AMERICAN SOCIOLOGICAL REVIEW

OFFICIAL JOURNAL OF THE AMERICAN SOCIOLOGICAL ASSOCIATION

ONLINE SUPPLEMENT
to article in
AMERICAN SOCIOLOGICAL REVIEW, 2018, VOL. 83

Is There a Male Marital Wage Premium? New Evidence from the United States

Volker Ludwig
TU Kaiserslautern

Josef Brüderl
University of Munich

Part A. Bias of FE and FEGS Models

We show here, analytically and by Monte Carlo simulation, why FE and FEGS are inconsistent estimators for the MWP under general patterns of marital selection that are allowed in the framework of FEIS estimation. The main points are (1) the FE estimator is biased if the decision of whether to marry is related to the steepness of the wage trajectory, (2) the FEGS estimator allows for an association of the steepness of the wage trajectory to the time-constant propensity of marriage, but rules out any relation to marriage timing, and (3) the FEIS estimator is consistent in both cases.

Statistical Framework

Write the model given in Equation 2 in the main text in more general terms as

$$y_{it} = \boldsymbol{x}_{1it}\boldsymbol{\beta} + \alpha_{1i} + \boldsymbol{x}_{2it}\boldsymbol{\alpha}_{2i} + \varepsilon_{it}, \qquad \text{(S1)}$$

where y_{it} is the wage, $\boldsymbol{x}_{1it}$ is a $(1\ X\ K)$ vector of covariates, including an indicator for marriage m_{it}, and $\boldsymbol{x}_{2it}$ is a $(1\ X\ J)$ vector of variables that interact with unobservables (work experience exp_{it} and its square, in our case). Everything else is defined as in the main text.

(1) Under the model given in Equation S1, the standard FE estimator is not consistent. Let $\boldsymbol{\alpha}_{2i} \equiv \boldsymbol{\alpha}_2 + \boldsymbol{a}_i$, where individual-specific trends $\boldsymbol{a}_i$ are deviations from the common trajectory for the whole population. If we run a standard FE model, we get

$$\ddot{y}_{it} = \ddot{\boldsymbol{x}}_{1it}\boldsymbol{\beta} + \ddot{\boldsymbol{x}}_{2it}\boldsymbol{\alpha}_2 + \ddot{\boldsymbol{x}}_{2it}\boldsymbol{a}_i + \ddot{\xi}_{it},$$

where the dots indicate that all variables have been de-meaned. Due to de-meaning, α_{1i} dropped out of the equation. However, individual deviations from the common wage trajectory are now contained in the error term. Hence, even if the usual strict exogeneity assumption that $E(\ddot{\boldsymbol{x}}'_{it}\xi_{it}) = \boldsymbol{0}$ holds, consistency of FE hinges on the additional assumption that $E(\ddot{\boldsymbol{x}}'_{1it}\ddot{\boldsymbol{x}}_{2it}\boldsymbol{a}_i) = \boldsymbol{0}$.

The assumption is certainly violated if marriage is more likely for men with higher-than-average wage growth. Since $\ddot{m}_{it} = m_{it} - \overline{m}_i$ is part of $\ddot{\boldsymbol{x}}_{1it}$, the assumption rules out any relation of individual wage trajectories to the person-specific time-average of the marriage dummy, $\overline{m}_i$. Hence, FE estimates a biased effect of marriage if the decision whether to marry is related to the steepness of the wage career. A sufficient assumption for consistency of FE is that individual deviations from the common slope of $\ddot{\boldsymbol{x}}_{2it}$ (work experience) are not related to any of the de-meaned covariates (including the marriage dummy). Formally, a sufficient condition is $E(\boldsymbol{a}_i|\ddot{\boldsymbol{x}}_{it}) = \boldsymbol{0}$, which is a version of the well-known assumption of parallel trends in outcomes.

(2) Similarly, the standard strict exogeneity assumption does not guarantee consistency of the FEGS estimator in the situation described by Equation S1. To see why, define $\boldsymbol{\alpha}_{2i} \equiv E(\boldsymbol{\alpha}_{2i} \mid treat_i) + \boldsymbol{d}_i$, where the expectation of $\boldsymbol{\alpha}_{2i}$ is either $\boldsymbol{\alpha}_{20i}$ or $\boldsymbol{\alpha}_{21i}$, depending on the treatment group (ever-married or never-married), and $\boldsymbol{d}_i$ are individual deviations from the treatment group–specific trajectories. If we estimate an FEGS model, we get

$$\ddot{y}_{it} = \ddot{\boldsymbol{x}}_{1it}\boldsymbol{\beta} + \ddot{\boldsymbol{x}}_{2it}\boldsymbol{\alpha}_{20i} + treat_i \cdot \ddot{\boldsymbol{x}}_{2it}\boldsymbol{\alpha}_{21i} + \ddot{\boldsymbol{x}}_{2it}\boldsymbol{d}_i + \ddot{\psi}_{it},$$

or, using more compact notation,

$$\ddot{y}_{it} = \ddot{\boldsymbol{x}}_{it}\boldsymbol{\gamma} + \ddot{v}_{it},$$

where, for simplicity, $\ddot{\boldsymbol{x}}_{it}$ includes $\ddot{\boldsymbol{x}}_{1it}$, $\ddot{\boldsymbol{x}}_{2it}$, and $treat_i \cdot \ddot{\boldsymbol{x}}_{2it}$. Again, α_{1i} is eliminated by de-meaning the data. While $\boldsymbol{\alpha}_{20i}$ and $\boldsymbol{\alpha}_{21i}$ are controlled explicitly, such that the assumption of parallel trends in the two treatment groups is no longer needed, individual deviations from the group-specific slopes are contained in the error term. Thus, $\ddot{v}_{it} = \ddot{\boldsymbol{x}}_{2it}\boldsymbol{d}_i + \ddot{\psi}_{it}$ is a composite error term. Following the reasoning for the FE model, strict exogeneity of ψ_{it} is not sufficient for consistency of FEGS. In addition, it must hold that $E(\ddot{\boldsymbol{x}}'_{1it}\ddot{\boldsymbol{x}}_{2it}\boldsymbol{d}_i) = \boldsymbol{0}$. A sufficient assumption is that individual deviations from the group-specific slopes of $\ddot{\boldsymbol{x}}_{2it}$ (work experience) are not related to the de-meaned covariates (notably, the marriage dummy). Formally, $E(\boldsymbol{d}_i|\ddot{\boldsymbol{x}}_{it}) = \boldsymbol{0}$.

