The Relationship Between Spouses' Wages Over Time

Margaret Carroll
University of Wisconsin

Hyunjae Kang
Stony Brook University

Steven Stern*
Stony Brook University

July 2021

Abstract

Using predicted wages from 18 CPS data waves between 1962 and 2019, we use local linear regression estimates of each spouse's predicted wage as a function of his/her spouse's wage and show how they have changed over time. We look further into issues such as how much of assortative mating on wages can be explained by assortative mating on education and when, in a couple's lifetime, wages should be observed. We find strong evidence of assortative mating in both wages and education with assortative mating in education explaining much of the assortative mating in wages. We find that using wages at the time of marriage causes significant measurement problems as they do not reflect the true long-term expected variation in wages. Finally, we find that most of the assortative mating occurs at relatively low values of wages.

1 Introduction

In many ways, marriage has become more equal between partners, especially in labor market characteristics. Recent research supports the idea of positive assortative mating, that equilibrium in the marriage market results in positive correlation between the earnings of spouses (e.g., Cancian et al.,1993; Blackburn and Bloom, 1995; Cancian and Reed, 1999; Hyslop, 2001; Nakosteen and Zimmer, 2001; Schwartz, 2010; Bredemeier and Juessen, 2013), and this may be a result of a change in the structure of marriage itself. Instead of labor market specialization within families, we see more cooperative households that have more equal distributions of financial power (e.g., Cancian et al., 1993; Blackburn and Bloom, 1995; Gray, 1997; Cancian and Reed, 1999; Hyslop, 2001; Bianchi, Robinson, and Milkie, 2006). Alternatively, intra-family equality could be due to the changing role of women in the workforce; women's labor force participation has increased since the 1960's as has women's levels of education and wage levels (Goldin, 2006).

There are some papers in the literature that use only two or three points in time (Blackburn and Bloom, 1995; Hyslop, 2001; Zimmerman and Nakosteen, 2001; Bredemeier and Juessen, 2013); others use multiple years (Cancian and Reed, 1999; Schwartz, 2010). This paper examines assortative mating trends comprehensively, using frequent cross-sections back to 1962. The existing research predominantly focuses on earnings when determining the association between spouses' labor market characteristics (e.g., Nakosteen and Zimmer, 2001). However, as we discuss later, given the evidence for the effect of marriage on labor force participation for women especially in the early to mid-20th century (e.g., Bredemeier and Juessen, 2013), it is less useful to look at earnings as it is a function of time spent working.[1] In fact, assortative mating is a story about wages (Becker, 1973), and hours is a story about household specialization (Becker, 1973);[2] the two together determine earnings. We instead predict underlying wages of each spouse and measure association along labor market characteristics in terms of how spouses' underlying wages relate to each other.

We also examine how much sorting with respect to wages can be explained by education. Part of the motivation for this further analysis is the potential role education plays in marriage: that the social component of schooling leads many

*This project started as a research project by Carroll in Stern's undergraduate econometrics class. Carroll did all of the heavy lifting throughout the project until she became ill. We would like to thank Marina Azzimonti, Gabriel Mihalache, David Wiczer, and especially Yiyi Zhou for excellent suggestions. Stern is the corresponding author: (631)632-1328; steven.stern@stonybrook.edu.

[1] Lundberg, Pollak, and Wales (1997), Blundell et al. (2005), and Pollak (2005) argue that wages are also the most relevant measure associated with bargaining because earnings reflect hours decisions. Chiappori, Fortin, and Lacroix (2002) show how bargaining power can affect labor supply within a household. Phipps and Burton (1998) use income but spend no time justifying its use over wage. Lundberg, Startz, and Stillman (2003) use retirement (which is equivalent to using earnings in the context of this discussion) and also do not address income's endogeneity.

[2] The existence of household specialization and/or bargaining issues might affect female education and, therefore, wages (e.g., Konrad and Lommerud, 2000; Chiappori, Iyigun, and Weiss, 2009; Ge, 2011; Gihleb and Lifshitz, 2013) or might affect which women marry (e.g., Van der Klaauw, 1996; Mazzocco, Ruiz, and Yamaguchi, 2007).

couples to meet in that environment, leading to a natural correlation between spouses' education levels and therefore wages. If positive assortative mating along labor market characteristics is primarily driven by association in spouses' education, then it is possible that couples are actually sorting along social characteristics rather than marginal productivity in the labor market (Charles, Hurst, and Killewald, 2013).

Section 2 provides a literature review. Section 3 describes the data. It provides information on how the data sets for each year were constructed and shows moments and trends of some of the key variables. Section 4 describes the methodology of the paper. We start off using a standard Heckman (1979) 2-step estimation procedure to estimate log wage equations and predict log wages. Then, we describe how to use local linear regression to measure the relationship between the wages and education of husbands and wives and how they change over time. Finally, we propose how to test for positive assortative mating. Section 5 provides empirical results, and Section 6 concludes.

2 Literature Review

The changes in wage association come in part from changes in family roles: a movement away from a more traditional, specialized family structure towards a collaborative one. Becker's (1973) theoretical model of the traditional family structure included a head of household, typically male, who specializes in the labor market, and a partner, typically female, who specializes in home production. This model of family structure leads to economic characteristics of partners being negatively correlated, especially earnings, but non-economic characteristics that are positively correlated as people seek spouses similar to themselves in other ways (Becker, 1973, 1991). However, Becker (1973, 1991) provides no empirical support for the theory, so we look to other research to assess the validity of his claims.

Empirical analysis shows specialization becoming less common over time (Bianchi, Robinson, and Milkie, 2006; Bredemeier and Juessen 2013). Fertility, education, and labor market participation decisions have moved away from the gender roles outlined by Becker, and increasingly families are composed of dual-earners with spouses entering the marriage with similar earnings capabilities and maintaining that equality throughout the relationship (Schwartz, 2010). Goldin (2006) describes this change as one from women as secondary workers, whose labor supply was highly elastic with regard to household income, to one in which women's identities were formed before partnering, an identity that "placed career ahead, or on equal footing, as marriage."

These profound changes in 20th century women's economic lives have had implications for their social lives as well. The correlation between spouses' education levels has increased by 25% between 1960 and 2003 (Schwartz and Mare, 2005). Because high-productivity women get more education and have higher labor force participation rates, high-productivity couples have more opportunities to meet at school or work. Meanwhile, low-productivity men are less likely to meet high productivity women, and low-productivity women are less likely to meet high-productivity men. This naturally results in some similarities in the underlying wages among couples.

Much of the existing research on assortative mating on labor market characteristics focused on its ramifications for inequality. Schwartz (2010) estimates that spousal matchings across different earnings percentiles accounts for 25% to 30% of the earnings inequality among households; other estimates range from 17% to 51% (Cancian et al. 1993; Blackburn and Bloom 1994; Cancian and Reed 1999; Hyslop 2001, Schwartz 2010). Chiappori et al. (2020a, 2020b) and Greenwood et al. (2014) focus on constructing Gini coefficients under counterfactual matching patterns and find similar increases. However, this increase in inequality across families is paired with a decrease in inequality within a marriage (Chiappori, 1992; Cancian and Reed, 1998; Schwartz, 2010; Greenwood et al., 2014). Greenwood et al. (2014) perform a straightforward analysis of the change in inequality over time. Using the Gini index, they observe the changes in household income distribution in 2005 with 2005 mating preferences, 1960 mating preferences, and random mating, with either 2005 level labor force participation or 1960 level labor force participation. The authors observe that the positive assortative mating trends in 2005 increases the Gini coefficient relative to what it would be given 1960 mating patterns from 0.35 to 0.43. Furthermore, given the 1960 mating patterns, imposing the female labor force participation rates of the 1960's increases the Gini coefficient from 0.35 to 0.44, implying the 2005 female labor force participation rates had a dampening effect on household inequality. We focus on the similarity between spouses' underlying wages, rather than total earnings, to avoid the effects of joint decision making over hours. We do so comprehensively, taking advantage of the age of the ASEC and CPS surveys to look at repeated cross sections of data over time since 1960.

There is a large and growing literature discussing the implications of inequality within a household. Friedberg and Webb (2005) find that the distributions of wages, earnings, and unearned income between spouses affects allocation of chores and leisure. Some find that children in the family receive more care when the wife has more bargaining power (e.g., Schultz, 1990; Thomas, 1990, 1994; Haddad and Hoddinott, 1994; Duflo,2003; Duflo and Udry, 2004). Others find other effects of within-family bargaining power on household consumption, with increases in women's income increasing consumption of goods like children's clothing and women's clothing relative to men's (Lundberg, Pollack, and Wales, 1997; Phipps and Burton 1998; Lundberg, Startz, and Stillman, 2003; Duflo and Udry, 2004).

A small but growing literature discusses the dynamics of assortative mating. Gihleb and Lifshitz (2013) provide evidence that dynamic human capital accumulation plays an important role in the assortative mating issue. Van der Klaauw (1996) endogenizes the marriage decision and shows that it has significant effects on estimates of wage effects on female labor supply. Ge (2011) suggests that the effect of education on finding a good mate has significant explanatory power in explaining women's education choices, and Chiappori, Salanié, and Weiss (2017) show evidence of increasing returns to women's education in marriage markets.

Presently, Siow (2015) and papers such as Graham (2011) and Chiappori, Salanié, and Weiss (2011) provide methods to measure and test for assortative mating. We think there is still room for discussion about how to measure the amount of assortative mating, and we focus on many of the issues we see as unresolved. These include what variable to focus on (e.g., education, wage, earnings); how to measure the variable (e.g., levels, percentiles); when to measure the degree of assortative mating (e.g., the time of marriage, in a particular band of age, or not controlling for duration of marriage or age); and how to test for assortative mating. These are the issues discussed in this paper.

3 Data

The data are from the Annual Social and Economic Supplement (ASEC) to the Current Population Survey (CPS), a monthly survey conducted by the Bureau of Labor Statistics. They cover the years 1962, 1967, 1970, 1972, 1975, 1978, 1980, 1985, 1990, 1995, 2000, 2005, 2007, 2010, 2014, 2015, 2018, and 2019. We choose these years to capture incremental changes, with more detail for the years before and after the Great Recession. The year 1972 is included as a comparison point to deal with a change in the availability of important explanatory variables that year. We use the data sets available from the Integrated Public Use Microdata Series.

We include only heterosexual married couples within the ages 16 - 64.[3] Selection rules and their effects on sample size are detailed in Table 1. The marriage restriction is the most important source of selection. Due to the increasing non-married population, the effect of the selection rule concerning marital status rises over time. The impact of the selection rules changes significantly between 1967 and 1970. In 1967, the proportion of observations lost to *not married*[4] was 0.268, while, in 1970, it was 0.487. The proportion of observations lost to the age restrictions is relatively stable over time, which is 0.077 on average. The important selection rules concern age and marital status excluding more than half of the sample. The impact of the selection rules changes significantly between 1962 and 1970. In 1962, the proportions of observations lost to *<16 years*, *>64 years*, and *not married* were 0.056, 0.134, and 0.265, while, in 1970, they were 0.321, 0.093, and 0.196, respectively. From 1970 on, the proportion of observations lost to *<16 years* declines by 0.6% annually, the proportion lost to *>64 years* increases by 0.6% annually, and the proportion lost to *not married* rises by 1.1% annually. We exclude government workers, self-employed people, and people in the armed forces. The sample sizes are very large, resulting in precise estimation of the relevant parameters. For each year of data, it is necessary to construct the wage variable from last year's salary, weeks worked, and usual hours worked variables: hourly wage = annual earnings ÷ (hours/week*weeks/year).[5] Implausibly low wages, which we define as wages less than the real minimum wage in the given year, are omitted from the estimation sample. Wages are measured in 2014 real dollars.

Table 2 contains moments for each variable across each year, including a disaggregation by sex for the labor market variables. Real log wages are conditional on working. With respect to explanatory variables, there is a notable rise in the proportion of the sample that is Hispanic, rising from 4.5% in 1972 to 20.6% in 2019. The age distribution of Hispanics does not change much over the years. But the never-married/single proportion changes from 0.24 in 1972 to 0.41 in 2019. This is a bigger increase than the same change in the whole population which is from 0.23 to 0.35. The variable *Has Children Under 5* is not available until 1970, and *Hispanic* is not available until 1972. The increase of the Hispanic population is associated with the surge of Mexican immigrants (Borjas ,2006).

Also, the composition of education levels change significantly, as seen in Figures 1 and 2. Over the sample period, the

[3] The CPS data started capturing same-sex marriage presumably due to the legalization of same-sex marriage in 2014. Although there is no same-sex marriage observed prior to 2019, unmarried same-sex couples have been included in the data since 1990. The proportion of same-sex marriage in 2019 is around 0.008, and the proportion of same-sex couples is around 0.015. Jepsen and Jepsen (2002) use data from 1990 Census data to look at inequality within gay marriages.

[4] Variable names, such as *not married*, are displayed in a different font.

[5] There exists inconsistency within the variables used in constructing wage. The weeks-worked and the income variables are the information on the previous year while the hours-worked variable represents the average hours worked in the week before the survey.

Table 1: Data Selection

	1962	1967	1970	1972	1975	1978	1980	1985	1990
CPS Sample Size	71,741	68,676	145,023	140,432	130,124	155,706	181,488	161,362	158,079
Cause of Exclusion									
Age > 64 or Age of Spouse > 64	6,556	6,002	8,832	8,952	8,954	10,532	12,842	11,942	12,422
Age < 16 or Age of Spouse < 16	26	24	32	16	32	20	26	24	2
Not Married	18,521	18,436	70,648	67,896	62,181	75,159	87,796	78,371	77,382
Same-sex spouse	14	0	0	0	0	0	0	0	0
Unmarried Partner	0	0	0	0	0	0	0	0	0
Government, Self-Employed, and Armed Forces	13,587	12,586	18,961	18,951	17,595	22,225	26,369	22,169	23,598
Other Variable Missing	3,226	2,734	3,798	3,893	3,806	4,744	5,094	4,544	4,274
Remaining Sample Size									
Total*	25,282	25,214	36,214	34,326	31,390	35,620	40,626	35,942	31,876
# Employed	14,725	15,546	22,684	21,276	18,822	22,654	26,367	23,849	22,204
# Employed with Plausible Wage	8,498	12,606	19,455	18,621	16,133	19,372	22,683	21,535	20,770

Note: Total includes sample with implausible wage, which we exclude for estimation

Table 1: Data Selection (Continued)

	1995	2000	2005	2007	2010	2014	2015	2018	2019
CPS Sample Size	149,642	133,710	210,648	206,639	209,802	199,556	199,024	180,084	180,101
Cause of Exclusion									
Age > 64 or Age of Spouse > 64	11,714	10,344	13,404	13,556	14,674	16,926	17,060	17,502	18,398
Age < 16 or Age of Spouse < 16	8	18	12	4	10	12	16	8	4
Not Married	74,287	67,522	111,621	108,945	112,201	105,563	105,127	93,401	92,585
Same-sex spouse	7	5	15	8	38	54	69	172	505
Living with unmarried partner	183	231	377	363	441	341	389	344	346
Government, Self-Employed, and Armed Forces	21,629	17,920	29,742	29,150	27,747	24,427	23,933	21,013	20,725
Other Variable Missing	3,842	3,402	5,128	5,212	5,235	4,906	4,772	4,298	4,123
Remaining Sample Size									
Total	29,734	27,216	41,070	40,318	39,830	37,204	37,012	32,554	32,496
# Employed	20,957	19,684	29,530	29,190	27,563	26,148	26,112	23,542	23,764
# Employed with Plausible Wage	19,476	18,187	28,017	27,464	25,383	24,331	24,215	22,196	22,636

Table 2: Sample Moments

Variable	1962	1967	1970	1972	1975	1978	1980	1985	1990
Female Employment	0.223	0.242	0.281	0.298	0.303	0.350	0.385	0.464	0.546
Female Real log Wage	2.738	2.835	2.864	2.835	2.884	2.896	2.879	2.798	2.804
	(0.436)	(0.414)	(0.402)	(0.424)	(0.430)	(0.431)	(0.455)	(0.496)	(0.529)
Male Employment	0.891	0.909	0.902	0.883	0.822	0.844	0.839	0.817	0.825
Male Real log Wage	3.093	3.216	3.291	3.280	3.342	3.326	3.321	3.227	3.183
	(0.440)	(0.453)	(0.465)	(0.481)	(0.484)	(0.479)	(0.492)	(0.555)	(0.569)
National log(Real Minimum Wage)	2.208	2.325	2.334	2.235	2.310	2.336	2.313	2.031	1.855
Age	39.546	39.830	39.755	39.956	39.919	40.064	39.745	40.225	40.157
White	0.931	0.938	0.931	0.931	0.927	0.930	0.924	0.916	0.901
Black	0.063	0.055	0.062	0.060	0.060	0.051	0.051	0.053	0.059
Hispanic	---	---	---	0.045	0.050	0.089	0.089	0.109	0.130
Has Children Under 5	---	---	0.435	0.409	0.363	0.350	0.364	0.357	0.350
High School Diploma	0.349	0.405	0.423	0.434	0.438	0.431	0.440	0.439	0.428
Some College	0.095	0.106	0.119	0.122	0.137	0.154	0.161	0.171	0.188
College Degree	0.055	0.062	0.063	0.067	0.076	0.083	0.091	0.110	0.120
Post-Graduate Degree	0.019	0.022	0.025	0.029	0.036	0.041	0.045	0.056	0.058
Northeast	0.289	0.278	0.265	0.260	0.247	0.230	0.218	0.229	0.243
South	0.246	0.258	0.272	0.277	0.295	0.277	0.283	0.301	0.304
Midwest	0.304	0.299	0.300	0.293	0.291	0.262	0.246	0.244	0.234
Metropolitan	0.713	0.716	0.703	0.722	0.718	0.726	0.735	0.721	0.771
# Obs	22,398	21,002	31,222	30,314	27,276	30,426	34,712	32,380	29,858

Note: Numbers in parentheses are standard deviations.

