-0.010	0.153***	0.139***
(With Uncertainty)	(0.026)	(0.044)	(0.031)	(0.045)	(0.032)	(0.036)
γ^*	-0.107	-0.042	-0.012	-0.018	0.055	0.043
$\gamma^*/\hat{\gamma}_{ZFS}$	1.038	1.438	-2.597	-2.093	-2.710	-1.637
Observations	70,872	83,482	83,791	101,175	109,624	85,427

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. Year and state fixed effects are included in all regressions. The following additional controls are included: the Inverse Mills Ratio, dummies for highest level of educational attainment, dummies for year of potential experience, the number of children below the age of 5, the number of children aged 5-17, and a dummy for a child over the age of 18. The instrument in all specification is the proportion of individuals' reference group that are married. We specify $\hat{\gamma}$ as the reduced form estimate from the zero-first-stage sample. For the plausibly exogenous with uncertainty results, we set $\Omega_\gamma = (.125\sqrt{S_0^2 + S_{-0}^2})^2$, where S_0 is the standard error on $\hat{\gamma}$ for the zero-first-stage sub-sample and S_{-0} is the equivalent for the remainder of the sample. γ^* is the value of γ that yields a coefficient on marriage equal to zero. Data used: CPS, 1977-2018.

Imperfect Instrumental Variable To bring the key ideas into focus, consider a simplified version of the 2SLS equations (in which we abstract from other covariates):

$$Y = M\alpha + \epsilon, \tag{25}$$

$$M = Z\Pi + V, \tag{26}$$

where $E[M\epsilon] \neq 0$, $E[Z\epsilon] = 0$.

The crux of the *imperfect instrumental variable* (IIV) approach of Nevo and Rosen (2012) is that it allows for the possibility that Z is correlated with the error ϵ. However, this relaxation of the moment condition for Z implies losing point identification.

There are two key assumptions inherent in the IIV approach. To discuss these, we follow the notation of Nevo and Rosen (2012), using σ_{ab} and ρ_{ab} to respectively denote the covariance and correlation between any two random variables A and B, and σ_a to denote the standard deviation of A. First, we assume that

$$\rho_{m\epsilon}\rho_{z\epsilon} \geq 0. \tag{27}$$

That is, our instrument has the same direction of correlation with the error term in the wage equation as does being married. Note, the standard IV assumption is that $\rho_{z\epsilon} = 0$.

Next, we assume that

$$\mid \rho_{m\epsilon} \mid \geq \mid \rho_{z\epsilon} \mid . \tag{28}$$

This states that our imperfect instrument is, at most, as correlated with the wage error term as is our endogenous treatment variable of marriage.

Equation 28 provides the basis to define $\lambda^* = \rho_{z\epsilon}/\rho_{m\epsilon}$, the relative degree of correlation between the IV and the error term, and marriage and the error term. We do not know the true value of λ^*, but if we were to know this, we could construct a (valid) compound instrument - $V(\lambda^*) = \sigma_x Z - \lambda^* \sigma_z X$, and use this to form the required moment condition for the identification of α:

$$E[V(\lambda^*)\epsilon] = E[(\sigma_x Z - \lambda^* \sigma_z X)\epsilon] = \sigma_x \sigma_{z\epsilon} - \lambda^* \sigma_z \sigma_{x\epsilon} = 0. \tag{29}$$

We can now consider the bounding approach that Nevo and Rosen (2012) suggest, replace the valid instrument $V(\lambda^*)$ with $V(1)$, the limit case from Equation 28. To be clear, the use of $V(1)$ establishes bounds based on the case where the imperfect IV is as endogenous as marital status.

Using the IIV approach, one can retrieve two-sided bounds for the case where the instrument is negatively correlated with the endogenous regressor ($\hat{\Pi} < 0$ in Equation 26), but only one-sided bounds when the instrument is positively correlated with the endogenous regressor. The latter describes our case.