Obviously, this condition of common trends within treatment groups is a weaker condition than the assumption of parallel trends needed for FE, because FEGS does allow for mean differences in wage growth between the never-married and ever-married men. What FEGS rules out is further individual differences in wage growth between ever-married men that are related to marriage timing.

In practice, FEGS reduces the bias of FE. Intuitively, the bias is smaller with FEGS because including the interaction of a time-constant treatment indicator $treat_i$ and $\ddot{\boldsymbol{x}}_{2it}$ in the regression model controls for the relation of $\boldsymbol{d}_i$ to the person-specific mean of marriage ($\bar{m}_i$). However, $\boldsymbol{d}_i$ may still be related to deviations from the person-means of the variables in $\boldsymbol{x}_{1it}$ (e.g., to $m_{it} - \bar{m}_i$). Hence, the estimate of the treatment effect would still be biased, because $E(\ddot{\boldsymbol{x}}'_{1it}\ddot{\boldsymbol{x}}_{2it}\boldsymbol{d}_i) \neq \boldsymbol{0}$.

How can this happen? Consider a simple example with variation of marriage timing. Suppose we observe persons for the first three years of their careers, where all of them are never-married at $exp_{it} = 0$ and some of the treated persons marry at $exp_{it} = 1$, and others marry at $exp_{it} = 2$. Then $\bar{m}_i$ takes value 0 for each of the never-married persons, but values .67 and .33 for persons marrying "early" and "late," respectively. In this case, if the timing of marriage varies according to values of $\boldsymbol{\alpha}_{2i}$, $\boldsymbol{d}_i$ remains related to $(m_{it} - \bar{m}_i)$. Hence, it would not be sufficient to control for group-specific trends in just one control and one treatment group, as done by interacting $treat_i$ and exp_{it}. (An interaction with exp_{it} would be needed for each level of $\bar{m}_i$. In fact, in a simple setting with fully balanced panels and without further covariates, a possible remedy of the FEGS estimator is an extension of the specification, where the interaction of $treat_i$ and $\ddot{\boldsymbol{x}}_{2it}$ is replaced by a set of dummy interactions of each level of $\bar{m}_i$ and $\ddot{\boldsymbol{x}}_{2it}$. This extended specification would recover the true treatment effect, since it is equivalent to FEIS.)

From this discussion, FEGS would fail if selection into treatment (marriage) varies over time, and the pattern of selection on the individual level, that is, marriage timing, depends on the values of (some of the) individual-specific unobserved variables that also determine the growth of the outcome (the wage). In other words, unlike FE, FEGS allows that unobservables in $\boldsymbol{\alpha}_{2i}$ determine whether a man eventually marries. What FEGS rules out is that $\boldsymbol{\alpha}_{2i}$ determines when he marries.

(3) Neither of the additional conditions for $\boldsymbol{a}_i$ or $\boldsymbol{d}_i$ are needed for consistency of FEIS, however, because a more general within transformation is applied to the data on the individual level. As described in the methods section of the article, to estimate the effects of marriage and other covariates in $\boldsymbol{x}_{1it}$, for each variable only the variation that is not due to $\boldsymbol{x}_{2it}$ (experience) is used. Therefore, both α_{1i} and $\boldsymbol{\alpha}_{2i}$ drop out of Equation S1, along with $\boldsymbol{x}_{2it}$. The FEIS estimation equation is given by

$$\tilde{y}_{it} = \tilde{\boldsymbol{x}}_{1it}\boldsymbol{\beta} + \tilde{\epsilon}_{it},$$

where the tilde denotes that the respective variable has been de-trended. The FEIS estimator is given by

$$\widehat{\boldsymbol{\beta}}_{FEIS} = \boldsymbol{\beta} + \left(N^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}\tilde{\boldsymbol{x}}'_{1it}\,\tilde{\boldsymbol{x}}_{1it}\right)^{-1}\left(N^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}\tilde{\boldsymbol{x}}'_{1it}\,\tilde{\epsilon}_{it}\right) \quad .$$

In contrast to the FE (or FEGS) model, $\boldsymbol{a}_i$ (or $\boldsymbol{d}_i$) is not part of the error term. Strict exogeneity of ϵ_{it} is thus sufficient for consistency.

Monte Carlo Simulation

For each of 1,000 replications, panel data are set up with *i*=1,…,1000 units, each observed at *t*=0,…,9. The data are then generated according to the following process

$$y_{it} = \beta x_{it} + \alpha_{1i} + \alpha_{2i}\, t + \varepsilon_{it}\ ,$$
$$x_{it} = \mathrm{I}[P(\theta_{1i} + \theta_{2i} t + \theta_3 \alpha_{2i} + \theta_4 \alpha_{2i} t) > 0.5]\ ,$$

where y_{it} is the outcome of unit *i* at time *t* and ε_{it} is a Gaussian error term. The true treatment effect equals β. In the simulations, we set $\beta = 1$. Further, we set $\alpha_{1i} = 0$ to focus on the bias due to α_{2i}. α_{2i} is an individual-specific and time-constant random variable that is normally distributed with mean and variance equal to 1, that is, $\alpha_{2i} \sim N(1,1)$ for all scenarios.

Finally, selection into time-varying treatment x_{it} is modeled by a binary indicator variable I that equals 1 if the probability of treatment P is larger than .5, and 0 otherwise. We assume that P follows the logistic distribution function. The set of parameters $\theta_{1i}, \theta_{2i}, \theta_3, \theta_4$ is used to vary selection into treatment according to the following scenarios:

Scenario (1): $\theta_{1i} \sim N(-0.9, 0.1),\ \theta_{2i} \sim N(0.1, 0.1),\ \theta_3 = \theta_4 = 0.$
Scenario (2): $\theta_{1i} \sim N(-0.9, 0.1),\ \theta_{2i} \sim N(0.1, 0.1),\ \theta_3 = 0.1,\ \theta_4 = 0.$
Scenario (2'): as scenario (2), but $x_{it} = 1$ if $treat_i\ = 1$ and $t > U \sim [0,9]$
(timing of treatment assigned randomly).
Scenario (3): $\theta_{1i} \sim N(-0.9, 0.1),\ \theta_{2i} \sim N(0,0),\ \theta_3 = \theta_4 = 0.1.$

The simulation results (shown in Table S1) document the vulnerability of FE and FEGS models to individual-specific outcome trajectories that are related to treatment.