Table 2: Sample Moments (continued)

Variable	1995	2000	2005	2007	2010	2014	2015	2018	2019
Female Employment	0.570	0.586	0.588	0.592	0.576	0.575	0.573	0.593	0.609
Female Real log Wage	2.856	2.931	2.972	3.005	3.066	3.040	3.062	3.101	3.103
	(0.580)	(0.608)	(0.635)	(0.652)	(0.645)	(0.617)	(0.671)	(0.684)	(0.665)
Male Employment	0.815	0.837	0.836	0.836	0.783	0.809	0.816	0.835	0.840
Male Real log Wage	3.161	3.225	3.266	3.289	3.328	3.281	3.286	3.330	3.338
	(0.615)	(0.647)	(0.681)	(0.699)	(0.681)	(0.689)	(0.703)	(0.709)	(0.711)
National log(Real Minimum Wage)	1.914	1.990	1.864	1.926	2.079	1.996	1.981	1.946	1.921
Age	40.827	41.655	42.008	42.508	43.128	43.899	43.925	44.088	44.156
White	0.873	0.893	0.851	0.846	0.833	0.823	0.820	0.809	0.814
Black	0.054	0.059	0.066	0.068	0.073	0.072	0.071	0.072	0.070
Hispanic	0.147	0.195	0.165	0.178	0.174	0.188	0.196	0.207	0.206
Has Children Under 5	0.342	0.322	0.345	0.348	0.331	0.295	0.305	0.295	0.303
High School Diploma	0.362	0.343	0.324	0.311	0.300	0.282	0.284	0.275	0.270
Some College	0.260	0.263	0.27	0.267	0.273	0.27	0.262	0.258	0.257
College Degree	0.150	0.165	0.19	0.198	0.210	0.219	0.222	0.238	0.240
Post-Graduate Degree	0.061	0.066	0.083	0.09	0.098	0.117	0.117	0.125	0.129
Northeast	0.241	0.198	0.204	0.195	0.203	0.186	0.170	0.155	0.159
South	0.306	0.312	0.299	0.308	0.305	0.314	0.348	0.368	0.362
Midwest	0.233	0.235	0.248	0.238	0.235	0.232	0.204	0.198	0.201
Metropolitan	0.778	0.795	0.803	0.819	0.823	0.833	0.832	0.844	0.845
# Obs	27,530	24,994	38,976	37,874	36,536	34,542	34,206	30,698	31,020

Note: Numbers in parentheses are standard deviations.

percentage of the sample with some post-high school education increases from 13.7% in 1962 to 65.1% in 2019 for women and from 20.1% in 1962 to 59.9% in 2019 for men. From 1962 to 1970, the big increase is in the proportion of people who finish high school. In later years, the increase is all of the other education levels. For example, the change in the percentage with at least a college degree is from 4.5% in 1962 to 38.4% in 2019 for women and from 10.2% in 1962 to 35.3% in 2019 for men.

Figures 3 and 4 show how employment rates and mean log real wages change over the sample period based on the moments in Table 2. Women's employment rates rise dramatically between 1962 and 1990 as documented in the literature (e.g., Goldin, 2006, Blau and Kahn, 2007; Bailey, Hershbein, and Miller, 2012). But, this by itself does not affect our analysis because our measure of analysis is wage rather than employment. On the other hand, Figure 4 shows large increases in wages for women, both absolutely and relative to men. However, unlike employment, most of the increase occurs between 1985 and 2010. The change in the log wage gap is shown in Figure 5 and is consistent with the literature (e.g., Blau, 1998; Eckstein and Nagypal, 2004; Blau and Kahn, 2017). The trend in log wages is the relevant one for this paper.

The weeks-worked variable in the CPS data before 1978 are collected by categories: $1-13$ weeks, $14-26$ weeks, $27-39$ weeks, $40-47$ weeks, $48-49$ weeks, and $50-52$ weeks. Papers using the CPS use the midpoint of each interval to represent the weeks-worked (e.g., Bailey, DiNardo, and Stuart, 2020). Derenoncourt and Montialoux (2020) use the midpoint and smooth the measure using random numbers generated from the uniform distribution. Using the fact that weeks-worked variables of later years are collected both as continuous and categorical variables, we identify the measurement errors from using the bracketed variable. Then we adjust the measurement errors in the weeks-worked variable nonparametrically. Using local-linear regression, we regress the measurement errors on total hours worked using later years' data, then use the prediction for the adjustment. For both female and male respondents, there exist unignorable errors whose annual hours worked were less than 2000. The magnitude of the measurement errors is displayed in the appendix.

4 Methodology

The goal of our work is to examine how the relationship between spouses' wages has changed over time. Our procedure does this in two stages separately for each year of investigation: first, we predict of underlying wages using a Heckman two-step procedure, and, second, we use local linear regression methods to examine the relationship between spouses' predicted wages. We consider at what point in across couples' marriage is a good time to measure assortative mating and consider measuring it in log wage levels and in percentiles from the joint wage distribution. Afterwards, we estimate how much of the sorting over wages is due to sorting over education and test for assortative mating against the null hypothesis of random sorting.

4.1 Heckman Procedure

We need to compare wage rates of all spouses, including those not in the labor force. The decision to work is determined in part by the wage a person would earn if she worked, thus causing the wage distribution of those who work to differ from

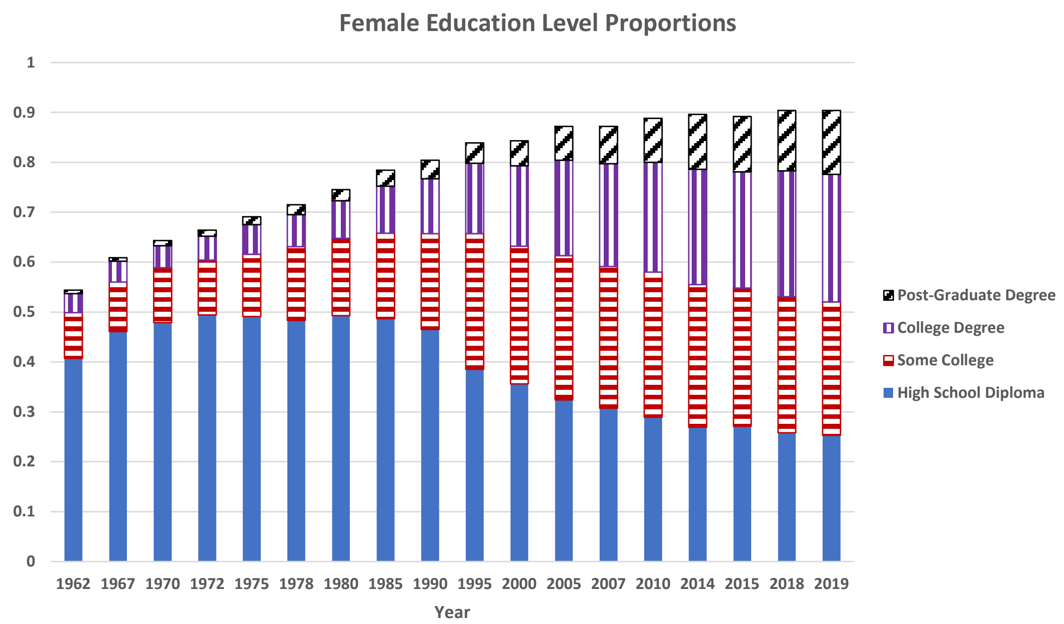


Figure 1: Female Education Level Proportions

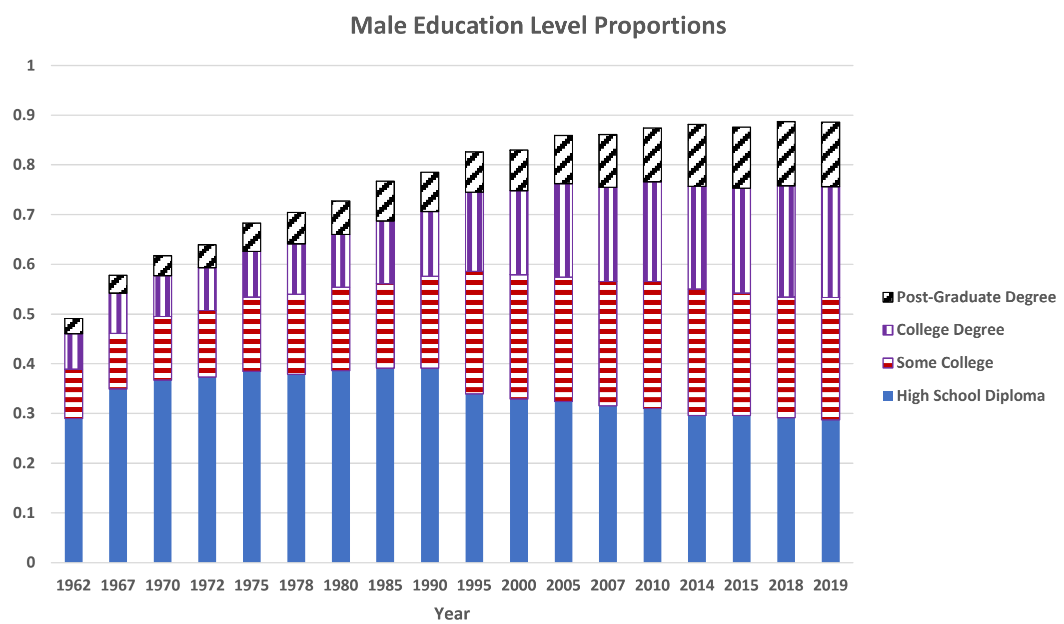


Figure 2: Male Education Level Proportions

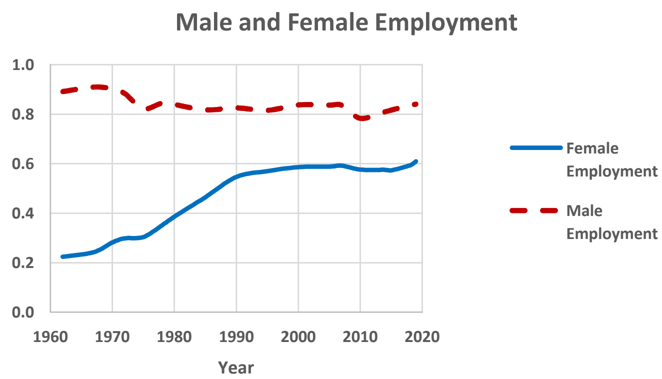


Figure 3: Male and Female Employment

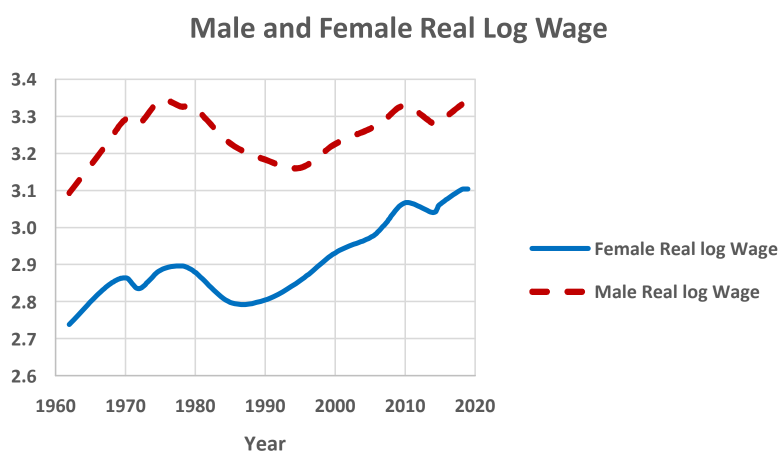


Figure 4: Male and Female Real log Wages

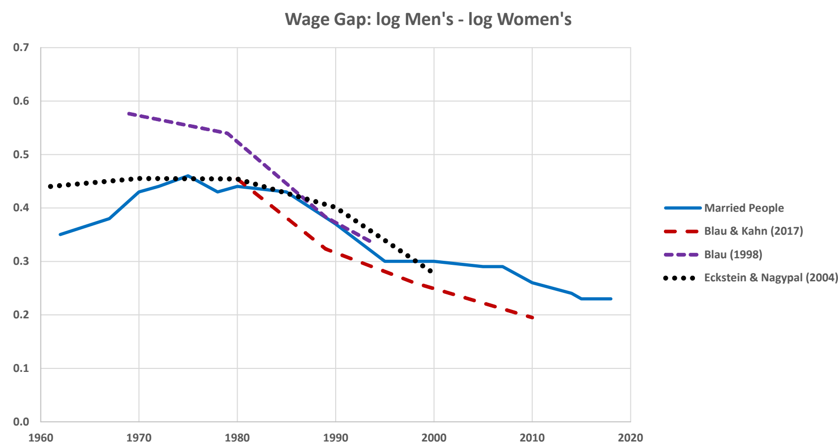


Figure 5: Wage Gap: log Men's - log Women's

Notes: The solid (blue) curve, labeled "Married People," is based on our data. The other three curves, each covering part of the relevant period, come from the literature. Blau and Kahn (2017) is based on the second panel of Table 1. Blau (1998) is based on Table 4. Eckstein and Nagypál (2004) is based on means in Figure 2. Eckstein and Nagypál (2004) requires a translation from the index in Eckstein and Nagypál (2004) to log wag gaps using 1980 as a matching year where the index is translated to log wages using Blau and Kahn (2017) as a match.

the wage distribution of those who do not work. Estimating wage equations using solely the people who work would lead to biased estimates (Heckman, 1979), and predicting potential wages of those who do not work even with consistent estimates still would lead to upward-biased predictions (e.g., Maddala, 1983). Thus, it is necessary to correct for selection bias during both estimation and prediction.

We use the Heckman (1979) two-step procedure to correct for selection bias. There is some evidence that estimates from the Heckman procedure are sensitive to the necessary joint normality assumption (e.g., Arabmazar and Schmidt, 1982; Lee, 1982; Olsen, 1982; Duncan and Leigh, 1985; Lalonde, 1986; Mroz, 1987; Newey, Powell, and Walker, 1990; Coslett, 1991; Ichimura, 1993; Stern, 1996). Some researchers have proposed semi-parametric methods to avoid the sensitivity (Lee, 1982; Olsen, 1982; Newey, Powell, and Walker, 1990; Coslett, 1991; Ichimura and Lee, 1993; Ichimura, 1993; Stern, 1996; Blundell et al., 2007; Escanciano, Jacho-Chavez, and Lewbel, 2016). However, the focus of this study is on the wage curves themselves, and we try to use standard methods to construct the wage curves. The Heckman procedure is used also, for example, by Hersch (1991) and Mulligan and Rubinstein (2008).

Despite it ubiquity, we provide notation for the Heckman procedure to aid in further discussion. Let

$$y^*_{tij} = X_{tij}\beta_t + u^y_{tij} \tag{1}$$

be the value of working for person j in family i from year t where X_{tij} is a set of exogenous explanatory variables and $u_{tij} \sim iidN(0,1)$. Let $j = 1$ for husband and $j = 2$ for wife. Included in X_{tij} are a constant, a quadratic polynomial in age, dummies for race and education, the presence of any children under 5 years of age,[6] regional dummies, and an urban dummy.[7] Recall that our data are repeated cross-sections; thus, family i in year t is not the same family as family i in a different year. We ignore the joint decision-making nature of the family's labor force participation decision (e.g., Gustman and Steinmeier, 1986; Chiappori, 1988; Blundell et al., 2005) and assume the binary participation decision for family member j is

$$y_{tij} = 1\left(y^*_{tij} > 0\right). \tag{2}$$

We estimate using probit.

The log wage one would receive if she[8] worked is modeled as

$$w_{tij} = Z_{tij}\gamma_t + u^w_{tij} \tag{3}$$

[6] Mulligan and Rubinstein (2008) use number of children times a dummy for marital status.

[7] One might worry about the endogeneity of young children (e.g., Rosenzweig and Wolpin, 1980; Schultz, 1990, 2010; Huber and Mellace, 2011) or education (e.g., Gihleb and Lifshitz, 2013). For education, there is much more concern about the endogeneity of education with respect to wage (e.g., Card, 2001). But neither issue is the point of the paper.