We suspect that marriage is positively correlated with unobservables in the wage equation, which implies $\rho_{m\epsilon} > 0$. This means we will calculate an upper bound using the IIV approach. Denoting α_z^{IV} as the 2SLS estimator using the imperfect IV and $\alpha_{v(1)}^{IV}$ as the 2SLS estimator using the limit case for the valid IV $V(1)$, and assuming $\rho_{m\epsilon} > 0$, we can provide an upper bound on α:

$$\alpha \leq min\{\alpha_z^{IV}, \alpha_{v(1)}^{IV}\}. \tag{30}$$

We present the results of this bounding exercise in Table F6. Even with the pessimistic assumption that the IV is as endogenous as marriage, our IV estimates fall always below the upper bounds. We interpret the results as evidence that the patterns that we document in the main body of the paper are robust to a moderate violation of the exclusion restriction.

Table F6: Imperfect Instrument Approach

A: Men	**(1) '77-'83**	**(2) '84-'90**	**(3) '91-'97**	**(4) '98-'04**	**(5) '05-'11**	**(6) '12-'18**
Upper Bound	0.170	0.183	0.204	0.204	0.210	0.190
95% Upper CI	[0.187]	[0.199]	[0.221]	[0.218]	[0.227]	[0.205]
B: Women						
Upper Bound	-0.149	-0.058	-0.018	-0.033	0.067	0.078
95% Upper CI	[-0.094]	[0.031]	[0.050]	[0.066]	[0.087]	[0.094]

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 table presents the upper bound for the coefficient on the married dummy, thus giving an estimate for α in Equation 30, and the corresponding 95% confidence interval for the upper bound. Year and state fixed effects are included in all regressions. The following additional controls are included: dummies for highest level of educational attainment, dummies for year of potential experience, the number of children below the age of 5, the number of children aged 5-17, and a dummy for a child over the age of 18. The instrument in all specification is the proportion of individuals' reference group that are married. In Panel B, the Inverse Mills Ratio is also included as a regressor. The exclusion restrictions for the employment equation are a series of dummies for age of youngest child in the household from 1-18, where age less that 1 is the base category, a dummy for ages 19-24, a dummy for 25 and over, and no children are also included. Data used: CPS, 1977-2018.

F.2 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. Our approach follows that of Dahl et al. (2014).

F.2.1 Complier Types

In this section, we 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:

$$\pi_c \equiv Pr(M_i = 1|Z_{M,i} = \overline{Z}) - Pr(M_i = 1|Z_{M,i} = \underline{Z}) = Pr(M_i(\overline{Z}) > M_i(\underline{Z})), \tag{31}$$

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:

$$\pi_a \equiv Pr(M_i = 1|Z_{M,i} = \underline{Z}) = Pr(M_i(\overline{Z}) = M_i(\underline{Z}) = 1), \tag{32}$$

and the proportion of never takers:

$$\pi_n \equiv Pr(M_i = 1|Z_{M,i} = \overline{Z}) = Pr(M_i(\overline{Z}) = M_i(\underline{Z}) = 0). \tag{33}$$

Table F7 presents these proportions of compliers for each sample, 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 2 for the non-selection corrected 2SLS approach and Equation 5 for the selection-corrected counterpart. After residualizing both marriage and our instrument with respect to our control variables, we run a local linear regression of resizualized marriage on our residualized instrument.[42]

We can also calculate these proportions using a linear model (i.e. Equations 2 and 5). 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. 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.

[42] Figure F1 presents the local linear regression relation between residualized marriage and local social norms in marriage underlying these calculations.

Table F7: Sample Share of Compliers

Model:	**Local Linear**			**Linear**		
Top/Bottom:	(1) 1%	(2) 2%	(3) 5%	(4) 1%	(5) 2%	(6) 5%
Men						
1977 - 1983	0.54	0.49	0.42	0.60	0.52	0.42
1984 - 1990	0.60	0.53	0.44	0.63	0.56	0.44
1991 - 1997	0.58	0.51	0.42	0.60	0.53	0.42
1998 - 2004	0.51	0.45	0.38	0.55	0.48	0.39
2005 - 2011	0.51	0.45	0.37	0.54	0.47	0.38
2012 - 2018	0.46	0.41	0.34	0.50	0.44	0.35
Women						
1977 - 1983	0.35	0.31	0.25	0.39	0.33	0.24
1984 - 1990	0.36	0.32	0.26	0.39	0.33	0.26
1991 - 1997	0.29	0.27	0.23	0.34	0.29	0.23
1998 - 2004	0.24	0.21	0.18	0.27	0.23	0.18
2005 - 2011	0.22	0.20	0.17	0.25	0.22	0.17
2012 - 2018	0.18	0.18	0.17	0.25	0.21	0.17