Scenario (1): All three estimators (FE, FEGS, and FEIS) provide the true value, as expected. Here, the outcome (e.g., the wage) follows an individual-specific time-trend, α_{2i}, but this trend is not systematically related to treatment (e.g., marriage), because $\theta_3 = \theta_4 = 0$. In this situation, it is sufficient to control for a common time-trend as in a standard FE model. The results for this scenario show that all three models can handle individual time-trends in the treatment variable (θ_{2i}), as long as these trends are not systematically related to the outcome.

Scenario (2): FE and FEGS are biased, and FEIS is unbiased. In this setting, individual time-trends in the outcome are related to treatment (because $\theta_3 \neq 0$). From the discussion above, it is clear that FE cannot handle this situation. It may be surprising that the FEGS estimates are also biased in this setting. After all, FEGS should be inconsistent only if α_{2i} affects the slope of the time-trend of treatment (i.e., the change of the probability of treatment depends on α_{2i}). FEGS should be consistent if α_{2i} merely shifts the level of the treatment variable up or down. However, the timing of treatment does vary by the value of α_{2i} due to construction of the treatment indicator. (We set $x_{it} = 1$ after the treatment probability exceeds .5. Because x_{it} is a discrete variable, units with high α_{2i} get treatment earlier even though the time-trend of the continuous treatment probability P does not depend on α_{2i}.) However, note that the bias of the FEGS estimator is much smaller than the bias returned by the standard FE model, because FEGS controls for mean differences in the treatment propensity between the treatment groups. Furthermore, the FEIS model identifies the true treatment effect. Results for scenario (2') show that the bias of FEGS is really due to the relation of α_{2i} to the timing of treatment. If we assign the timing of treatment at random, FEGS provides the true effect on average.

Scenario (3): The FE and FEGS estimators are inconsistent, and FEIS is consistent. Here, the time-trend of the treatment variable explicitly varies by the individual level of α_{2i} (because $\theta_4 \neq 0$). Thus, in both the outcome equation and the equation for selection into treatment, α_{2i} interacts with t. In other words, the same unobserved variables that drive the growth of the outcome also affect the timing of treatment. For example, persons with higher values on α_{2i} may not only experience stronger wage growth, but may also marry earlier. In this scenario, the bias of FE is huge and even FEGS is far from the truth.

Taken together, the simulations confirm the analytic results. In short, both standard FE and FEGS are biased if the timing of marriage is related to the same unobserved variables that also drive the steepness of the individual wage trajectory. As we note in the methods section of the main text, FEGS reduces the bias of the standard FE model considerably, because the model allows for group-specific trends in the wage and in marital selection. However, the model is still biased if not only the decision of whether to marry, but also the decision of when to marry, is related to the steepness of the career. FEIS does not need additional assumptions for individual deviations from an estimated wage trajectory, because each man is allowed to have his own trajectory. As a result, FEIS provides consistent estimates of the MWP even if the marriage propensity or marriage timing is related to individual wage growth. This is why FEIS is our preferred choice for estimation of the MWP.

Table S1. Simulation results demonstrating the bias of FE and FEGS due to individual-specific slopes

Estimator	(1)	(2)	(2’)	(3)
FE	1.004	1.689	1.569	6.768
	(.247)	(.235)	(.232)	(.165)
FEGS	.997	1.131	1.002	1.329
	(.152)	(.138)	(.168)	(.108)
FEIS	.998	1.001	.998	1.000
	(.054)	(.052)	(.051)	(.051)
Number of units (*N*)	1,000	1,000	1,000	1,000
Number of time points (*T*)	10	10	10	10
Number of observations (*NT*)	10,000	10,000	10,000	10,000
Prop. of treated units	50.0	54.3	54.3	54.0
Prop. of treated observations	20.1	23.9	29.4	25.5

Note: Simulation results (1,000 replications) for standard fixed-effects model (FE), fixed-effects model with group-specific slopes (FEGS), and fixed-effects model with individual-specific slopes (FEIS). The table shows mean of regression coefficients and mean of panel-robust standard errors (in parentheses).

Part B. Simulation results showing consistency of the FEIS estimator for a time-varying treatment

Table S2. Monte Carlo simulations for FEIS model with time-varying binary treatment

	Linear impact function		Dummy impact function	
	b	s.e.	b	s.e.
Treated (ref.: not treated)	.999	.181		
Time since treatment	.101	.172		
Time since treatment 0 (ref.: not treated)			.998	.215
Time since treatment 1			1.101	.369
Time since treatment 2			1.197	.596
Time since treatment 3			1.299	.898
Time since treatment 4			1.399	1.281
Time since treatment 5			1.498	1.746
Time since treatment 6			1.596	2.296
Time since treatment 7			1.691	2.935
Time since treatment 8			1.790	3.674

Note: Simulation results (1,000 replications) for fixed-effects model with individual-specific slopes (FEIS), mean of regression coefficients (b), and panel-robust standard errors (s.e.). Table shows estimates for two specifications of time-varying treatment effects: *linear impact function* specifies effect of binary treatment indicator and linear effect of time since treatment; *dummy impact function* specifies effect of dummies for each time period after treatment.

Simulation setup. For each replication, panel data are set up with *i*=1,…,200 units observed at *t*=1,…,10 points in time. 100 units get a binary treatment. Timing of the treatment is assigned randomly (uniformly distributed over *t*).