[8] We use "she" as the generic pronoun throughout.

where Z_{tij} is a set of exogenous explanatory variables,

$$u^w_{tij} \mid u^y_{tij} \sim iidN\left(\rho_t \sigma_{et} u^y_{tij}, \sigma^2_{et}\left(1-\rho^2_t\right)\right),$$

and ρ is the correlation of (u^w_{tij}, u^y_{tij}). Included in Z_{tij} are the variables in X_{tij} excluding the presence of young children.[9] As in Heckman (1979),

$$E(w_{tij} \mid y_{tij}=1) = Z_{tij}\gamma_t + \alpha_t \tau(X_{tij}\beta_t) \tag{4}$$

where $\tau(\cdot) = \phi(\cdot)/\Phi(\cdot)$ is the inverse Mills ratio and $\alpha_t = \rho_t\sigma_{et}$. Following Heckman (1979), we estimate the log wage equation parameters for each year (γ_t, α_t) using OLS and then adjust standard errors appropriately.

Given probit estimates $\widehat{\beta}_t$ and OLS estimates $(\widehat{\gamma}_t, \widehat{\rho}_t, \widehat{\sigma}_{et})$, we can predict log wages for each person in the sample, whether or not she works, as

$$w_{tij} = \begin{cases} Z_{tij}\widehat{\gamma}_t + \widehat{\rho}_t\widehat{\sigma}_{et}\tau(X_{tij}\widehat{\beta}_t) & \text{if } y_{tij}=1 \\ Z_{tij}\widehat{\gamma}_t + \widehat{\rho}_t\widehat{\sigma}_{et}\psi(X_{tij}\widehat{\beta}_t) & \text{if } y_{tij}=0 \end{cases} \tag{5}$$

where

$$\psi(X_{tij}\widehat{\beta}_t) = E\left(u^w_{tij} \mid y_{tij}=0\right) = -\frac{\phi\left(X_{tij}\widehat{\beta}_t\right)}{1-\Phi\left(X_{tij}\widehat{\beta}_t\right)}.$$

We use predicted log wages for each sample person, even if log wage is observed. In some sense, this has the opposite flavor of Nakosteen and Zimmer (2001) and Zhang and Lui (2003). Both papers measure whether there is assortative mating associated with residuals because Becker (1973) focuses on assortative mating over wages after controlling for other characteristics. However, residuals capture effects not observed by the researcher. There is no reason why potential couples would distinguish between those characteristics observed by the researcher and those not observed by the researcher. If the data became a little better (e.g., one more variable was added), do we think couples' behavior would change? On the other hand, it might be that couples sort based on some characteristics not observed by the researcher. Consider a wage equation similar to equation (3),

$$w_{tij} = Z_{tij1}\gamma_{1t} + Z_{tij2}\gamma_{2t} + Z_{tij3}\gamma_{3t} + u^w_{tij} \tag{6}$$

where Z_{tij} in equation (3) is decomposed into $(Z_{tij1}, Z_{tij2}, Z_{tij3})$ and where Z_{tij1} is observed in the data and observed by potential partners, Z_{tij2} is not observed in the data but is observed by potential partners, and Z_{tij3} is neither observed in the data nor by potential partners. The true object used for sorting is $Z^*_{tij} = Z_{tij1}\gamma_1 + Z_{tij2}\gamma_2$ which is different from both $Z_{tij1}\widehat{\gamma}_{1t}$ (which we use) and $w_{tij} - Z_{tij1}\widehat{\gamma}_{1t}$ (which is what Nakosteen and Zimmer (2001) and Zhang and Lui (2003) use).

4.2 Local Linear Estimation

After performing the two-step Heckman procedure to predict log wages, we use a local linear regression estimator to predict expected spouse's wage conditional on own predicted wage and spouse's gender, using the predicted wages for both those who work and those who do not.[10] Using a local linear regression estimator is a simple and flexible way to examine the relationship between spouses' wages and still see the variation in association over wage levels.[11] In particular, let $\widehat{w}^o_{tij}$ be j's "own" predicted log wage and $\widehat{w}^s_{tij}$ be j's spouse's predicted log wage. We define, for some arbitrary value $\left(\widehat{w}^s_{tj}, \widehat{w}^o_{tj}\right)$ of a couple's predicted log wage as

$$\widehat{E}\left[\widehat{w}^s_{tj} \mid \widehat{w}^o_{tj}\right] = \widehat{\varkappa}_{0t}\left(\widehat{w}^o_{tj}\right) \tag{7}$$

where[12]

$$\begin{pmatrix} \widehat{\varkappa}_0\left(\widehat{w}^o_{tj}\right) \\ \widehat{\varkappa}_1\left(\widehat{w}^o_{tj}\right) \end{pmatrix} = \left[\sum_i K\left(\frac{\widehat{w}^o_{tij}-\widehat{w}^o_{tj}}{b}\right)\begin{pmatrix} 1 \\ \widehat{w}^o_{tij} \end{pmatrix}\begin{pmatrix} 1 & \widehat{w}^o_{tij} \end{pmatrix}\right]^{-1} \cdot \left[\sum_i K\left(\frac{\widehat{w}^o_{tij}-\widehat{w}^o_{tj}}{b}\right)\begin{pmatrix} 1 \\ \widehat{w}^o_{tij} \end{pmatrix}\widehat{w}^s_{tij}\right] \tag{8}$$

[9] Common examples of variables included only in the participation equation are husband's income (e.g., Mroz, 1987; Zabel, 1993; Huber and Mellace, 2011) and characteristics of children (e.g., Martins, 2001; Mulligan and Rubinstein, 2008; Lee, 2009; Chang, 2011; Huber and Mellace, 2011).

[10] Suen and Lui (1999) and Fernandez, Guner, and Knowles (2005) use correlations between husbands' and wives' education levels to test for positive assortative mating.

[11] One could also test hypotheses of interest using nonparametric testing methods (e.g., Hardle, 1992).

[12] See, for example, Ichimura and Todd (2007) or Hansen (2021).

and

$$K(x) = \exp\left\{-.5x^2\right\} \tag{9}$$

is the kernel function and b is the bandwidth.[13] $\widehat{E}[\cdot \mid \cdot]$ is an estimate of the expected wage of a spouse given one's wage. $\widehat{E}[\cdot \mid \cdot]$ can be evaluated over all $\widehat{w}^o_{tj} \in W^o_{tj}$ where W^o_{tj} is a closed set bounded within the set of own predicted log wages observed in the year-t/gender-j sample. The $\widehat{E}[\cdot \mid \cdot]$ curves can vary across gender and across years thus allowing us to observe how assortative mating changes over time.[14]

4.3 Measuring the Amount of Wage Sorting Due to Education Sorting

As wage is highly correlated with other factors over which couples sort, we next determine how much of sorting over wages can be explained by sorting over education. We do this by regressing each individual's predicted wage given their spouse's predicted wage on (from the local linear regression estimation exercise) evaluated at deciles of the distribution of gender-specific education.

First, we order husbands by their level of education in each year t, and then we define husband k as the husband with the kth decile of education, $k = 1, 2, .., 20$. Next, we define (with a slight abuse of notation) $\widehat{w}^f_t(\widehat{w}^m_{tk})$ as the predicted log wage of the wife of husband k conditional on his predicted log wage $\widehat{w}^m_{tk}$ from equation (8). We define $\widehat{w}^m_t\left(\widehat{w}^f_{tk}\right)$ analogously for the husbands. Similarly, we define $\widehat{e}^f_t(e^m_{tk})$, using a local linear regression estimator similar to equation (8) as the wife's predicted level of education conditional on her husband's level of education[15] and $\widehat{e}^m_t\left(e^f_{tk}\right)$ analogously.[16]

The OLS regression equations are then

$$\begin{aligned} \widehat{w}^f_t(\widehat{w}^m_{tk}) &= \lambda^f_0 + \lambda^f_1 \widehat{e}^f_t(e^m_{tk}) + \varepsilon^f_{tk}, \;\; \varepsilon^f_{tk} \sim iid\left(0, \sigma^{f2}_{\varepsilon}\right); \\ \widehat{w}^m_t\left(\widehat{w}^f_{tk}\right) &= \lambda^m_0 + \lambda^m_1 \widehat{e}^m_t\left(e^f_{tk}\right) + \varepsilon^m_{tk}, \;\; \varepsilon^m_{tk} \sim iid\left(0, \sigma^{m2}_{\varepsilon}\right). \end{aligned}$$

These allow us to estimate how much of the assortative mating trends are due to sorting over wages and how much is due to sorting over education. Because education is highly correlated with wages as well as between spouses due to the social context of marital sorting, we expect to see a positive relationship between the two.

4.4 Testing for Positive Assortative Mating

Siow (2015) and associated papers (e.g., Graham, 2011; Chiappori, Salanié, and Weiss, 2011) suggest methods based on empirical representations of supermodularity to test whether a matching process exhibits positive assortative matching. Siow (2015) suggests that a "positive correlation test" is a weak test because it is unclear how low the correlation should be before the theory is rejected."

To understand our proposed test, let n_{mf} be the number of observations in a sample where a type-m male, $m = 1, 2, .., M$, is matched with a type-f female, $f = 1, 2, .., F$, and define $N = \sum_{mf} n_{mf}$ as the total number of sampled marriages. Let p_{mf} be the probability that a randomly sampled marriage is of type mf. The inclusion of some unobserved preferences causes the sample proportions to be random.[17] The unrestricted log likelihood function is

$$L = \sum_{m,f} n_{mf} \log p_{mf}.$$

The unrestricted MLE of p is $\widehat{p}^u = \left\{\widehat{p}^u_{mf}\right\}$ with $\widehat{p}^u_{mf} = n_{mf}/N$. Siow (2015) and others argue that the way to test for positive assortative mating is to also estimate $\widehat{p}$ under the restrictions imposed by positive assortative mating and then perform a

[13] We produce results both using fixed bandwidth (0.5) and the bandwidth based on least square cross-validation (the rule of thumb). This process chooses the constant bandwidth which minimizes the conditional weighted mean integrated squared error (Fan and Gijbels, 1996). While the cross-validation approach is used to provide the minimized integrated mean squared errors, the generated curve is sometimes hard to interpret.

[14] Asadi and Zarezadeh (2020) suggest an alternative measure to analyze the relationship between two random variable, X and Y, with marginal distributions F_x and F_y. They measure the correlation between X and $F_x^{-1}(F_y(Y))$.

[15] We use the same kernel function as in equation (8), and the bandwidth be is chosen using the cross-validation method.

[16] For this procedure, we define education as the semi-continuous variable highest grade level, as opposed to the distinct categorical variables used in equations (1) and (4).

[17] Siow (2015) uses iid Extreme Value errors. Papers such as Graham (2011) show that the methodology can be used even without specifying a functional form for the distribution of the errors.

likelihood ratio test. A simpler approach is to construct a "pseudo-Wald test," requiring only the unrestricted estimates (Kang and Stern, 2021). This avoids the cost associated with imposing the positive assortive mating restrictions, and it allows for more general tests. The condition for positive assortative mating is[18]

$$R_{mf} = (\log p_{mf} + \log p_{m+1,f+1}) - (\log p_{m+1,f} + \log p_{m,f+1}) > 0 \quad \forall mf,$$

Note that one could test the null hypothesis that $R_{mf} = 0 \ \forall m, f$ with power against the alternative of (positive or negative) assortative mating as a Wald Test with

$$T = \sum_{m,f} \left(T_{mf}^2 / \widehat{s}_{mf}^2\right) \sim \chi_K^2 \tag{10}$$

where

$$T_{mf} = \left(\log \widehat{p}_{mf}^u + \log \widehat{p}_{m+1,f+1}^u\right) - \left(\log \widehat{p}_{m+1,f}^u + \log \widehat{p}_{m,f+1}^u\right),$$

$$\begin{aligned} s_{mf}^2 &= Var\left[\left(\log \widehat{p}_{mf}^u + \log \widehat{p}_{m+1,f+1}^u\right) - \left(\log \widehat{p}_{m+1,f}^u + \log \widehat{p}_{m,f+1}^u\right)\right] \\ &= \left[\frac{(1-p_{mf})}{np_{mf}} + \frac{(1-p_{m+1,f+1})}{np_{m+1,f+1}} + \frac{(1-p_{m+1,f})}{np_{m+1,f}} + \frac{(1-p_{m,f+1})}{np_{m,f+1}}\right], \end{aligned}$$

and

$$K = (M-1)(F-1) \tag{11}$$

is the number of restrictions (Kang and Stern, 2021). The results of this test statistic are reported in Section 5.5.

If one were interested in testing the null hypothesis of no assortative mating versus the alternative of positive assortative mating (one-sided), then the distribution of T under either the null or alternative would be difficult to evaluate. However, the same is true of Siow (2015) because the restrictions are inequalities instead of equality restrictions (see, for example, Kudo, 1962; Gourieroux, Holly, and Monfort, 1982; Kodde and Palm, 1986). Siow (2015) uses parametric bootstrapping to simulate critical values. The same can be done for T. Under positive assortative mating, $T > 0$, and, with total positive assortative mating (TP2 in Siow, 2015), $T \to \infty$ because $p_{m+1,f}p_{m,f+1} = 0$. Under random sorting, where $\log p_{mf} = \log x_m + \log y_f$, $R_{mf} = 0 \ \forall m, f$. In general, one can simulate a critical value $\widehat{c}$ such that

$$\Pr[T > \widehat{c}_A \mid R_{mf} = 0 \quad \forall m, f] = \alpha.$$

The advantages of the Siow (2015) test over our suggested pseudo-Wald test are a) it provides an estimate of p under the restriction of positive assortative mating (and therefore also allows for simulation using $\widehat{p}$) and b) it allows one to think of the positive assortative mating model as the null hypothesis. Advantage (a) has significant value, but advantage (b) is not so valuable since it is not clear which of the two models should be taken as the null hypothesis model. In our case, since we do not focus on the positive assortative mating parameters, our test seems advantageous, given its ease of use.

We can also construct a likelihood ratio (LR) test. The unrestricted estimate of p_{mf} is $\widehat{p}_{mf}^u = n_{mf}/n$ where n_{mf} is the number of observations in cell mf and $n = \sum_{m,f} n_{mf}$. Under the null hypothesis of random sorting, the restricted estimate of p_{mf} is $\widehat{p}_{mf}^r = \widehat{p}_m^r \widehat{p}_f^r$ where $\widehat{p}_m^r = \sum_f \widehat{p}_{mf}^u$ and $\widehat{p}_f^r = \sum_m \widehat{p}_{mf}^u$. Then, the LR test statistic is

$$\begin{aligned} 2\log \frac{L_u}{L_r} &= 2\left[\sum_{m,f} n_{mf} \log \widehat{p}_{mf}^u - \sum_{m,f} n_{mf} \log \widehat{p}_{mf}^r\right] \\ &= 2\left[\sum_{m,f} n_{mf} \log \frac{\widehat{p}_{mf}^u}{\widehat{p}_{mf}^r}\right] \\ &= 2\left[\sum_{m,f} n_{mf} \log \frac{\widehat{p}_{mf}^u}{\left(\sum_f \widehat{p}_{mf}^u\right)\left(\sum_m \widehat{p}_{mf}^u\right)}\right] \sim \chi_K^2 \end{aligned}$$

[18]Note that $R_{ij} > 0 \ \forall ij \Rightarrow$

$$\log p_{ij} + \log p_{i+k,j+m} > \log p_{i+k,j} + \log p_{i,j+m} \ \forall k, m \geq 1.$$

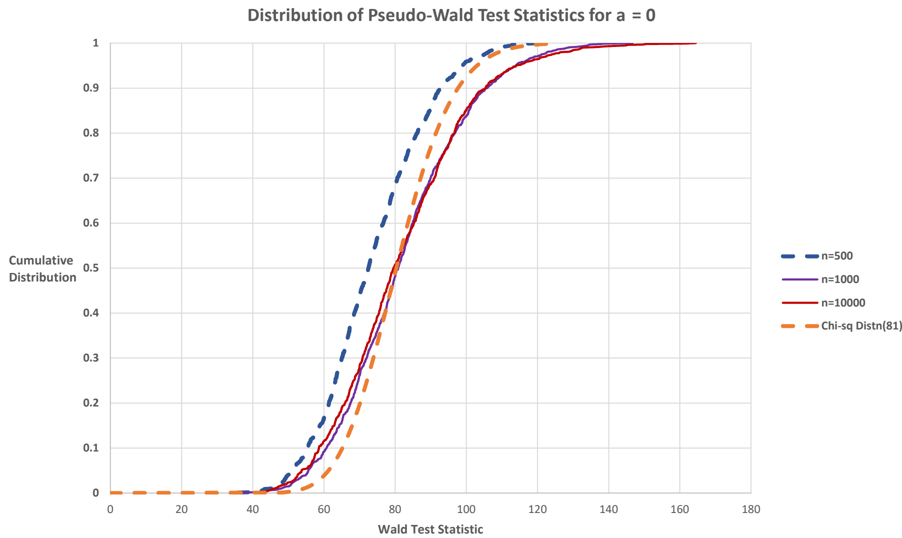


Figure 6: Distribution of Pseudo-Wald Test Statistic for $\alpha = 0$

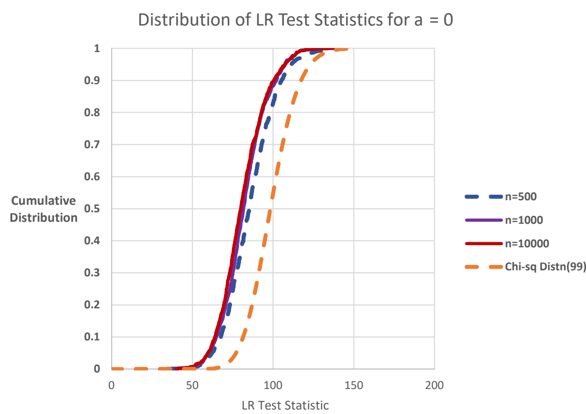


Figure 7: Distribution of LR Test Statistic for $\alpha = 0$

with

$$K = MF - 1$$

which is larger than the degrees of freedom in equation (11) (Kang and Stern, 2021). This difference exists because, for the pseudo-Wald test, one loses degrees of freedom at $m = M$ and $f = F$ which does not happen for the LR test. The results of these test statistics are reported in Section 5.5.