Notes: The dependent variable in Columns 1-3 is the residualized dummy for married, and in Columns 4-6 is a dummy for marriage. Year and state fixed effects are included in all regressions. The following additional controls are included: dummies for highest level of educational attainment, dummies for year of potential experience, the number of children below the age of 5, the number of children aged 5-17, and a dummy for a child over the age of 18. The instrument in Columns 1-3 the residualized proportion of individuals' reference group that are married. In Columns 4-6 the instrument is the proportion of individuals' reference group that are married. For women, the Inverse Mills Ratio is also included as a regressor. The exclusion restrictions for the employment equation are a series of dummies for age of youngest child in the household from 1-18, where age less that 1 is the base category, a dummy for ages 19-24, a dummy for 25 and over, and no children are also included. The local linear approach uses an Epanechnikov kernel with bandwidth .15 and polynomial of degree 2. Data used: CPS, 1977-2018.

F.2.2 Characterizing Compliers

Although we cannot identify compliers individually, we are able to characterize the compliant sub-population. This is useful, given that we can only identify the 2SLS estimate of the MWP based on the compliers, and not the full sample. The statistic of interest to characterize the compliers is $\frac{P[X=x|complier]}{P[X=x]}$. This tells us how much more likely a complier is to have a given characteristic compared to the sample as a whole. In order to calculate the numerator, we calculate several ancillary statistics:

$$P[X = x|complier] = \frac{P[complier|X = x] \times P[X = x]}{P[complier]}, \tag{34}$$

where $P[complier] = \hat{\pi}_1(\overline{Z} - \underline{Z})$ is calculated as described in the Section F.2.1 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 $X = x$. We create 6 mutually exclusive sub-groups based on education (college and no college) and age (three 10-year bands) group categories. Table F8 presents the results.

Table F8: Complier Characteristics

Education:	**No College**			**College**		
Age Group:	(1) 25-34	(2) 35-44	(3) 45-54	(4) 25-34	(5) 35-44	(6) 45-54
Men						
1977 - 1983	1.070	0.813	0.346	0.980	0.889	0.366
1984 - 1990	1.119	0.833	0.280	0.926	0.847	0.325
1991 - 1997	1.147	0.945	0.284	0.940	0.962	0.369
1998 - 2004	1.078	1.042	0.286	0.938	1.119	0.459
2005 - 2011	0.965	1.083	0.441	0.939	1.116	0.681
2012 - 2018	0.851	1.086	0.569	0.875	1.155	0.746
Women						
1977 - 1983	0.943	0.688	0.279	1.041	1.024	0.544
1984 - 1990	0.959	0.620	0.167	1.079	0.973	0.340
1991 - 1997	1.153	0.656	0.192	1.163	0.976	0.322
1998 - 2004	0.837	0.528	0.117	1.151	1.318	0.447
2005 - 2011	0.661	0.412	0.245	1.165	1.255	0.402
2012 - 2018	0.458	0.530	0.144	0.900	1.176	0.485

Notes: In order to generate the conditional probabilities above, we run a series of first stage equations for both the full sample, and for the six mutually exclusive sub-groups. The dependent variable in all specifications is a dummy for married. Year and state fixed effects are included in all regressions. The following additional controls are included: dummies for highest level of educational attainment, dummies for year of potential experience, the number of children below the age of 5, the number of children aged 5-17, and a dummy for a child over the age of 18. The instrument is the proportion of individuals' reference group that are married. For women, the Inverse Mills Ratio is also included as a regressor. The exclusion restrictions for the employment equation are a series of dummies for age of youngest child in the household from 1-18, where age less that 1 is the base category, a dummy for ages 19-24, a dummy for 25 and over, and no children are also included. The values for $\underline{Z}$ and $\overline{Z}$ are defined for each respective sub-sample and are based on the 5% and 95% percentile of the instrument respecitively. Data used: CPS, 1977-2018.