Data generating process. Data are generated according to the following equation

$$y_{it} = \alpha_{2i}\, t_{it} + \alpha_{3i}\, t_{it}^2 + \beta_1\, x_{it} + \beta_2\, xd_{it} + \alpha_{1i} + \varepsilon_{it}\ ,$$

where y_{it} is the outcome of unit *i* at time *t*, x_{it} is a binary treatment indicator (=1 after treatment), xd_{it} is time since treatment, and ε_{it} is a Gaussian error term. The true treatment effect is time-varying. Variation of the effect over time is specified by parameters β_1 and β_2. At zero duration ($xd_{it} = 0$), the treatment effect is $\beta_1 = 1$; each period after treatment, the effect increases by $\beta_2 = 0.1$.

$\alpha_{1i}, \alpha_{2i}, \alpha_{3i}$ are time-constant (unobserved) variables that are normally distributed with means that differ by treatment group (variances are set to the same values).

Never-treated units: $\alpha_{1i} \sim N(0{,}0.1)$, $\alpha_{2i} \sim N(0{,}0.01)$, $\alpha_{3i} \sim N(0{,}0.001)$.

Ever-treated units: $\alpha_{1i} \sim N(1{,}0.1)$, $\alpha_{2i} \sim N(0.1{,}0.01)$, $\alpha_{3i} \sim N(-0.01{,}0.001)$.

Thus, α_{1i} are unit-specific constants and produce a time-constant difference in mean outcomes between treatment groups. α_{2i} and α_{3i} are unit-specific slopes for time. They produce a time-varying difference in mean outcomes between treatment groups (non-parallel trends).

Part C. Full regression results for descriptive evidence on wage profiles (Figure 2) and time-path of the MWP (Figure 4)

Table S3. FE, FEGS, and FEIS models with time-varying estimate of the MWP

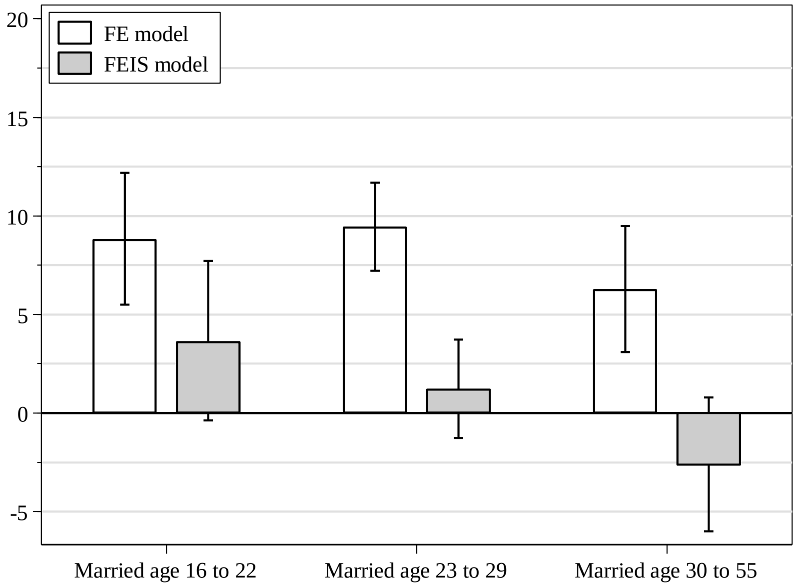
	FE	FEGS	FEIS	FEIS w/o controls
Years in 1st marriage: 1 year	.054***	.028**	–.003	.011
(ref.: never-married)	(.009)	(.009)	(.010)	(.010)
2 years	.078***	.046***	.000	.016
	(.010)	(.011)	(.013)	(.013)
3 years	.083***	.045***	–.007	.011
	(.011)	(.012)	(.015)	(.015)
4 years	.101***	.057***	–.002	.022
	(.012)	(.013)	(.018)	(.018)
5 years	.108***	.058***	–.015	.013
	(.013)	(.014)	(.020)	(.020)
6 years	.100***	.043**	–.027	.005
	(.014)	(.016)	(.022)	(.022)
7 years	.120***	.058***	–.026	.013
	(.015)	(.017)	(.026)	(.025)
8 years	.121***	.053**	–.038	.006
	(.016)	(.019)	(.028)	(.028)
9 years	.137***	.063**	–.041	.008
	(.017)	(.020)	(.031)	(.031)
10 years	.114***	.035	–.064*	–.008
	(.018)	(.021)	(.032)	(.033)
11 to 15 yrs.	.138***	.045	–.071	–.001
	(.018)	(.024)	(.037)	(.038)
Currently enrolled	–.201***	–.199***	–.123***	–.123***
	(.010)	(.010)	(.010)	(.010)
Years of education	.068***	.066***	.007	.058***
	(.004)	(.004)	(.006)	(.004)
One child (ref.: no child)	.009	.010	–.015	
	(.009)	(.009)	(.010)	
Two children	.012	.015	–.023	
	(.012)	(.012)	(.016)	
Three or more children	–.028	–.025	–.045	
	(.019)	(.019)	(.025)	
Tenure (years)	.011***	.011***	.008***	
	(.001)	(.001)	(.001)	
Work experience (years)	.043***	.035***		
	(.003)	(.003)		
Experience ^ 2	–.061***	–.045***		
(divided by 100)	(.009)	(.010)		
Ever-married X Experience		.015***		
		(.003)		
Ever-married X Exp. ^ 2		–.026*		
(div. by 100)		(.011)		
Within *R* squared	.34	.34	.02	.05
Number of persons	4,287	4,287	4,287	4,287
Number of person-years	49,801	49,801	49,801	49,801

Source: NLSY79.
Note: Table shows regression coefficients and panel-robust standard errors (in parentheses) from fixed-effects (FE), fixed-effects group-specific slopes (FEGS), and fixed-effects individual-slopes (FEIS) models. All models include indicators of grouped survey years (coefficients not shown). *FEIS w/o controls*: specification does not include controls for tenure and number of children; individual slopes specified for potential work experience (age – years of education – 5) instead of actual experience. *FEGS*: F-Test for joint significance of interactions "Ever-married X Experience" and "Ever-married X Exp. ^ 2": $F(2, 4286) = 16.95$, $p < .001$.
*$p < .05$; **$p < .01$; ***$p < .001$ (two-sided test).