The application in Siow (2015) is about education which has a natural discreteness to it, while our application is about log wage which is continuous. Dagsvick (1994) and Dupuy and Galichon (2014) suggest how to amend the Choo and Siow (2006a, 2006b) model for continuous variables. To operationalize the test statistic in equation (10), we divide the $\mathbb{R} \times \mathbb{R}$ space of log wages into cells where the boundaries of each cell are the deciles for men and women in each of our samples. By construction, under the null hypothesis of random sorting, this implies that $\widehat{p}^{r}_{mf} = 1/100 \; \forall m, f$. We also report results of the test statistics for education using 5 levels of education for men and women.

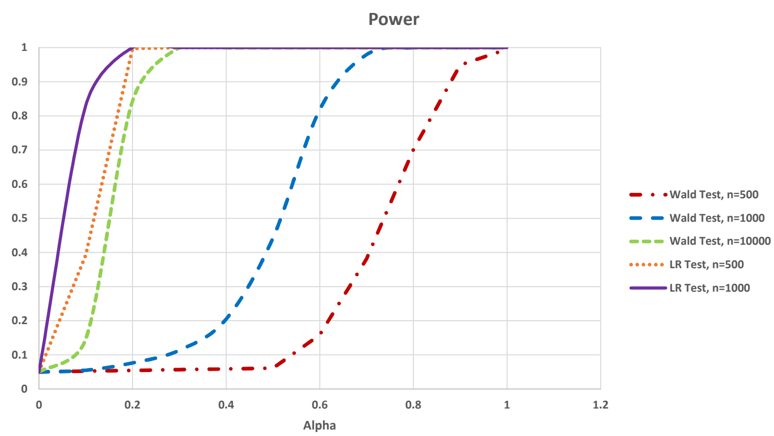


Figure 8: Power

Table 3: Female Participation Estimates

Variable	1962	1967	1970	1972	1975	1978	1980	1985	1990
Constant	-2.772 ***	-1.290 ***	-0.620 **	-0.487 **	-0.479 **	-0.344 *	-0.699 ***	-0.695 ***	-1.399 ***
White	-0.161	-0.354 *	-0.011	-0.138	-0.111	-0.164 *	-0.128 *	-0.036	0.086
Black	-0.202	-0.356 *	0.246	0.037	0.050	-0.008	0.007	0.154 *	0.169 *
Hispanic				-0.121 *	-0.002	-0.070	-0.009	-0.109 **	-0.117 ***
Age/100	7.199 ***	2.810 ***	0.919	0.750	0.930	1.818 *	3.537 ***	3.976 ***	7.966 ***
$(\text{Age}/100)^2$	-8.235 ***	-3.124 **	-2.587 **	-2.566 **	-3.150 ***	-4.883 ***	-7.150 ***	-7.903 ***	-12.761 ***
HS Diploma	0.208 ***	0.331 ***	0.279 ***	0.304 ***	0.302 ***	0.362 ***	0.467 ***	0.444 ***	0.485 ***
Some College	0.221 ***	0.268 ***	0.294 ***	0.284 ***	0.416 ***	0.422 ***	0.597 ***	0.645 ***	0.647 ***
College Degree	0.134	-0.014	0.060	0.153 **	0.381 ***	0.437 ***	0.489 ***	0.660 ***	0.678 ***
Post Grad Degree	0.562 ***	0.216	0.386 ***	0.312 **	0.709 ***	0.671 ***	0.558 ***	0.750 ***	0.757 ***
# Children < 5	-		-0.598 ***	-0.631 ***	-0.615 ***	-0.601 ***	-0.567 ***	-0.503 ***	-0.471 ***
Northeast	0.170 ***	0.100 *	0.095 **	0.128 ***	0.061	0.049	0.061 *	0.079 *	0.116 ***
South	0.078	0.047	0.039	0.148 ***	0.057	0.033	0.001	0.046	0.072 *
Midwest	0.040	0.053	0.062	0.122 ***	0.056	0.046	0.098 ***	0.106 ***	0.137 ***
Metro	0.187 ***	0.051	0.035	0.095 ***	0.083 **	0.059 *	0.080 ***	0.095 ***	0.064 *
# Obs	9,961	10,175	15,246	14,792	13,274	14,877	17,013	15,886	14,677

Notes:
1) Single-starred items are statistically significant at the 5% level, double-starred items at the 1% level, and triple-starred at the 0.1% level.
2) Standard errors are available from the corresponding author.

5 Results

In this section, we provide results using the methods described above. First, we show results from our use of the Heckman 2-stage procedure. We do this separately for men and women as we need to predict log wages for both and gender has a large effect on how other variables affect labor force participation and wages. Next, we display, mostly graphically how assortative mating works, both for wages and for education, and how it has changed from 1962 up to the present. In particular, we look at the characteristics of men who marry women with specified log wages and education, and then we look at the characteristics of women who marry men with specified log wages and education. We also perform a cohort analysis to better measure the relationship between men and women at the time they are likely to marry. We finish with the results of our proposed test statistics for positive assortative mating with respect to log wages and education.

5.1 Labor Force Participation Estimates

Tables 3 and 4 show estimates for the male and female participation equations defined in equations (1) and (2). The results in Tables 3 and 4 are largely as we would expect and are fairly consistent with probit labor force participation estimates, for example, for 1986 from Hersch (1991) and for the 1970's and 1990's from Mulligan and Rubinstein (2008). Labor force participation increases with education level and with age up to around 50, depending on the year, after which it decreases. An exception to the general results for education is that, for years 1970 and 1972, men with post-graduate degrees have lower labor force participation rates relative to other education levels.

Table 3: Female Participation Estimates (continued)

Variable	1995	2000	2005	2007	2010	2014	2015	2018	2019
Constant	-1.610 ***	-1.963 ***	-1.629 ***	-1.895 ***	-1.755 ***	-2.039 ***	-2.524 ***	-1.697 ***	-1.561 ***
White	0.120 **	0.171 **	0.087 **	0.084 *	0.093 **	0.128 ***	0.103 **	0.203 ***	0.195 ***
Black	0.196 **	0.298 ***	0.139 **	0.194 ***	0.127 **	0.175 ***	0.123 *	0.262 ***	0.256 ***
Hispanic	-0.176 ***	-0.124 ***	-0.138 ***	-0.158 ***	-0.171 ***	-0.130 ***	-0.118 ***	-0.166 ***	-0.146 ***
Age/100	8.656 ***	9.911 ***	8.636 ***	9.868 ***	8.071 ***	8.707 ***	10.618 ***	7.411 ***	6.734 ***
$(Age/100)^2$	-13.158 ***	-14.137 ***	-12.291 ***	-13.465 ***	-10.995 ***	-11.301 ***	-13.319 ***	-10.108 ***	-9.194 ***
HS Diploma	0.454 ***	0.478 ***	0.470 ***	0.442 ***	0.508 ***	0.504 ***	0.526 ***	0.415 ***	0.451 ***
Some College	0.656 ***	0.658 ***	0.653 ***	0.629 ***	0.689 ***	0.698 ***	0.763 ***	0.622 ***	0.656 ***
College Degree	0.656 ***	0.696 ***	0.649 ***	0.589 ***	0.698 ***	0.809 ***	0.882 ***	0.759 ***	0.793 ***
Post Grad Degree	0.710 ***	0.743 ***	0.664 ***	0.747 ***	0.765 ***	0.912 ***	0.981 ***	0.951 ***	0.976 ***
# Children < 5	-0.423 ***	-0.411 ***	-0.382 ***	-0.356 ***	-0.331 ***	-0.287 ***	-0.279 ***	-0.367 ***	-0.327 ***
Northeast	0.083 *	0.099 **	0.152 ***	0.117 ***	0.139 ***	0.164 ***	0.125 ***	0.146 ***	0.124 ***
South	0.036	0.064 *	0.017	0.018	-0.023	-0.004	-0.031	-0.026	-0.048
Midwest	0.241 ***	0.209 ***	0.211 ***	0.202 ***	0.236 ***	0.246 ***	0.279 ***	0.243 ***	0.271 ***
Metro	0.081 **	0.045	-0.029	-0.004	0.027	-0.027	0.035	0.055	0.037
# Obs	13,561	12,271	19,222	18,664	18,049	17,050	16,870	15,129	15,291

Notes:

1) Single-starred items are statistically significant at the 5% level, double-starred items at the 1% level, and triple-starred at the 0.1% level.

2) Standard errors are available from the corresponding author.

Table 4: Male Participation Estimates

Variable	1962	1967	1970	1972	1975	1978	1980	1985	1990
Constant	-1.902 ***	-1.653 ***	-2.079 ***	-2.296 ***	-2.670 ***	-2.421 ***	-2.109 ***	-2.010 ***	-2.094 ***
White	0.277	0.236	0.313	0.286 *	0.263 *	0.328 ***	0.324 ***	0.347 ***	0.232 ***
Black	-0.242	-0.101	-0.013	-0.097	-0.156	0.013	0.016	-0.012	-0.117
Hispanic	-	-	-	-0.112	-0.127 *	-0.082	-0.048	-0.103 *	0.048
Age/100	12.902 ***	14.093 ***	15.129 ***	15.809 ***	16.024 ***	16.596 ***	14.912 ***	14.004 ***	15.456 ***
$(Age/100)^2$	-16.786 ***	-19.288 ***	-19.995 ***	-20.505 ***	-20.405 ***	-22.305 ***	-20.292 ***	-19.701 ***	-21.867 ***
HS Diploma	0.449 ***	0.409 ***	0.389 ***	0.389 ***	0.418 ***	0.404 ***	0.397 ***	0.325 ***	0.409 ***
Some College	0.264 ***	0.357 ***	0.265 ***	0.336 ***	0.456 ***	0.531 ***	0.460 ***	0.460 ***	0.448 ***
College Degree	0.551 ***	0.521 ***	0.410 ***	0.549 ***	0.786 ***	0.727 ***	0.807 ***	0.777 ***	0.685 ***
Post Grad Degree	0.385 **	0.041	0.187 *	0.332 ***	0.662 ***	0.503 ***	0.699 ***	0.721 ***	0.693 ***
# Children < 5	-	-	0.011	0.061 *	0.060 *	0.016	-0.072 **	-0.039	-0.024
Northeast	0.120 *	0.281 ***	0.318 ***	0.127 **	0.158 ***	0.063	0.152 **	0.225 ***	0.124 ***
South	-0.035	0.120 *	0.197 ***	0.205 ***	0.138 ***	0.135 *	0.088	0.211 ***	0.116 ***
Midwest	0.201 ***	0.384 ***	0.381 ***	0.212 ***	0.175 ***	0.186 ***	0.130 *	0.075	0.103 *
Metro	0.318 ***	0.298 ***	0.259 ***	0.273 ***	0.205 ***	0.167 ***	0.144 ***	0.175 ***	0.171 ***
# Obs	7,809	10,258	15,456	15,042	13,536	15,102	17,232	16,088	14,825

Notes:

1) Single-starred items are statistically significant at the 5% level, double-starred items at the 1% level, and triple-starred at the 0.1% level.

2) Standard errors are available from the corresponding author.

Table 4: Male Participation Estimates (continued)

Variable	1995	2000	2005	2007	2010	2014	2015	2018	2019
Constant	-1.849 ***	-1.692 ***	-2.108 ***	-2.356 ***	-2.351 ***	-2.455 ***	-2.553 ***	-2.025 ***	-2.064 ***
White	0.305 ***	0.291 ***	0.122 **	0.183 ***	0.140 ***	0.195 ***	0.150 ***	0.139 ***	0.111 **
Black	0.056	0.022	-0.278 ***	-0.168 **	-0.226 ***	-0.098	-0.146 **	-0.170 **	-0.208 ***
Hispanic	0.016	0.102 *	0.151 ***	0.160 ***	0.083 **	0.152 ***	0.212 ***	0.219 ***	0.203 ***
Age/100	13.416 ***	13.516 ***	15.182 ***	16.134 ***	14.401 ***	15.175 ***	16.180 ***	14.330 ***	14.840 ***
$(Age/100)^2$	-19.651 ***	-19.898 ***	-20.794 ***	-22.188 ***	-18.979 ***	-19.716 ***	-20.780 ***	-18.950 ***	-19.613 ***
HS Diploma	0.420 ***	0.366 ***	0.332 ***	0.412 ***	0.363 ***	0.313 ***	0.225 ***	0.230 ***	0.254 ***
Some College	0.540 ***	0.444 ***	0.434 ***	0.545 ***	0.512 ***	0.378 ***	0.324 ***	0.349 ***	0.334 ***
College Degree	0.728 ***	0.788 ***	0.679 ***	0.728 ***	0.852 ***	0.722 ***	0.612 ***	0.572 ***	0.526 ***
Post Grad Degree	0.850 ***	0.664 ***	0.798 ***	0.926 ***	1.074 ***	0.830 ***	0.803 ***	0.670 ***	0.680 ***
# Children < 5	-0.037	0.063 *	0.086 ***	0.067 **	0.070 **	0.071 **	0.113 ***	0.079 **	0.133 ***
Northeast	0.030	0.059	0.011	0.044	0.071 *	0.102 **	0.034	0.037	0.028
South	0.096 *	0.001	0.048	0.037	0.049	0.056	0.043	0.049	0.030
Midwest	0.144 ***	0.110 *	0.081 *	0.069 *	0.052	0.133 ***	0.161 ***	0.115 **	0.129 **
Metro	0.202 ***	0.203 ***	0.186 ***	0.206 ***	0.072 *	0.105 ***	0.115 ***	0.146 ***	0.162 ***
# Obs	13,655	12,415	19,336	18,797	18,087	17,106	16,948	15,213	15,391

Notes:

1) Single-starred items are statistically significant at the 5% level, double-starred items at the 1% level, and triple-starred at the 0.1% level.

2) Standard errors are available from the corresponding author.

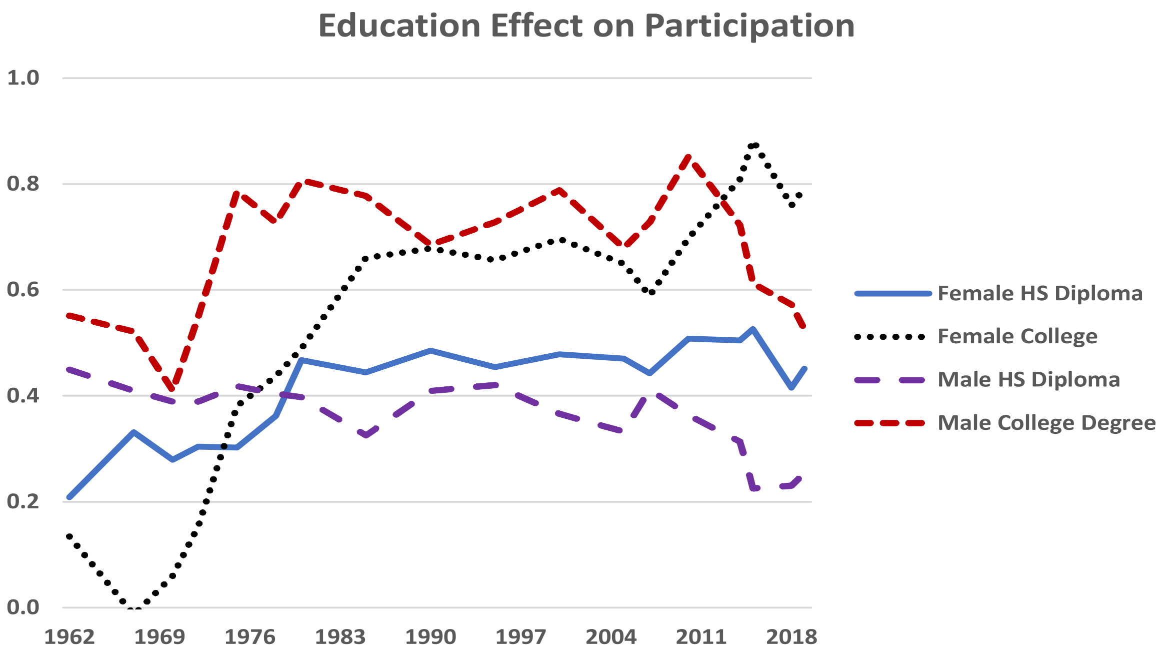


Figure 9: Education Effect on Participation

For women, whites tend to have a lower propensity to work than blacks; while, for men, whites and other races as well as Hispanics have a higher propensity to work than blacks. This could be associated with the fact that the spouses of white women earn more than the spouses of black women on average, which is likely to increase the reservation wage of white women. The presence of children under age five greatly decreases a woman's propensity to work, but it has no effect on men. Living in a metro area increases propensity to work for men and women except for women in the most recent years in which it decreases propensity to work. With respect to changes in estimates over time, we focus on the education estimates and *has child under 5*, and we focus on estimates in 1972 and after. Prior to 1972, *has child under 5* was not available, causing there to be a large break in estimates in that year. For *has child under 5*, while the effect for women is large, over time, it has declined (in absolute value) at a surprisingly steady annual rate.[19]

The changes in the estimated effect of education on labor force participation, reported in Figure 9, are quite interesting. Women's labor force participation is more responsive to education than men's, and women's is growing at a faster rate than men's. Figure 10 shows how the predicted variation in labor force participation for men can be decomposed into variation in parameter estimates across years and variation in covariates across and within years. Table 2 shows the variation across years. The only two covariates with significant variation are *Hispanic* and *metropolitan*. However, there is significant variation in all of the covariates within each year that causes variation in participation rates within each year. Meanwhile, Tables 3 and 4 and Figure 9 show the variation in parameter estimates across years. Besides variation in education estimates across years, there is significant variation in estimates especially for racial effects, age effects, and metro effects. The first two curves from the bottom are for men who did not complete high school, the next two are for men with a high school degree, and the last two are for men with a college degree. The solid lines are expected participation rates using 1990 sample means but allowing parameter estimates to vary across years $E\left(y^*_{tij} \mid \overline{X}_{1990,j}\beta_t\right)$ from equation (1) conditional on gender and education, and the dashed lines use both variation in estimates across years and variation in covariates across and within years $E\left(y^*_{tij} \mid X_{tij}\beta_t\right)$. The variation in the parameter estimates alone explains almost all of the variation in participation rates across years for men with at least a high school diploma. For men who did not finish high school, the variation in covariates other than just education explains a significant amount of the variation in participation prior to 2005. For women, the variation in parameter estimates alone (using 1990 sample means) explains almost all of the variation in labor force participation rates.[20]

[19] The slope of the OLS regression line is 0.0116, and $R^2 = 0.989$.

[20] The graph for women is available from the corresponding author.

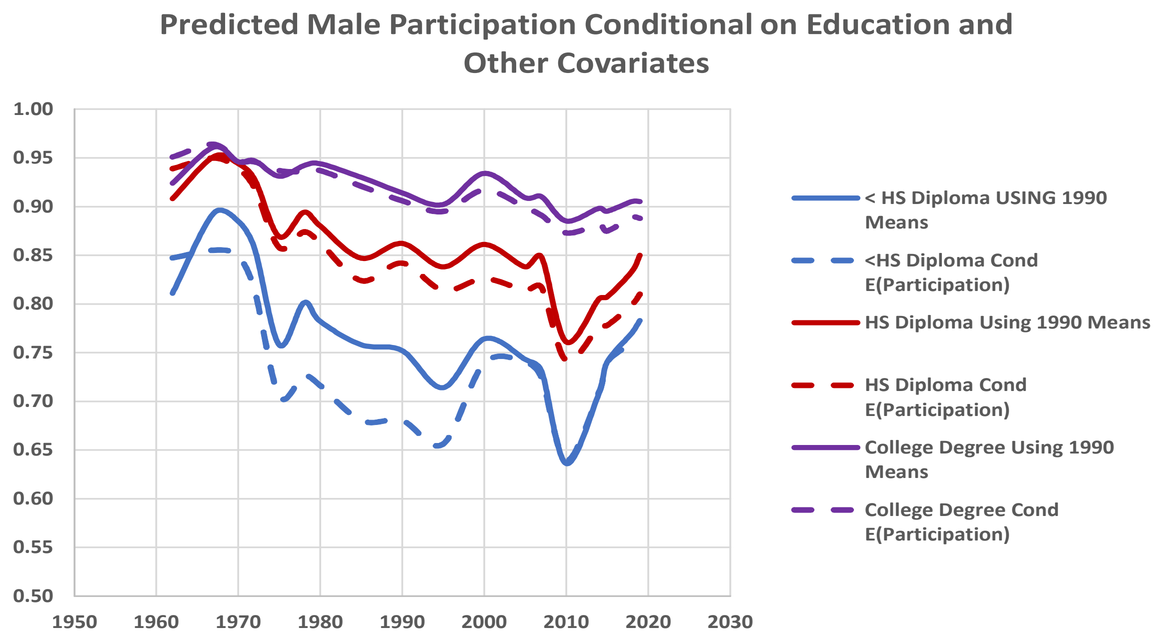


Figure 10: Predicted Male Participation Conditional on Eduction and Other Covariates

5.2 Log Wage Estimates

Log wage estimates are reported in Table 5 (for women) and Table 6 (for men). As is the case for the labor force participation estimates in Tables 3 and 4, the log wage results are largely as we would expect and are fairly consistent with, for example, Hersch (1991) and Mulligan and Rubinstein (2008). Log wages increase at a decreasing rate with age, maximized at around 50 to 60 years. As seen in Figure 11, log wages increase with education level with higher returns to education in more recent years. Effects for women are slightly higher than for men, possibly helping to explain the large increase in the proportion of women getting more education (e.g., Goldin, 1990; Blau and Kahn, 2007; Bailey, Hershbein, and Miller, 2012). Ashworth and Ransom (2019) and Ashworth et al. (2020) find evidence of flattening of education returns in the last 20 years, and our results are somewhat consistent with their findings. The year 1962 for women is an exception for all of these results, which may be due to the vastly different role education and work played in the lives of women then (e.g., Goldin, 1990).

The imprecise results in 1962-1972 are likely the consequence of the lack of information about household children in those years. Without the variable about existing children under 5, when distinguishing the selection equation from the wage equation, there is no exclusion restriction. As we have the ability to observe the dependent variable in the selection equation, this does not interfere with the identification of the inverse Mills ratio term. However, the lack of exclusion restriction increases the degree of multicollinearity between the inverse Mills ratio and the other covariates. This multicollinearity increases the standard errors on the estimated coefficients, decreasing the precision of the estimates. In particular, the results of the Heckman two-step procedure in 1962 and 1967 suffer from severe multicollinearity. The problem is fixed from 1970 onward with *has child under 5* becoming available.[21]

The effect of race on log wages is much stronger for men than for women. White men earn more than black men who earn more than Hispanic men. For women, the results are similar but weaker, and the race terms are insignificant over several years. Living in a metro area increases log wages for both genders, especially for the years after 1972.

For women, the inverse Mills ratio τ coefficient estimates are significant essentially in all years while the τ coefficient estimates are generally insignificant for men. Selection bias is much stronger for women; i.e. more women self-select out of the workforce because their wages would be lower than their unobserved reservation wage. Our results are in contrast to the 1980s results in Smith and Ward (1989) but are broadly similar to the results of Nakosteen and Zimmer (2001) and Zhang

[21] Returning to the potential use of semiparametric estimation methods, for example, Escanciano, Jacho-Chavez, and Lewbel (2016) provide a proof and Monte Carlo evidence that exclusion restrictions are not necessary for identification in semiparametric two-step estimators. However, it is not clear how relaxing a functional form assumption in a model that is behaving poorly without an exclusion restriction will perform better than with the functional form assumption.

Table 5: Female log Hourly Wage Estimates

Variable	1962	1967	1970	1972	1975	1978	1980	1985	1990
Constant	2.019 ***	2.570 ***	2.272 ***	2.200 ***	2.321 ***	2.099 ***	2.225 ***	1.758 ***	1.411 ***
White	0.182 *	0.179 **	-0.001	0.094	-0.101 *	0.026	-0.014	0.069 *	0.029
Black	0.074	0.114	-0.011	0.007	-0.146 *	0.027	-0.010	0.044	-0.017
Hispanic	---	---	---	-0.106 **	-0.069	-0.087 ***	-0.070 ***	-0.099 ***	-0.096 ***
Age/100	2.967 ***	0.063	1.390 ***	1.738 ***	2.219 ***	2.825 ***	1.639 ***	3.111 ***	4.207 ***
$(Age/100)^2$	-3.494 ***	-0.194	-1.702 ***	-2.154 ***	-2.783 ***	-3.528 ***	-2.118 ***	-3.705 ***	-5.030 ***
HS Diploma	0.039	0.072 ***	0.123 ***	0.096 ***	0.078 ***	0.114 ***	0.134 ***	0.181 ***	0.231 ***
Some College	0.086 *	0.211 ***	0.230 ***	0.202 ***	0.178 ***	0.186 ***	0.249 ***	0.345 ***	0.446 ***
College Degree	0.121 *	0.244 ***	0.278 ***	0.271 ***	0.261 ***	0.281 ***	0.400 ***	0.472 ***	0.623 ***
Post Grad Degree	0.675 ***	0.299 **	0.534 ***	0.369 ***	0.417 ***	0.371 ***	0.510 ***	0.693 ***	0.767 ***
Northeast	-0.127 **	0.004	-0.011	-0.016	-0.015	-0.010	-0.020	-0.046 **	0.073 ***
South	-0.166 ***	-0.084 **	-0.085 ***	-0.096 ***	-0.067 **	-0.077 ***	-0.074 ***	-0.067 ***	-0.055 **
Midwest	-0.146 ***	-0.045	-0.028	-0.031	-0.012	-0.021	0.006	-0.074 ***	-0.039 *
Metro	0.045	0.078 ***	0.082 ***	0.125 ***	0.079 ***	0.062 ***	0.053 ***	0.071 ***	0.132 ***
Inverse Mills Ratio			0.168 ***	0.073 *	0.121 ***	0.118 ***	0.212 ***	0.163 ***	0.202 ***
Residual SE			0.000	0.000	0.000	0.000	0.000	0.000	0.000
# Obs	1,259	2,213	15,246	14,792	13,274	14,877	17,013	15,886	14,677

Notes:

1) Single-starred items are statistically significant at the 5% level, double-starred items at the 1% level, and triple-starred at the 0.1% level.

2) Standard errors are available from the corresponding author.

Table 5: Female log Hourly Wage Estimates (continued)

Variable	1995	2000	2005	2007	2010	2014	2015	2018	2019
Constant	1.127 ***	1.152 ***	1.068 ***	0.946 ***	1.069 ***	0.548 **	0.826 ***	1.401 ***	1.298 ***
White	0.065 *	0.096 **	0.050 *	0.054 *	0.027	0.076 ***	0.044 *	0.034	0.058 **
Black	0.032	0.065	0.001	-0.002	-0.054	0.031	-0.002	-0.041	-0.040
Hispanic	-0.076 ***	-0.160 ***	-0.146 ***	-0.182 ***	-0.158 ***	-0.136 ***	-0.159 ***	-0.187 ***	-0.181 ***
Age/100	5.211 ***	5.050 ***	5.588 ***	5.978 ***	5.362 ***	6.334 ***	5.784 ***	4.058 ***	4.108 ***
$(Age/100)^2$	-6.245 ***	-5.915 ***	-6.482 ***	-6.895 ***	-5.888 ***	-6.806 ***	-6.341 ***	-4.315 ***	-4.339 ***
HS Diploma	0.263 ***	0.257 ***	0.265 ***	0.272 ***	0.311 ***	0.330 ***	0.289 ***	0.252 ***	0.236 ***
Some College	0.423 ***	0.444 ***	0.467 ***	0.493 ***	0.515 ***	0.552 ***	0.505 ***	0.452 ***	0.444 ***
College Degree	0.705 ***	0.702 ***	0.765 ***	0.767 ***	0.772 ***	0.880 ***	0.846 ***	0.790 ***	0.786 ***
Post Grad Degree	0.894 ***	1.011 ***	1.050 ***	1.065 ***	1.132 ***	1.161 ***	1.141 ***	1.151 ***	1.094 ***
Northeast	0.032	0.063 **	0.044 *	0.045 *	0.021	0.038	0.013	0.011	0.019
South	-0.065 ***	-0.010	-0.061 ***	-0.036 *	-0.073 ***	-0.070 ***	-0.068 ***	-0.103 ***	-0.084 ***
Midwest	0.005	0.017	-0.053 **	-0.054 **	-0.042 *	-0.002	0.009	-0.050 *	-0.008
Metro	0.145 ***	0.136 ***	0.132 ***	0.163 ***	0.121 ***	0.138 ***	0.152 ***	0.091 ***	0.160 ***
Inverse Mills Ratio	0.255 ***	0.254 ***	0.293 ***	0.314 ***	0.347 ***	0.495 ***	0.407 ***	0.308 ***	0.323 ***
Residual SE	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
# Obs	13,561	12,271	19,222	18,664	18,049	17,050	16,870	15,129	15,291

Notes:

1) Single-starred items are statistically significant at the 5% level, double-starred items at the 1% level, and triple-starred at the 0.1% level.

2) Standard errors are available from the corresponding author.

Table 6: Male log Hourly Wage Estimates

Variable	1962	1967	1970	1972	1975	1978	1980	1985	1990
Constant	1.413 ***	1.955 ***	1.844 ***	1.481 ***	2.017 ***	2.138 ***	2.004 ***	1.858 ***	0.995 ***
White	0.212 **	0.219 ***	0.259 ***	0.098 *	0.083 *	0.089 **	0.124 ***	0.061 *	0.145 ***
Black	-0.049	0.057	0.058	-0.122 **	-0.052	-0.052	0.006	-0.148 ***	-0.045
Hispanic	---	---	---	-0.196 ***	-0.186 ***	-0.145 ***	-0.157 ***	-0.147 ***	-0.198 ***
Age/100	5.688 ***	3.879 ***	4.572 ***	6.912 ***	4.846 ***	4.314 ***	4.736 ***	4.751 ***	7.469 ***
$(Age/100)^2$	-6.125 ***	-3.933 ***	-4.983 ***	-7.740 ***	-4.894 ***	-4.007 ***	-4.579 ***	-4.075 ***	-7.923 ***
HS Diploma	0.148 ***	0.132 ***	0.142 ***	0.149 ***	0.119 ***	0.127 ***	0.119 ***	0.159 ***	0.217 ***
Some College	0.218 ***	0.245 ***	0.238 ***	0.249 ***	0.195 ***	0.172 ***	0.160 ***	0.214 ***	0.327 ***
College Degree	0.446 ***	0.416 ***	0.475 ***	0.464 ***	0.307 ***	0.337 ***	0.317 ***	0.408 ***	0.554 ***
Post Grad Degree	0.464 ***	0.430 ***	0.448 ***	0.471 ***	0.349 ***	0.448 ***	0.409 ***	0.521 ***	0.658 ***
Northeast	-0.024	-0.049 **	-0.053 ***	-0.022	-0.023	-0.079 ***	-0.088 ***	-0.068 ***	0.057 ***
South	-0.125 ***	-0.121 ***	-0.139 ***	-0.111 ***	-0.079 ***	-0.106 ***	-0.094 ***	-0.090 ***	-0.073 ***
Midwest	0.004	-0.018	0.006	0.010	-0.006	-0.043 ***	-0.018	-0.058 ***	-0.007 ***
Metro	0.138 ***	0.137 ***	0.150 ***	0.161 ***	0.109 ***	0.065 ***	0.064 ***	0.058 ***	0.162 ***
Inverse Mills Ratic	0.269	-0.012	0.174	0.282 *	-0.055	-0.199 *	-0.097	-0.192	0.200
Residual SE	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
# Obs	7,809	10,258	15,456	15,042	13,536	15,102	17,232	16,088	14,825

Notes:

1) Single-starred items are statistically significant at the 5% level, double-starred items at the 1% level, and triple-starred at the 0.1% level.

2) Standard errors are available from the corresponding author.

Table 6: Male log Hourly Wage Estimates (continued)

Variable	1995	2000	2005	2007	2010	2014	2015	2018	2019
Constant	0.583 **	1.500 ***	1.095 ***	1.073 ***	1.280 ***	1.431 ***	1.631 ***	2.146 ***	2.150 ***
White	0.108 ***	0.136 ***	0.144 ***	0.145 ***	0.139 ***	0.077 ***	0.095 ***	0.058 **	0.033
Black	-0.073 *	-0.017	-0.081 **	-0.087 **	-0.068 *	-0.115 ***	-0.096 ***	-0.122 ***	-0.154 ***
Hispanic	-0.176 ***	-0.219 ***	-0.211 ***	-0.192 ***	-0.169 ***	-0.159 ***	-0.189 ***	-0.218 ***	-0.195 ***
Age/100	9.119 ***	5.178 ***	6.960 ***	7.131 ***	6.273 ***	5.607 ***	4.700 ***	3.066 **	2.866 **
$(\text{Age}/100)^2$	-10.108 ***	-5.265 ***	-7.497 ***	-7.786 ***	-6.774 ***	-5.537 ***	-4.565 ***	-2.381	-2.126
HS Diploma	0.236 ***	0.175 ***	0.181 ***	0.221 ***	0.205 ***	0.182 ***	0.197 ***	0.171 ***	0.175 ***
Some College	0.355 ***	0.320 ***	0.340 ***	0.352 ***	0.359 ***	0.310 ***	0.320 ***	0.301 ***	0.322 ***
College Degree	0.651 ***	0.611 ***	0.657 ***	0.701 ***	0.710 ***	0.630 ***	0.660 ***	0.600 ***	0.633 ***
Post Grad Degree	0.827 ***	0.871 ***	0.944 ***	1.017 ***	0.999 ***	0.950 ***	0.923 ***	0.868 ***	0.898 ***
Northeast	0.048 **	0.008	0.033 *	0.016	0.033 *	-0.002	-0.001	-0.021	-0.010
South	-0.041 **	-0.013	-0.020	-0.055 ***	-0.050 ***	-0.025	-0.025	-0.073 ***	-0.051 **
Midwest	-0.003	0.008	-0.024	-0.050 ***	-0.106 ***	-0.042 *	-0.075 ***	-0.070 ***	-0.062 ***
Metro	0.158 ***	0.146 ***	0.176 ***	0.172 ***	0.140 ***	0.106 ***	0.106 ***	0.076 ***	0.107 ***
Inverse Mills Ratio	0.364 **	0.070	0.179	0.207	0.199	0.042	0.039	-0.298	-0.300 *
Residual SE	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
# Obs	13,655	12,415	19,336	18,797	18,087	17,106	16,948	15,213	15,391

Notes:

1) Single-starred items are statistically significant at the 5% level, double-starred items at the 1% level, and triple-starred at the 0.1% level.