Part D. Robustness checks

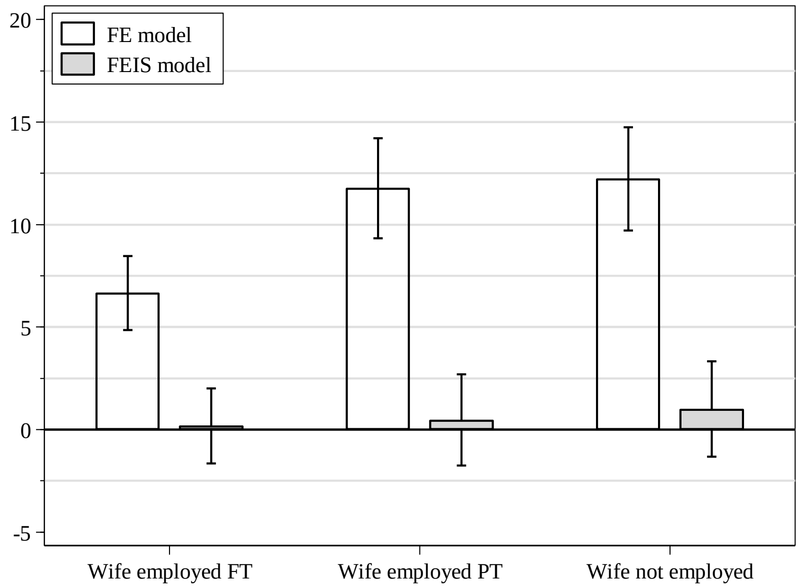
Figure S1. Test for heterogeneous effects of marriage by age at marriage, wife's employment, educational achievement, ethnicity, and urbanicity (FE and FEIS results)

a. Age at marriage



Killewald and Lundberg (2017) show that men who marry early experience the strongest wage growth. This finding is in line with our argument that these men are promising candidates for marriage. Our own results confirm this interpretation: With FEIS, the MWP is at 3.6 percent for those marrying very young (before age 23), 1.3 percent for those marrying at typical age, and -2.5 percent for those marrying late (after age 30). Although these results suggest there may be some effect heterogeneity by age at marriage, none of the effects is significantly different from zero. Even for men marrying early, the MWP found with FEIS is small compared to the 8.8 percent estimated with FE. We would therefore argue that, regardless of the timing of marriage, marital selection on wage growth mainly explains the wage benefit found with FE models.

b. Wife's employment



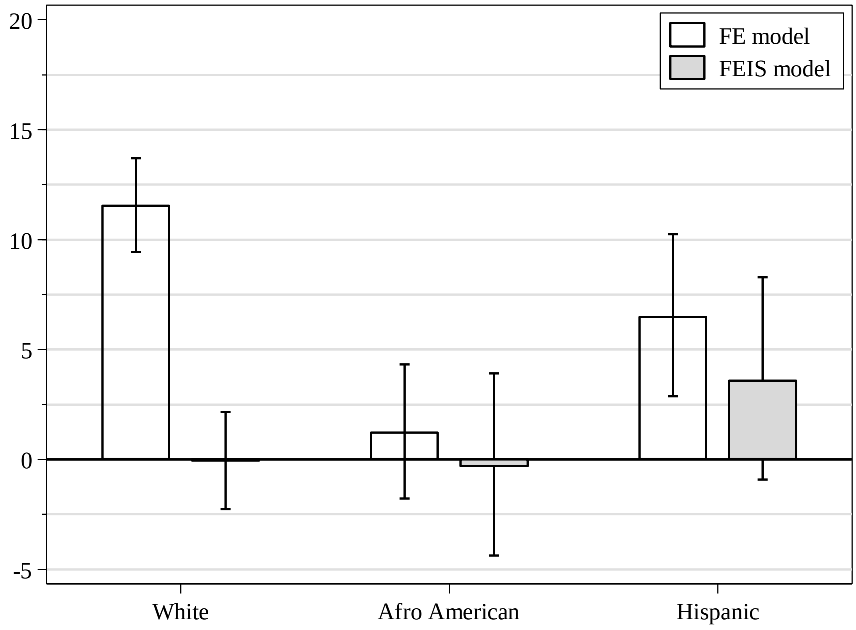
The literature consistently shows a stronger MWP for men whose wife is employed less than full-time (including the case of a non-working wife). Killewald and Gough (2013) report a premium that is 4.2 percent higher in this case. Similarly, Budig and Lim (2016) estimate an MWP of 11 percent for male breadwinners, which is about twice as large as the premium for men in dual-earner and female breadwinner households. However, these studies used conventional FE models. We suspect that wife's employment status is endogenous because women married to a man with strong wage growth are more likely to reduce their own market work hours. Indeed, our results are consistent with this explanation.

As reported in earlier studies, we find that wife's employment moderates the marriage premium in the FE model. Controlling for men's own employment status, the MWP is higher while the spouse is not working (12.4 percent) or working part-time (12 percent) than if she is working full-time (6.7 percent). As expected by the specialization argument, the MWP seems to increase as the wife reduces her market work. However, the strong premium for husbands even in dual-career couples is an unexpected finding that contradicts the claim that the MWP is limited to traditional marriages.