2) Standard errors are available from the corresponding author.

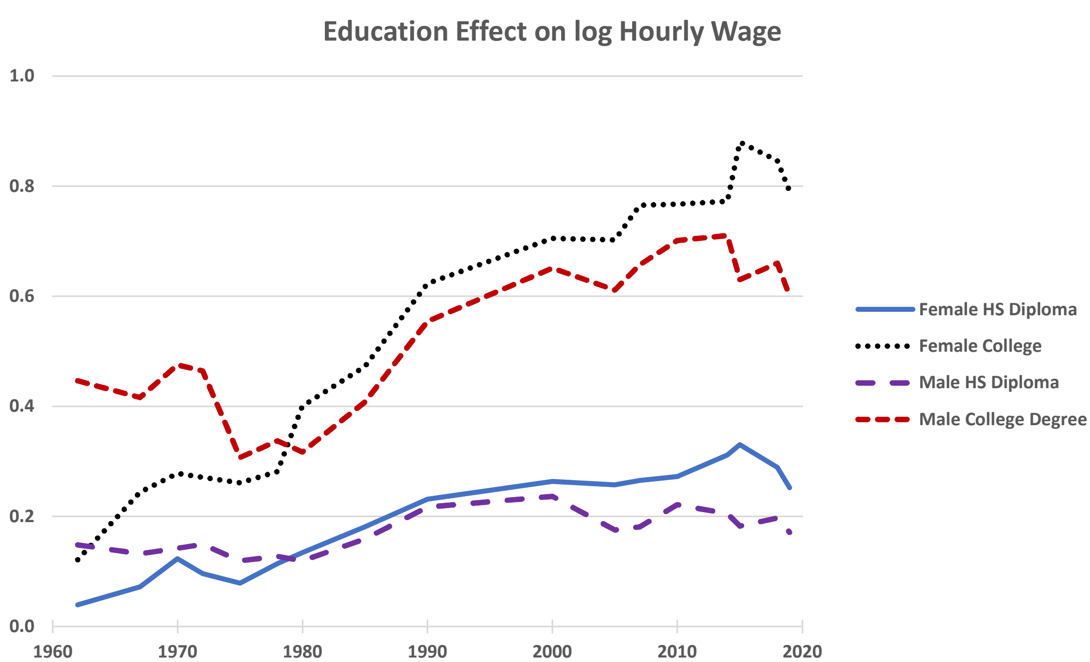


Figure 11: Education Effect on log Hourly Wage

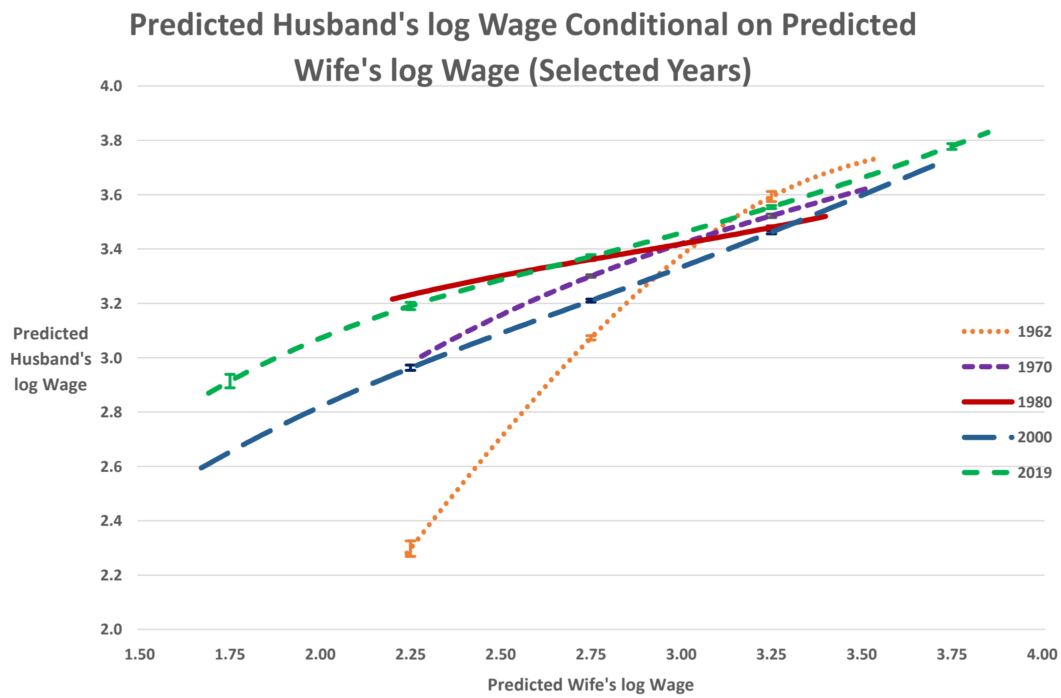


Figure 12: Predicted Husband's log Wage Conditional on Predicted Wife's log Wage (Selected Years)

and Lui (2003), which also used a Heckman selection model. They found no evidence for selection bias among men and evidence for positive selection bias for women for the years 1983-1991. While Nakosteen and Zimmer (2001) then went on to use an OLS regression for the men in their sample due to these results, we continue to use the estimates generated by the Heckman two-step method to maintain consistency across years and gender.[22]

5.3 Local Linear Regression Estimation Results

5.3.1 log Wage Relationships

Using the wages predicted by equation (5) and the estimates from Tables 5 and 6, we next use equation (8) to construct $w^s(w^o)$ curves in equation (7). The results are displayed in Figures 12 and 13 with 95% percentile intervals (see Cattaneo and Jansson, 2018) at discrete points.[23] One can see that all of the curves are measured with high precision. Figure 12 shows that, even in 1962, there is a positive relationship between spouses' wages. The relationship becomes a little weaker in more recent years, which we see in the change in the general slope of each successive year. The positive relationship is constant across all levels of income. Except for 1962, husbands' wages are generally higher than their wives', which we also see in Figure 13.

One might wonder why the curves in Figures 12 and 13 are not transposed mirror images of each other. To gain some intuition, consider a simple example where

$$\begin{pmatrix} w^f \\ w^m \end{pmatrix} \sim N\left[\begin{pmatrix} \mu^f \\ \mu^m \end{pmatrix}, \begin{pmatrix} \sigma^{ff} & \sigma^{fm} \\ \sigma^{fm} & \sigma^{mm} \end{pmatrix}\right].$$

[22] Zhang and Lui (2003) use 2nd-stage log wages but only for the women who did not work. The wage equation they estimate is for men, so it is also not clear how substituting the selection-corrected woman's log wage has the properties described in Heckman (1979).

[23] The standard errors used in constructing confidence intervals ignore the fact that the predicted log wages are themselves random variables. However, the standard errors associated with these estimates are quite small making the relevant adjustment negligible.

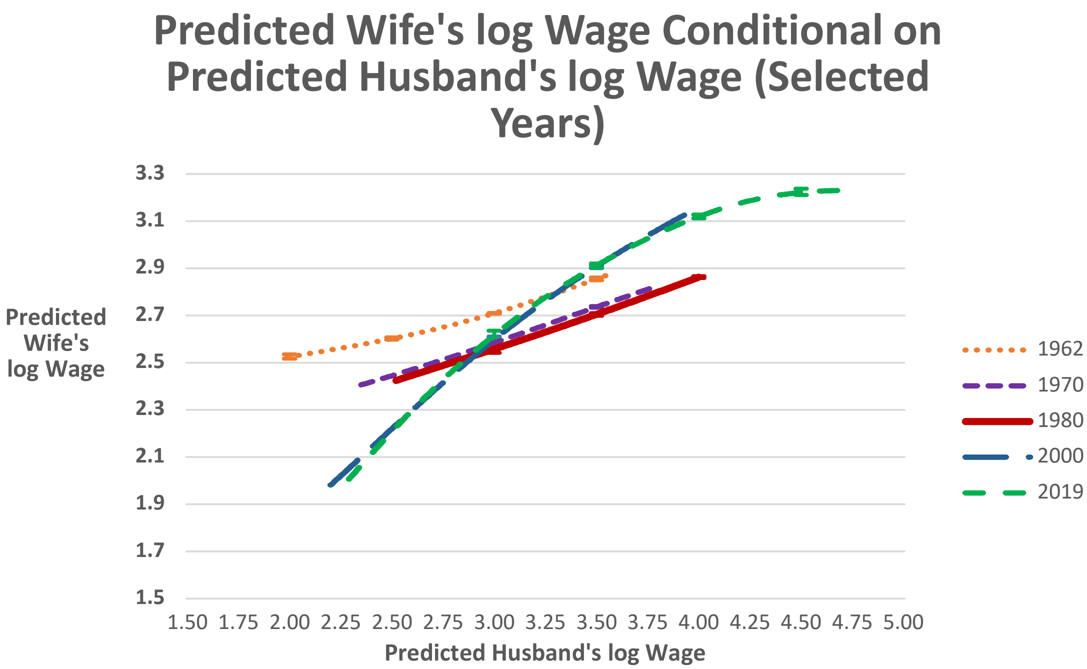


Figure 13: Predicted Wife's log Wage Conditional on Predicted Husband's log Wage (Selected Years)

Then

$$\begin{aligned} Ew^m \quad &| \quad w^f = \mu^m + \frac{\sigma^{fm}}{\sigma^{ff}}\left(w^f - \mu^f\right); \\ Ew^f \quad &| \quad w^m = \mu^f + \frac{\sigma^{fm}}{\sigma^{mm}}\left(w^m - \mu^m\right), \end{aligned} \tag{12}$$

and

$$E\left(w^f \mid \left(Ew^m \mid w^f\right)\right) = \frac{\sigma^{fm}}{\sigma^{mm}}\frac{\sigma^{fm}}{\sigma^{ff}}w^f + \left(1 - \frac{\sigma^{fm}}{\sigma^{mm}}\frac{\sigma^{fm}}{\sigma^{ff}}\right)\mu^f$$

is a weighted average of w^f and μ^f. This is shown clearly in Figure 14 with

$$\begin{pmatrix} w^f \\ w^m \end{pmatrix} \sim N\left[\begin{pmatrix} 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 1 & 2\rho \\ 2\rho & 2 \end{pmatrix}\right]$$

with $\rho = 0.4$. The horizontal axis has w^f and $E\left(w^f \mid w^m\right)$, and the vertical axis has w^m and $E\left(w^m \mid w^f\right)$. The blue dashed curve is $E\left(w^m \mid w^f\right)$, and the red solid curve is $E\left(w^f \mid w^m\right)$. Each of the conditional expectations move towards the mean. Thus, a wife with a log wage of 4.0 has a husband with a conditional expected log wage of 4.13, and a husband with a log wage of 4.13 has a wife with a conditional expected log wage of 2.32 (as seen by the arrows).

An interesting feature of Figure 12 is that, for all years, when the predicted wife's log wage (PWLW) reaches 2.75 (wage = \$15.64), the predicted husband's log wage (PHLW) flattens with respect to the PWLW and converges across years. Since \$15.64/hour is not that high a wage, this suggests that assortative mating is a stronger phenomenon at lower wages than at higher wages. Figure 13 shows a somewhat different pattern where assortative mating exists over the whole period but is significantly stronger in later years (2000 and after). We still see some flattening out at log wages about 4.0 (wage = \$54.60/hour). One might wonder how two figures, both about matching over the same period of time and the same group of people, can tell such different stories. The reason is that the two figures are sorting the data in fundamentally different ways. Figure 12 aggregates PWLWs that have similar PHLWs, while Figure 13 aggregates PHLWs that have similar PWLWs. Let $A_f(w_m)$ be the function that aggregates PWLWs with similar PHLWs equal to w_m, and let $A_m(w_f)$ be the function that aggregates PHLWs with similar PWLWs equal to w_f. It is only in special cases that $A_m(A_f(w_m)) = w_m$ which is what is required for Figures 12 and 13 to provide equivalent curves.

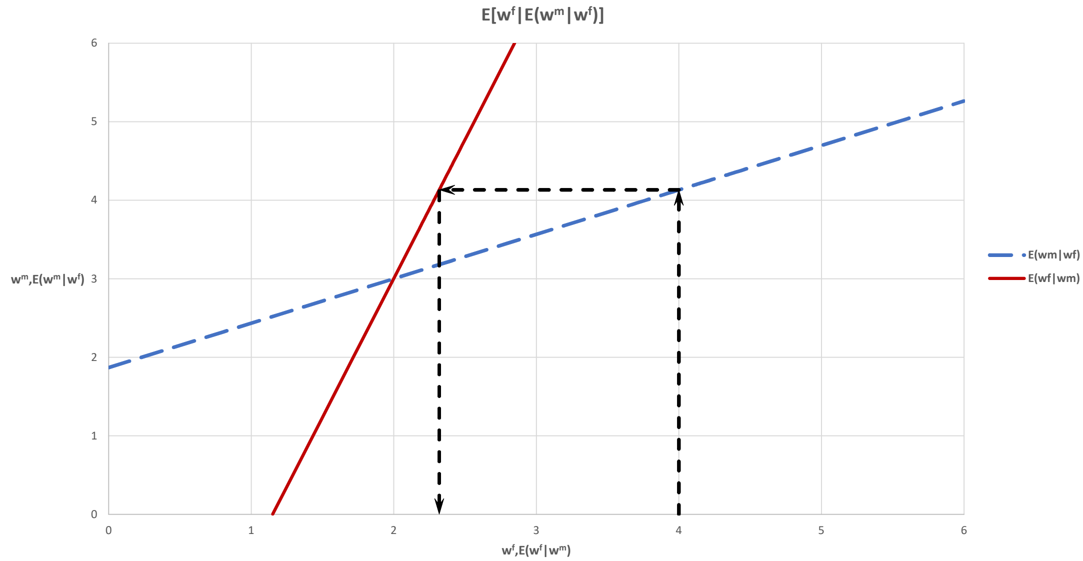


Figure 14: $E\left[w^f \mid E\left(w^m \mid w^f\right)\right]$

5.3.2 Percentile Relationships

Another way to look at the data is to turn predicted log wages of husbands and wifes into percentiles. Choo and Siow (2006b) assume this will happen implicitly and Burdett and Coles (1997) assume so explicitly. In particular, let $\widehat{F}_j\left(w^j\right)$ be the empirical distribution of predicted log wages of j = husband or wife, and use local linear regression for $\widehat{F}_j\left(w^j\right)$ on $\widehat{F}_j\left(w^{-j}\right)$ where $-j$ is the other member of the couple. The results of these regressions are displayed in Figures 15 and 16. Figure 15 shows a monotone increasing relationship between the household member percentiles with the exception of low income husbands in 1962. The relationship is very stable over time. The curves all come close to passing through $(50, 50)$, and they all have slopes less than 1. This implies sorting by percentile on average with some significant deviations due to other sorting factors (e.g., emotional attraction). Barnett (1980) presents some bivariate uniform distributions of interest that can be used to compare to Figure 15. For example, $f(u, v) = 1 + \alpha(1 - 2u)(1 - 2v)$ for $0 \leq \alpha \leq 1$, discussed by Schucany, Parr, Boyer (1978), has marginal standard uniform densities and exhibits positive assortative mating. For this case, $E(U \mid v) = (1/2) + (\alpha(1 - 2v)/6)$, implying a smaller slope than in Figure 15. An alternative is a bivariate Gaussian coppula with density,

$$
\begin{aligned}
f(u, v) &= \frac{1}{2\pi\sqrt{1-\rho^2}\phi(u)\phi(v)} \cdot \\
&\quad \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\Phi^{-1}(u)\right)^2 - 2\rho\Phi^{-1}(u)\Phi^{-1}(v) + \left(\Phi^{-1}(v)\right)^2\right]\right\}
\end{aligned} \tag{13}
$$

and conditional expectation,

$$
E(U \mid v) = \int_0^1 \Phi\left(\frac{\rho\Phi^{-1}(v) - \Phi^{-1}(u)}{\sqrt{1-\rho^2}}\right) du
$$

with average $\partial E(U \mid v)/\partial v$ varying from 0 (when $\rho = 0$) to 1 (when $\rho = 1$). The curves in Figure 15 are very similar to those that would be generated by the joint density in equation (13) with $\rho \approx 0.4$.

Figure 16 is mostly a monotone increasing relationship with exceptions for low income husbands in 2019 and middle income husbands in 1970 and 1980. The curves are not as similar as are those in Figure 15 and are not centered at $(50, 50)$ as in Figure 15. Also, globally, the slopes in Figure 16 are slightly steeper than in Figure 15. The large dip in the 1980 curve in the middle percentiles is due to the bimodal nature of the wive's expected log wage density which is unique to 1980. The other deviations from the smooth and similar behavior of curves in Figure 15 may be due to other, more subtle characteristics of the marginal or bivariate distributions.

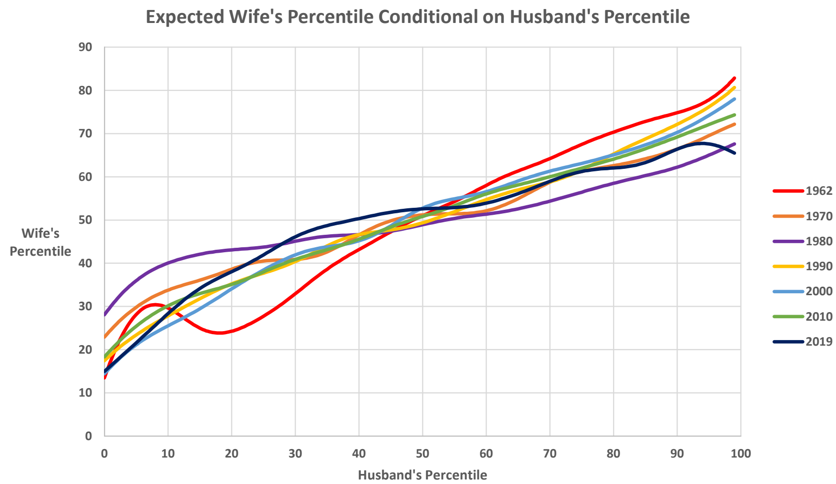


Figure 15: Expected Wife's Percentile Conditional on Husband's Percentile

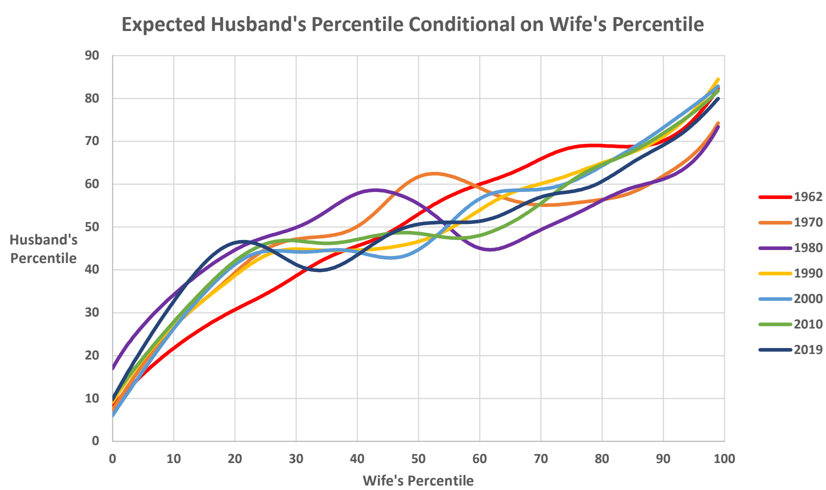


Figure 16: Expected Husband's Percentile Conditional on Wife's Percentile

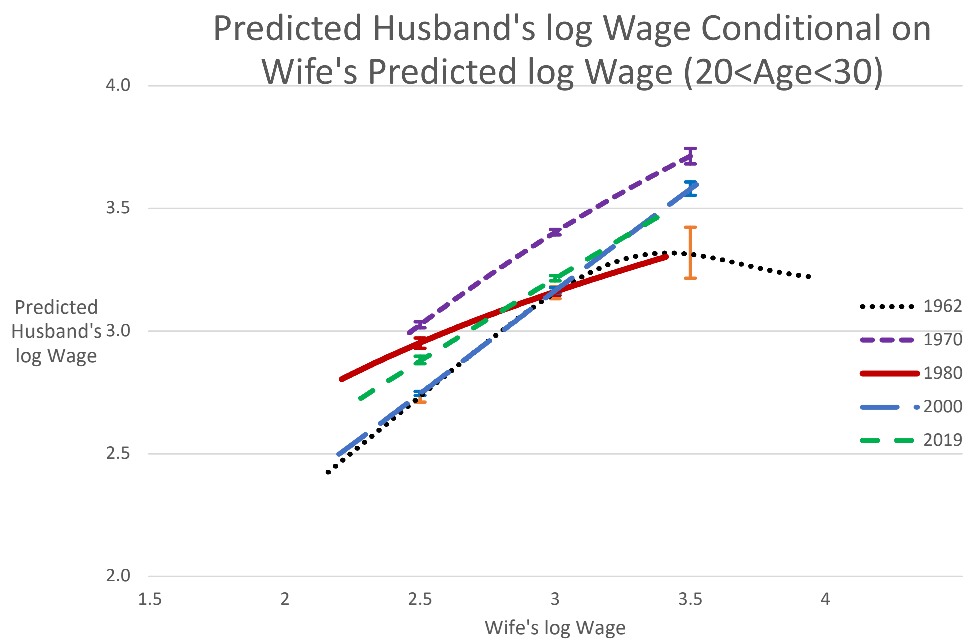


Figure 17: Predicted Husband's log Wage Conditional on Wife's Predicted log Wage (20<Age<30)

5.3.3 Cohort Effects

A potentially serious problem with the curves in Figures 12 and 13 is that they are based on all couples in the data while only a small portion of couples in the data were married recently. Assortative mating is a statement about similarities with who one marries. If people change over time, then a set of curves from a representative sample of people may not be similar to the curves from a representative sample of people who recently married. Also, since the marriage market changes as people age, it is probably important to focus on a cohort of people who recently married. Unfortunately, there is no exact way to construct a cohort of people who recently married using our data because, except for prior to and including 1970, we do not know people's age when they got married. The second best feasible alternative is to focus on cohorts of people within an age range when people usually marry (e.g., see Pencavel, 1998). Figures 17 and 18 focus on cohorts of men and women who were between the ages of 20 to 30 in five different years. For example, the dark blue dashed curve (lower than the other curves) in Figure 17 displays how PHLW changes with PWLW for the 20 to 30 year olds in 2000. The curves for the five selected years have different vertical intercepts but similar slopes, suggesting that assortative mating among people ages 20 to 30 has not changed much over the 60 years included in the data. This result differs from Greenwood and Guner (2008) and Santos and Weiss (2016) but is consistent with Gayle, Golan and Soytas (2020) and Gihleb and Lang (2020).

The curves in Figure 18 show a more interesting picture. The earliest curve (1962) is an outlier showing no wage gaps between men and women but still a significant amount of assortative mating. This is probably caused by the poor performance of the Heckman 2-step estimator for 1962 due to lack of plausible instruments. From 1970 to 2019, the curves are bunched together, but the patterns of assortative mating vary by year. In 1970, the slope of the curve declines with the PHLW. In 2000, it increases with the PHLW. In 1990 it is flat, then declines, and then increases. In 2019, it is almost linearly increasing in PHLW. Overall, we still see positive assortative mating almost always, but in more varied patterns. The positive slope confirms earlier research. Positive assortative mating extends to wages, not just non-labor market characteristics. The general trend towards stronger positive relationships between spouses' wages supports our earlier hypothesis that spouses' wages are more similar over time. However, the results contradict some of the theoretical discussions of marital sorting in the 1960's and 1970's. For example, Becker (1985) and Becker(1991) both predict negative assortative mating over wages holding other characteristics constant, in the earlier years of the analysis. The results indicate that positive assortative mating in wages existed even in 1962, if weakly. But they have nothing to say about correlation after controlling for other characteristics. It is possible that negative assortative mating in wages existed prior to 1962 or that negative partial assortative mating has always existed, but we do not have the data to consider the possibility.[24]

[24] It is worth noting that Becker's analysis was significantly after 1962, the earliest year used in our analysis.

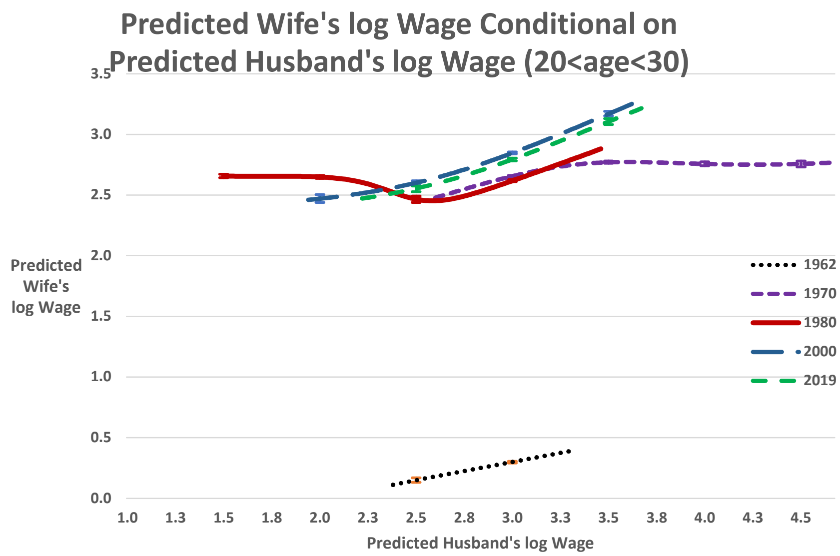


Figure 18: Predicted Wife's log Wage Conditional on Predicted Husband's log Wage (20<Age<30)

The main reason for looking at cohorts in Figures 17 and 18 instead of Figures 12 and 13 was assortative mating is a statement about similarities when marrying and the characteristics of couples can change as they age.[25] To see how important this concern is, we can look at how PHLWs and PWLWs change as they age. Figure 19 shows how the relationship between PWLW and PHLW changes as a couple from the 1970 cohort ages from $20-30$ years of age to $40-50$ years of age. Figure 19 shows PHLW steadily increasing as a function of PWLW.

Figure 20 shows how the relationship between PHLW and PWLW changes as a couple from the 1970 cohort ages from $20-30$ years of age to $40-50$ years of age. This figure shows little assortative mating in 1970 among the $20-30$ year old cohort; PWLW is almost flat as a function of PHLW. But, by 1980, there is significantly more assortative mating. This suggests that there was some characteristic, such as education, that had a small effect on PWLW in 1970, at least for new job entrants, but then later led to significant increases in PWLW 10 years later and beyond. There is also a dip in the curve for $40-50$ years; we do not know how to interpret the dip. Altogether, this figure suggests caution in using wages of newly married couples to measure assortative mating as much of the variation in wages occurs later in life and in a somewhat predictable way. A large literature on income generating processes, including Macurdy (1982), Baker (1997), Baker and Solon (2003), and Guvenen (2009), finds significant evidence of heterogeneous wage growth. An alternative set of work, including Abowd and Card (1989), Deaton and Paxson (1994), Gottschalk and Moffitt (1995), and Meghir and Pistaferri (2004), model the income process with permanent shocks and/or random walks. The first approach suggests that individuals have much information about their future wages, and the second assumes future wages exhibit significant shocks unknown early in an individual's life. Our results lend some support for the first approach.

5.3.4 Sharing Within Families

Chiappori, Fortin, and Lacroix (2002) use a collective model to estimate a transfer function from the husband to the wife,

$$\begin{aligned}\phi &= -56.64\log w_f - 25.94\log w_m + 20.06\log w_f\log w_m \\ &\quad +0.70v + 216.28s_1 + 4.31s_2 + \kappa(z) + \varepsilon\end{aligned}$$

where $\log w_f$ is the log wage of the wife, $\log w_m$ is the log wage of the husband, v is non-labor income, s_1 is the age/race/state-specific sex ratio, s_2 is a measure of divorce law characteristics, $\kappa(z)$ is an unidentified effect of other household characteristics, and ε is the error.[26] Figure 21 shows how the share of household resources going to the husband $(w_m+\phi)/(w_m+w_f+v)$

[25] Also, divorce rates varying over joint realizations of the husband's and wife's wages would change the amount of assortative mating as couples age. For example, Bertrand, Kamenica, and Pan (2015) provide evidence that couples with the husband earning less than the wife are more likely to divorce. This changes the nature of observed assortative mating as the divorces occur.

[26] It is not clear how to interpret the error in Chiappori, Fortin, and Lacroix (2012). In fact, the structural equations in Chiappori, Fortin, and Lacroix (2012) are the labor supply equations, and they have no errors attached to them. The addition of errors might represent unobserved

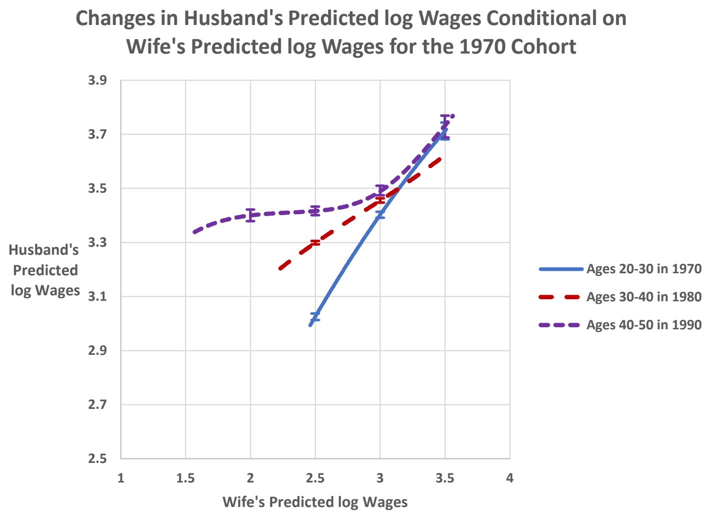


Figure 19: Changes in Husband's Predicted log Wages Conditional on Wife's Predicted log Wages for the 1970 Cohort

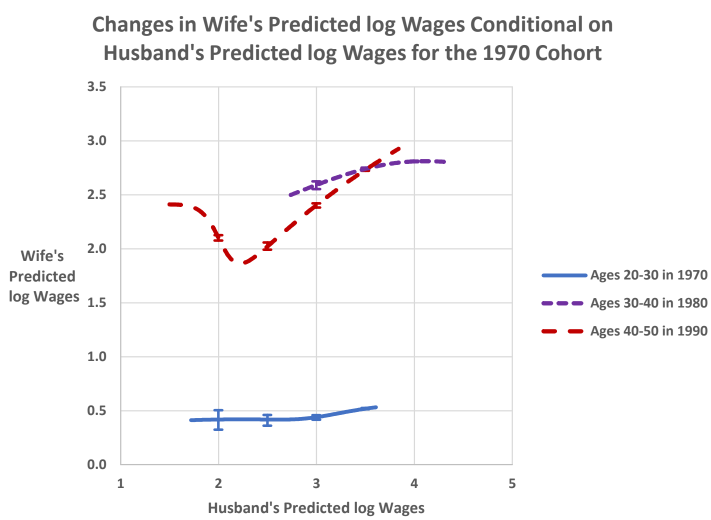


Figure 20: Changes in Wife's Predicted log Wages Conditional on Husband's Predicted log Wages for the 1970 Cohort

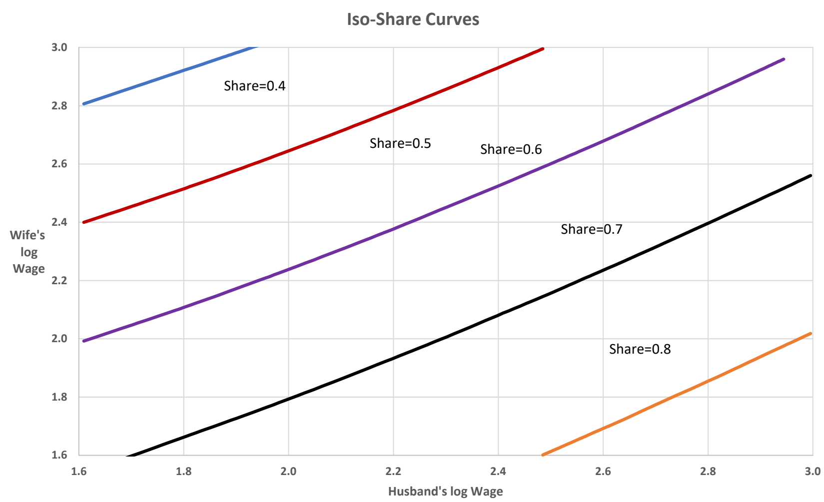


Figure 21: Husband's Share

varies with the wife's log wage and the husband's log wage.[27] The contour curves presented in the figure have slopes varying from 0.6 when the husband's wage is low relative to the wife's wage to 0.85 when the husband's wage is high relative to the wife's wage. This implies a moderately small amount of sharing (and caring) consistent with results in Chiappori, Fortin, and Lacroix (2002) and in Friedberg and Stern (2014).