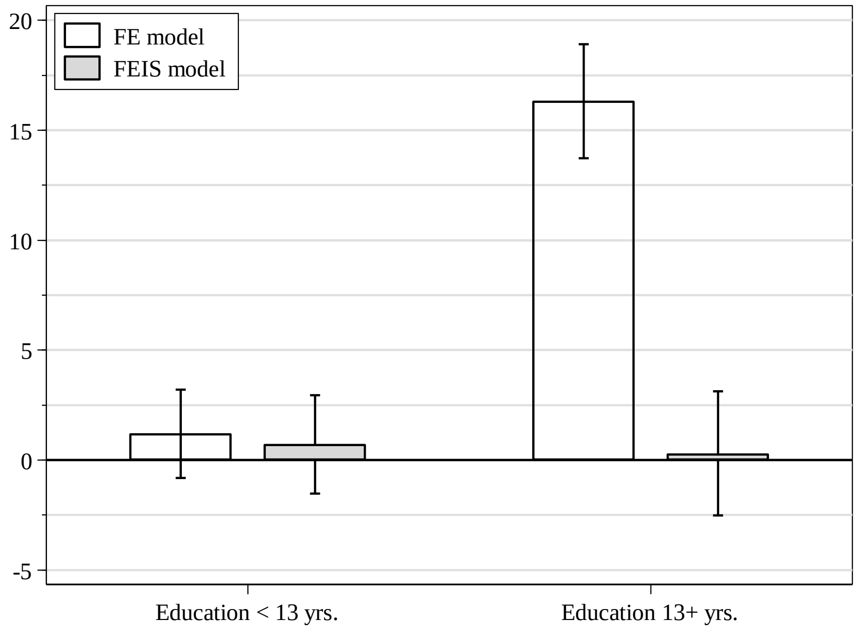
The results of the FEIS model help to interpret the FE results. If women marry men who are on a steep wage trajectory (promising men) regardless of their own career, this produces a spurious MWP also in dual-career couples. According to FEIS results the premium is close to zero and not significant regardless of the wife's employment status. Hence, even a homemaking spouse does not raise husbands' wages. Rather, it seems that causality runs the other way around: being married to a man with strong wage growth induces women to reduce employment (perhaps because of the opportunity to devote more time to childcare). This explains why FE estimates a larger MWP if the wife works less than full-time. Again, the evidence contradicts the specialization argument.

It has been suggested to extend the FE model using an instrumental variables (IV) approach to deal with endogeneity of wives' employment (Jacobsen and Rayak 1996). Our results indicate that the strong assumptions of this approach are not necessary if we use the FEIS model.

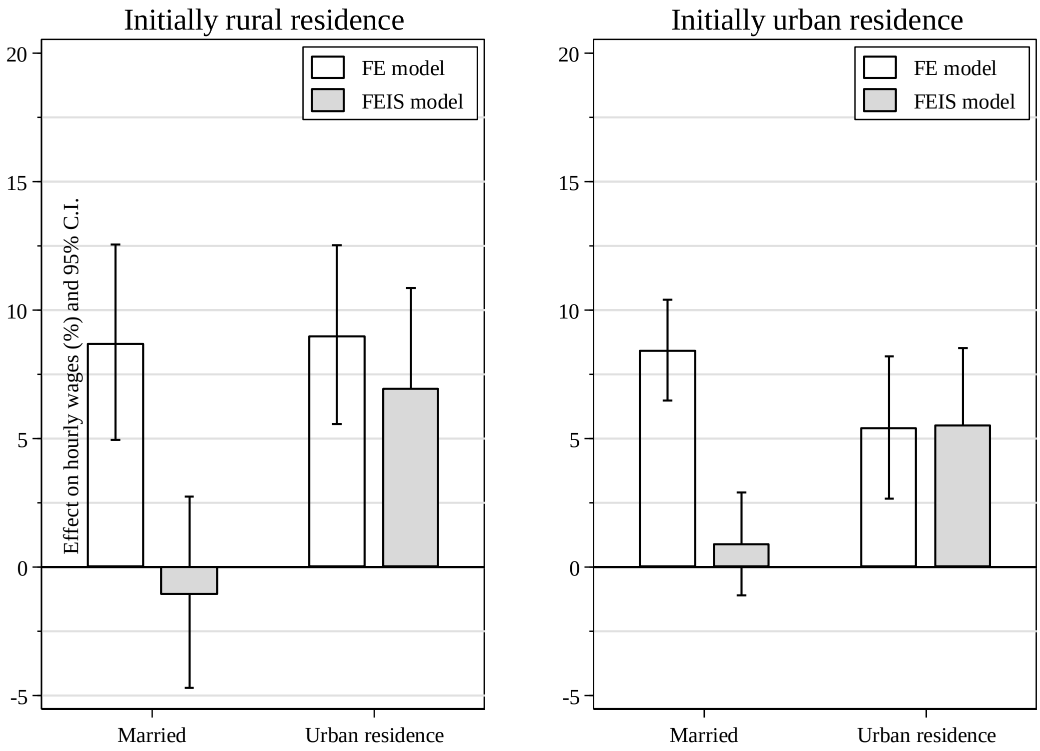
c. Ethnicity



d. Educational achievement



e. Urbanicity



Source: NLSY79 data.
Note: Marital wage premium estimated by fixed-effects (FE) and fixed-effects individual-specific slopes (FEIS) models including interaction effects. Regression models include as further covariates number of biological children (four categories), tenure with current employer, years of education, indicator for persons currently enrolled in education (reference: not in education), and survey year dummies (grouped, seven categories). Confidence intervals (C.I.) based on panel-robust standard errors. *Figure b. (Wife's employment)*: Husband's employment status is controlled in the models (indicators for full-time employment [reference], part-time employment, marginal employment, enrolled in education); wife's employment is derived from her annual work hours in the past calendar year. *Figure e. (Urbanicity)*: Models are estimated separately by first observed (initial) status of a man having rural or urban residence.

Table S4. Regression models including selection indicator, test of sample selection

	POLS	FE	FEGS	FEIS
Selection indicator *t*+1	–.027***	–.014*	–.014*	.005
(ref.: person included *t*+1)	(.008)	(.007)	(.007)	(.007)
Married	.163***	.084***	.047***	.006
(ref.: never-married)	(.011)	(.009)	(.009)	(.010)
One child (ref.: no child)	–.000	.017	.011	–.015
	(.011)	(.009)	(.009)	(.010)
Two children	.035*	.032**	.014	–.025
	(.014)	(.012)	(.012)	(.016)
Three or more children	–.030	.000	–.023	–.057*
	(.021)	(.019)	(.019)	(.026)
Currently enrolled	.077***	.066***	.064***	.006
(ref.: not enrolled)	(.002)	(.004)	(.004)	(.006)
Years of education	–.194***	–.195***	–.194***	–.120***
	(.009)	(.010)	(.010)	(.010)
Tenure (years)	.018***	.012***	.011***	.009***
	(.002)	(.001)	(.001)	(.002)
Work experience (years)	.050***	.044***	.035***	
	(.003)	(.003)	(.003)	
Experience ^ 2	–.071***	–.058***	–.042**	
(divided by 100)	(.012)	(.010)	(.013)	
Ever-married X Exp.			.017***	
			(.003)	
Ever-married X Exp. ^ 2			–.031*	
(divided by 100)			(.013)	
R-squared	.35	.33	.33	.02
Number of persons	3,990	3,990	3,990	3,990
Number of person-years	44,623	44,623	44,623	44,623

Source: NLSY79 data.

Note: Regression coefficients and panel-robust standard errors (in parentheses). Pooled OLS (POLS), fixed-effects (FE), fixed-effects group-specific slopes (FEGS), and fixed-effects individual-specific slopes (FEIS) models including a selection indicator. Selection indicator for *t*+1 equals 1 if a person is not contained in the estimation sample in the next year (0 otherwise). Models further include indicators of grouped survey years (seven categories). Estimation sample excludes the last wave of the NLSY79, for which the selection dummy is not defined; persons with fewer than four person-years excluded. Reported *R*-squared is overall *R*-squared for POLS and within *R*-squared for FE, FEGS, and FEIS models.

*$p < .05$; **$p < .01$; ***$p < .001$ (two-sided test).

Table S5. Logistic regression model for test of attrition bias in the NLSY79

	b	se	AME
Later married (ref. never-married)	–.662***	.189	–.138***
Ever-married	–.741***	.087	–.152***
Log hourly wage 2nd quintile (ref. 1st quintile)	.152	.120	.029
3rd quintile	.100	.120	.019
4th quintile	.103	.125	.019
5th quintile	–.034	.130	–.006
Number of children	–.015	.144	–.003
Currently enrolled	.042	.093	.008
Years of education	.006	.028	.001
Tenure (years)	.043	.062	.008
Work experience (years)	.031	.063	.006
Age (years)	–.036	.029	–.007
Birth cohort 1958 (ref. 1957)	–.237	.168	–.051
Cohort 1959	–.394*	.174	–.082*
Cohort 1960	–.423*	.175	–.088*
Cohort 1961	–.528**	.184	–.107**
Cohort 1962	–.396*	.188	–.082*
Cohort 1963	–.625**	.191	–.124**
Cohort 1964	–.561**	.195	–.113**
Male white poor subsample (ref. male white)	–.310	.213	–.059
Male black	–.683***	.181	–.118***
Male Hispanic	.121	.181	.025
Supplementary male black	–.351***	.106	–.066***
Supplementary male Hispanic	–.111	.116	–.022
Military male white	1.684***	.249	.388***
Military male black	.648	.368	.145
Military male Hispanic	–.146	.660	–.029
Constant	.537	.587	
Pseudo *R* square		.041	
Number of observations		3,788	

Source: NLSY79 data.
Note: Dependent variable is an attrition indicator, equals 1 if person drops out of the NLSY79 before 2012. Table shows logit coefficients (b), standard errors (se), and average marginal effects (AME). Cross-sectional sample of men included in sample 1; 499 men from the supplementary male white poor subsample are excluded because subsample was discontinued after 1990.
**p* < .05; ***p* < .01; ****p* < .001 (two-sided test).

Table S6. The average marital wage premium, test of attrition bias by weighted estimation

	FE	FEIS
Married	.092***	.006
(ref.: never-married)	(.009)	(.01)
Number of children	.002	–.019**
	(.005)	(.007)
Currently enrolled	–.208***	–.125***
	(.011)	(.010)
Years of education	.068***	.006
	(.004)	(.006)
Tenure	.116***	.086***
(years, divided by 10)	(.012)	(.014)
Work experience	.445***	
(years, divided by 10)	(.026)	
Experience squared	–.059***	
(divided by 100)	(.009)	
Within *R*-squared	.34	.02
Number of persons	3,788	3,788
Number of person-years	45,878	45,878

Source: NLSY79 data.

Note: Regression coefficients and panel-robust standard errors (in parentheses). Fixed-effects (FE) and fixed-effects individual-specific slopes (FEIS) models with inverse probability weights used to correct for attrition bias. Attrition weights are computed from models shown in Table S5. Individual attrition weights are computed as $\frac{(1-\hat{p}_r)}{(1-\hat{p}_u)}$, where $\hat{p}_u$ are predicted probabilities from the (unrestricted) model shown in Table S5, and $\hat{p}_r$ are predicted probabilities from a (restricted) model including indicators for NLSY79 subsamples only. Values of all time-varying covariates are taken from the first person-year included in the NLSY79 estimation sample. Models further include indicators of grouped survey years (seven categories). 499 men from the supplementary male white poor subsample are excluded from the NLSY79 sample because subsample was discontinued after 1990.

**p* < .05; ***p* < .01; ****p* < .001 (two-sided test).

Killewald and Gough (2013) were concerned with sample selectivity due to non-employment and item non-response on wages. They applied simple longitudinal imputation using wages of men observed one year later (or earlier if not observed). Although this is not a perfect test, their results suggest that selectivity does not bias the marital premium. However, panel attrition might bias our results if it is systematically related to wages and marriage. We ran a test for attrition bias in two steps.

First, we estimated a cross-sectional logit model where we included the first person-year of each man contained in our estimation sample. A binary attrition indicator is the dependent variable of the model. We find that attrition strongly depends on treatment status, with ever married men, on average, being 15 percentage points less likely to drop out than never-married men (see Table S5). However, there is no evidence of an association of wages and attrition.

Second, we computed attrition weights and used them to correct for attrition bias in FE and FEIS models. The results show only minor changes if we take attrition weights into account, with estimates for the MWP still insignificant and close to zero (see Table S6).

Finally, we estimated FE and FEIS models with a less restricted sample than we used in the main analyses. For this larger sample, we still required men to be once observed never-married, but we did not apply further restrictions on employment, duration of first marriage or marital status. To model the MWP for first marriage using this larger sample, we also modified the specification. We introduced additional indicators for divorce and remarriage in the models. In addition, we included person-years of self-employed men and we did not restrict the sample to men for whom we know they are currently working. Otherwise the same restrictions apply as for the sample used in the main analyses. We added dummy variables to the specification that capture wage differences of the self-employed and persons who are currently not working.

Table S7. The average marital wage premium, test for sample selection bias using less restricted sample

	Less restricted sample		Polynomials work experience, tenure, education		At least four years pre-treatment	
	FE	FEIS	FE	FEIS	FE	FEIS
Married	.090***	.028***	.082***	.017	.078***	.014
(ref.: never-married)	(.008)	(.009)	(.008)	(.009)	(.010)	(.010)
Separated / divorced	–.002	.004	.000	–.005	–.006	.000
	(.013)	(.014)	(.013)	(.015)	(.016)	(.018)
Widowed	–.044	.043	–.025	.050	–.016	.022
	(.058)	(.056)	(.060)	(.063)	(.085)	(.077)
Remarried	.072***	.019	.070***	.008	.071***	.019
	(.016)	(.018)	(.016)	(.020)	(.021)	(.025)
One child (ref.: no child)	.034***	–.010	.033***	–.006	.033***	–.018
	(.008)	(.009)	(.008)	(.009)	(.010)	(.010)
Two children	.070***	–.018	.067***	–.013	.073***	–.036*
	(.012)	(.013)	(.012)	(.014)	(.014)	(.015)
Three or more children	.049**	–.056**	.049**	–.046*	.046*	–.090***
	(.016)	(.019)	(.016)	(.021)	(.019)	(.023)
Currently enrolled	–.217***	–.124***	–.209***	–.109***	–.221***	–.131***
	(.009)	(.009)	(.009)	(.009)	(.011)	(.010)
Currently self-employed	–.062***	–.001	–.067***	.007	–.060**	–.003
	(.016)	(.016)	(.016)	(.017)	(.019)	(.019)
Currently working	.070***	.030***	.063***	.023**	.067***	.027***
	(.007)	(.007)	(.007)	(.007)	(.008)	(.008)
Years of education	.074***	.016***	–.117***	–.120***	.075***	.017***
	(.004)	(.005)	(.019)	(.030)	(.004)	(.005)
Years of education ^ 2			.689***	.495***		
(divided by 100)			(.068)	(.112)		
Tenure (years)	.012***	.011***	.026***	.018***	.012***	.009***
	(.001)	(.001)	(.001)	(.002)	(.001)	(.001)
Tenure ^ 2			–.070***	–.053***		
(divided by 100)			(.006)	(.012)		
Work experience (years)	.044***		.044***		.046***	
	(.002)		(.003)		(.003)	
Experience ^ 2	–.059***		–.093***		–.055***	
(divided by 100)	(.006)		(.022)		(.006)	
Experience ^ 3			.011**			
(divided by 1000)			(.004)			
R-squared	.30	.02	.32	.02	.31	.02
Number of persons	4,816	4,816	4,659	4,659	3,606	3,606
Number of person-years	78,611	78,611	77,983	77,983	60,026	60,026

Note: Regression coefficients and panel-robust standard errors (in parentheses). Estimation results from fixed-effects (FE) and fixed-effects individual-specific slopes (FEIS) models. Models further include indicators of grouped survey years (coefficients not shown).
Less restricted sample: Sample includes person-years after first marriage ends (due to separation, divorce, widowhood); person-years observed later than 15 years within first marriage; person-years while self-employed; person-years where respondent is currently not working, but hourly wage rate is recorded in the NLSY79 (based on earnings in the last job). Otherwise same selection criteria apply as for main estimation sample (see main text and Appendix, Table A1).
Polynomials experience, tenure, education: Less restricted sample, but men with fewer than five valid person-years excluded. Specification additionally includes cubic term for work experience, squared term for tenure, and education.
At least four years pre-treatment: Less restricted sample, but men with fewer than four valid person-years prior to first marriage excluded.
**p* < .05; ***p* < .01; ****p* < .001 (two-sided test).

The effect of first marriage returned by the FE model (Table S7, column 1) is similar to the estimate found with the main sample (9.4 percent). Further, we do find a significant positive effect of remarriage (7.5 percent), and no significant effect of divorce (-0.2 percent). These results seem to support the hypothesis that marriage promotes men's career. However, the findings do not hold with the FEIS model. Using FEIS (column 2), the effects of marriage and remarriage are much smaller (2.8 percent and 1.9 percent).

Although the MWP for first marriage is significant in the FEIS model, further tests imply that even this small effect is biased upwards. Firstly, the wage profile may not be approximated well by a quadratic function for work experience with a large proportion of person-years observed in later stages of the career. Hence, it may be necessary to specify effects of higher polynomials. If we extend the specification along these lines (including a cubic term for experience and a squared term for tenure and education), the MWP for first marriage is merely 1.7 percent and not significant (see column 4).

Secondly, if we use person-years observed over a long period after marriage, it follows that the panels of the respective persons are short pre-treatment. (In extreme cases, a person followed for 33 years after marriage in the NLSY79 can be observed only once prior to marriage.) Since the large, less restricted sample includes many persons with such treatment patterns, we cannot hope to get clean estimates of the marriage premium. Clearly, wages and covariates need to be observed pre-treatment to estimate an FE or FEIS model (This was the reason why we exclude men who enter the sample married.) However, since the FEIS model implicitly controls for individual wage growth that is independent of treatment, it might make sense to require even more than one person-year pre-treatment (see Morgan and Winship 2007). Four person-years pre-treatment would then be necessary to estimate an intercept and slopes for work experience (linear and squared) for each man. In fact, if we apply this restriction, the FEIS estimate of the MWP shrinks to 1.4 percent (and is no longer significant), even though later marriages are used for estimation (see column 6). In our main analyses, however, we did not want to apply this restriction, since it entails excluding nearly 40 percent of the men who actually are observed to marry. Instead, we included all these men, but restricted marriage duration to a maximum of 15 years. This restriction is sufficient to get reliable results. From a theoretical point of view, we also see no reason why it should take more than 15 years to discern an effect of household specialization on men's career. Moreover, the empirical pattern of the time-path of the effect of marriage contradicts the specialization argument during the first 15 years of marriage (see Table S3 and Figure 4).

Additional References

Jacobsen, Joyce P. and Wendy L. Rayack 1996. "Do Men Whose Wives Work Really Earn Less?" *The American Economic Review* 86(2) (Papers and Proceedings): 268-273.