A concern with this approach is whether we can interpret the estimates in Chiappori, Fortin, and Lacroix (2002) as the correct structure. Choo and Siow (2006a, 2006b) and Chiappori, Iyigun, and Weiss (2009) argue that one should not separately analyze assortative mating and marital sharing. For example, part of the results seen in Figures 12 and 13 is due to the increasing average educational attainment of women. On the one hand, this increases a woman's log wage causing her transfer to her husband to decline; this is the effect measured in Figure 21. On the other hand, other women's wages also increase, thus improving the bargaining power of the husband. Since neither the average log wage of women nor the average education level of women in the husband's marriage market is included in the model, the model is not controlling for this second effect.[28] To some degree, these types of effects are the meat of Greenwood et al. (2014). But neither Greenwood et al. (2014) nor its companion model in Greenwood et al. (2016) have anything to say about the household sharing rule.

A second concern is that some of the literature provides structural estimates showing that the gains to marriage for women decline with the woman's wage (e.g., Van der Klaauw (1996); Mazzocco, Ruiz, and Yamaguchi, 2007). This suggests that a woman does not receive as much from a higher wage in marriage as she would if single. This contradicts the results in Chiappori, Fortin, and Lacroix (2002), and it implies a surface for Figure 21 with a less steep slope in the wife's log wage, reducing the gains associated with higher female wages.

5.4 Assortative Mating and Education

Smits, Ultee, and Lammers (1998), Fernandez, Guner, and Knowles (2005), Hou and Myles (2008), Siow (2015), Eika, Mogstad, and Zafar (2017) find strong evidence of positive assortative mating in marriage by education.[29] Also, Mare (1991) and Pencavel (1998) provide evidence of changing degrees of assortative mating on education and suggest explanations for the changes, Gihleb (2013) and Siow (2015) find that the degree of assortative mating on education has not changed much over time, and Chiappori, Salanié, and Weiss (2011) argue that the US marriage market for individuals born between 1943 and 1972 can be described as having a constant value of a statistic (TP2) similar to T from equation (10) with time-varying marriage rates. Eika, Mogstad, and Zafar (2017) find stronger assortative mating at lower levels of education than higher levels of education; this is somewhat consistent with our results showing more assortative mating at low wages than at

variation in preferences, unobserved factors affecting the transfer rule, or measurement error in hours data. Each of these possibilities would affect the transfer equation, and therefore the share equation, in different ways.

[27] Wages are measured as hourly wages. All necessary assumptions are made based on sample averages provided in Table 1 of the paper.

[28] The increased log wage of the wife also changes the optimal sharing of household production. Chiappori, Fortin and Lacroix (2002) control for this effect.

[29] Arum, Roksa, and Budig (2008) present evidence that couples sort on the quality of the school where one attends college.

Table 7: Effects of Predicted Wife's Education on log Predicted Wife's Wage

Dependent variable: log(Predicted Wife's Wage)

Variable	1962	1967	1970	1972	1975	1980	1985	1990
Continuous Education								
Predicted Wife's	0.026 ***	0.034 ***	0.034 ***	0.039 ***	0.032 ***	0.025 ***	0.030 ***	0.092 ***
Education	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.004)	(0.009)
Constant	2.430 ***	2.449 ***	2.276 ***	2.303 ***	2.350 ***	2.355 ***	2.261 ***	1.482 ***
	(0.056)	(0.058)	(0.060)	(0.058)	(0.060)	(0.067)	(0.063)	(0.129)
R squared	0.820	0.825	0.889	0.922	0.871	0.839	0.917	0.916
Discrete Diploma								
Predicted Wife's	0.042 ***	0.044 ***	0.044 ***	0.047 ***	0.038 ***	0.035 ***	0.040 ***	0.100 ***
Education	(0.003)	(0.005)	(0.007)	(0.005)	(0.002)	(0.002)	(0.001)	(0.002)
Constant	2.234 ***	2.316 ***	2.143 ***	2.194 ***	2.278 ***	2.233 ***	2.131 ***	1.356 ***
	(0.036)	(0.061)	(0.079)	(0.057)	(0.029)	(0.032)	(0.017)	(0.031)
R squared	0.969	0.957	0.933	0.961	0.987	0.971	0.984	0.993

Notes:
1) Single-starred items are statistically significant at the 5% level, double-starred items at the 1% level, and triple-starred at the 0.1% level.
2) Standard errors are available from the corresponding author.

Table 7: Effects of Predicted Wife's Education on log Predicted Wife's Wage (Continued)

Dependent variable: log(Predicted Wife's Wage)

	Year							
Variable	1995	2000	2005	2007	2010	2014	2015	2019
Continuous Education								
Predicted Wife's	0.087 ***	0.121 ***	0.130 ***	0.134 ***	0.139 ***	0.152 ***	0.151 ***	0.105 ***
Education	(0.007)	(0.005)	(0.006)	(0.010)	(0.007)	(0.006)	(0.004)	(0.002)
Constant	1.527 ***	1.147 ***	0.992 ***	0.946 ***	0.878 ***	0.520 ***	0.611 ***	1.352 ***
	(0.097)	(0.066)	(0.079)	(0.135)	(0.096)	(0.100)	(0.058)	(0.030)
R squared	0.948	0.973	0.980	0.962	0.973	0.985	0.987	0.993
Discrete Diploma								
Predicted Wife's	0.099 ***	0.134 ***	0.134 ***	0.137 ***	0.131 ***	0.146 ***	0.149 ***	0.105 ***
Education	(0.003)	(0.003)	(0.006)	(0.005)	(0.007)	(0.008)	(0.006)	(0.004)
Constant	1.354 ***	0.946 ***	0.930 ***	0.889 ***	0.983 ***	0.617 ***	0.647 ***	1.355 ***
	(0.048)	(0.037)	(0.089)	(0.074)	(0.106)	(0.109)	(0.083)	(0.063)
R squared	0.990	0.998	0.993	0.993	0.988	0.988	0.990	0.994

Notes:
1) Single-starred items are statistically significant at the 5% level, double-starred items at the 1% level, and triple-starred at the 0.1% level.
2) Standard errors are available from the corresponding author.

high wages. Gayle, Golan and Soytas (2020) show that the characteristics of assortative mating in education vary by race. Meanwhile, Fernandez, Guner, and Knowles (2005) find some interesting variation in the amount of assortative mating across countries and tie it to variation in country-specific skill premia.

To further examine the sorting trends in each year in our data, we analyze how much of the sorting over wages can be explained by sorting over education. We use an OLS regression of the local linear regression predicted wages displayed in Figures 12 and 13 on an analogous local linear regression predicted education term (as described in Section 4.3), both given the spouse's wage or education. The regression results are provided in Tables 7 and 8 and then graphed in Figure 22. Wife's education has very small effects on log wage prior to 1980. From 1980 on, the effect grows and becomes large, and it becomes larger than for husbands.

5.5 Positive Assortative Mating Test Results

We perform two tests each for log predicted wages and education separately and for a range of sample years, displayed in Table 9. For the log predicted wage tests, we divide the space of log wages for men and women into deciles, implying that $M = F = 10$ which, using equation (11), leads to degrees of freedom of 81. Using a normal approximation for the χ^2_{81} distribution, the critical value for the LR test statistics is $81 + 1.96\sqrt{162} = 105.9$. For the education tests, $M = F = 5$ which implies that the degrees of freedom are 16. The critical value for the LR test statistics is 26.3. For all years and all tests, the test statistics are very large leading us to reject the null hypothesis of no assortative mating. The LR test statistics are an order of magnitude larger than the Wald test statistics (see Kang and Stern (2021) for more discussion).

Table 8: Effects of Predicted Husband's Education on log Predicted Husband's Wage

Dependent variable: log(Predicted Husband's Wage)

Variable	1962	1967	1970	1972	1975	1980	1985	1990
Continuous Education								
Predicted Husband'	0.081 **	0.048 ***	0.038 ***	0.069 ***	0.037 ***	0.019 ***	0.045 ***	0.080 ***
Education	(0.029)	(0.006)	(0.006)	(0.010)	(0.007)	(0.003)	(0.004)	(0.004)
Constant	2.148 ***	2.688 ***	2.808 ***	2.405 ***	2.904 ***	3.103 ***	2.704 ***	2.112 ***
	-0.312	(0.074)	(0.069)	(0.118)	(0.082)	(0.030)	(0.054)	(0.058)
R squared	0.711	0.839	0.934	0.940	0.873	0.917	0.959	0.946
Discrete Diploma								
Predicted Husband'	0.107 *	0.066 ***	0.044 **	0.084 ***	0.044 ***	0.023 ***	0.051 ***	0.093 ***
Education	(0.043)	(0.004)	(0.008)	(0.006)	(0.005)	(0.001)	(0.005)	(0.002)
Constant	1.745 **	2.425 ***	2.717 ***	2.183 ***	2.791 ***	3.045 ***	2.615 ***	1.905 ***
	(0.525)	(0.054)	(0.094)	(0.070)	(0.065)	(0.008)	(0.059)	(0.030)
R squared	0.702	0.971	0.915	0.986	0.962	0.996	0.975	0.998

Notes:

1) Single-starred items are statistically significant at the 5% level, double-starred items at the 1% level, and triple-starred at the 0.1% level.

2) Standard errors are available from the corresponding author.

Table 8: Effects of Predicted Husband's Education on log Predicted Husband's Wage (Continued)

Dependent variable: log(Predicted Husband's Wage)

	Year							
Variable	1995	2000	2005	2007	2010	2014	2015	2019
Continuous Education								
Predicted Husband'	0.092 ***	0.089 ***	0.099 ***	0.100 ***	0.092 ***	0.077 ***	0.084 ***	0.090 ***
Education	(0.004)	(0.007)	(0.006)	(0.005)	(0.007)	(0.006)	(0.003)	(0.006)
Constant	1.822 ***	2.037 ***	1.873 ***	1.854 ***	1.964 ***	2.193 ***	2.096 ***	2.161 ***
	(0.055)	(0.097)	(0.078)	(0.072)	(0.088)	(0.074)	(0.043)	(0.079)
R squared	0.969	0.969	0.972	0.983	0.973	0.963	0.987	0.973
Discrete Diploma								
Predicted Husband'	0.108 ***	0.102 ***	0.109 ***	0.110 ***	0.101 ***	0.075 ***	0.083 ***	0.086 ***
Education	(0.003)	(0.007)	(0.006)	(0.007)	(0.009)	(0.005)	(0.005)	(0.008)
Constant	1.593 ***	1.835 ***	1.718 ***	1.717 ***	1.846 ***	2.210 ***	2.112 ***	2.220 ***
	(0.041)	(0.083)	(0.073)	(0.086)	(0.112)	(0.064)	(0.066)	(0.108)
R squared	0.997	0.989	0.992	0.990	0.978	0.987	0.990	0.980

Notes:

1) Single-starred items are statistically significant at the 5% level, double-starred items at the 1% level, and triple-starred at the 0.1% level.

2) Standard errors are available from the corresponding author.

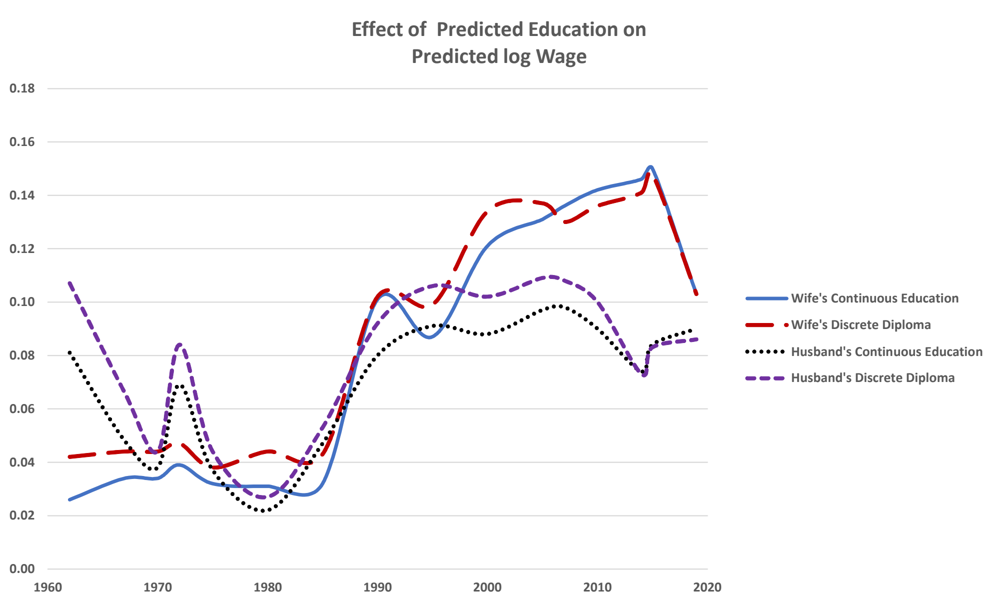


Figure 22: Effect of Predicted Education on Predicted log Wage

Table 9: Assortative Mating Test Statistics

	log Predicted Wage		Years of Education	
Year	Wald Test	LR test	Wald Test	LR test
1962	525.5	7314.0	2443.6	5642.5
1967	773.3	6298.7	2981.7	5968.1
1970	604.1	5317.9	4453.5	8447.1
1972	497.9	7318.3	3952.1	8404.5
1975	269.1	4329.7	3456.1	7756.3
1978	318.5	2382.0	3857.4	8507.5
1980	672.9	3119.2	5022.2	10139.1
1985	871.8	3101.7	5121.0	9977.7
1990	564.9	5818.5	6061.8	10134.3
1995	451.1	5098.8	4344.7	8250.6
2000	544.9	5764.1	3799.6	7642.9
2005	631.3	8087.4	5372.7	11930.9
2007	654.3	7502.8	5108.6	11536.9
2010	807.4	7046.4	4525.0	10983.2
2014	1004.0	8168.8	3654.2	10317.1
2015	935.5	7930.8	3946.9	10395.8
2018	1043.6	5283.8	2964.6	8854.5
2019	1183.9	5208.0	2897.4	8604.3
# Cells	10		5	
DF	81	99	16	24

6 Conclusion

This paper adds to the literature examining effects of economic characteristics on marital sorting by using Current Population Survey data with local linear regression to flexibly model the relationship between spouses' log wages and an OLS decomposition to examine the effects of education on this sorting over wage. Our primary departure from existing research is to focus on underlying wages, instead of earnings or labor supply, which we believe better captures the labor market characteristics of each spouse relevant to assortative mating. The local linear regression also allows us to examine how the association between spouses' wages varies given wage level and to avoid endogeneity issues associated with the spousal labor market participation decision.

We find evidence for positive assortative mating over wages, in concordance with the existing research, and extend that research to compare the assortative mating patterns of recent years to the 1960's and 1970's. Our findings indicate the existence of positive assortative mating even as early as 1962, contrary to Becker (1973). The education sorting analysis across years shows an increasing percentage of sorting over wages is driven by the association between spouses' education levels. Also, our results generally find more assortative mating at lower levels of education and wages than at higher levels.

We considered a number of other questions associated with estimating the degree of assortative mating. We explored what was the best time in a marriage to measure assortative mating. Theoretically, because assortative mating involves how couples sort themselves in marriage, it might seem that the relevant measure is assortativeness at the the time of marriage or very early in a marriage. On the other hand, there might be factors that are important for sorting at the beginning of a marriage that become observable to an outside observer only much later. In particular, there might be important factors, including level of and success in education that are observable to couples at the time of marriage that manifest themselves into significant heterogeneity in wages much later. We find strong evidence of this in that the degree of assortative mating is much greater later in couples' marriages than at the time of marriage.

We also explore the value of transforming predicted log wages into percentiles. Most of the theoretical literature implies that measuring assortative mating in terms of percentiles would be more informative than in terms of log wages. Our results have nothing to say about which is preferred. But the percentile results are interesting in and of themselves and lead to more understanding of the sorting process.

Finally, we suggest some new test statistics associated with assortative mating that are easier to use than those already existing in the literature. The test statistics imply rejection of the null hypothesis of no assortative mating in favor of the existence of positive assortative mating, both for log predicted wages and level of education, for every CPS sample year.

7 Appendix: Weeks Worked Measurement Error

Figure 7 presents the measurement error of the *weeks worked* variable varying with *annual hours worked* of the person. Maximum weeks worked overestimates the actual weeks worked. One can see that there exists unignorable measurement errors for people whose annual hours worked were less than 2000 hours. Using the fact that *weeks worked* variables of later years are collected both as continuous and categorical variables, we identify the measurement errors from using the bracketed variable. Then we adjust the measurement errors using local-linear regression. In particular, denote h as an arbitrary value of total hours worked by an individual from year t, w_{it}^c as a *weeks worked* variable for person i in year t collected in a continuous manner, and w_{it}^d as a *weeks worked* variable for person i in year t collected in a discrete manner. We calculate the *weeks worked* error as

$$E(\widehat{m} \mid h) = \frac{1}{T} \sum_t \widehat{\mu}_0(h),$$

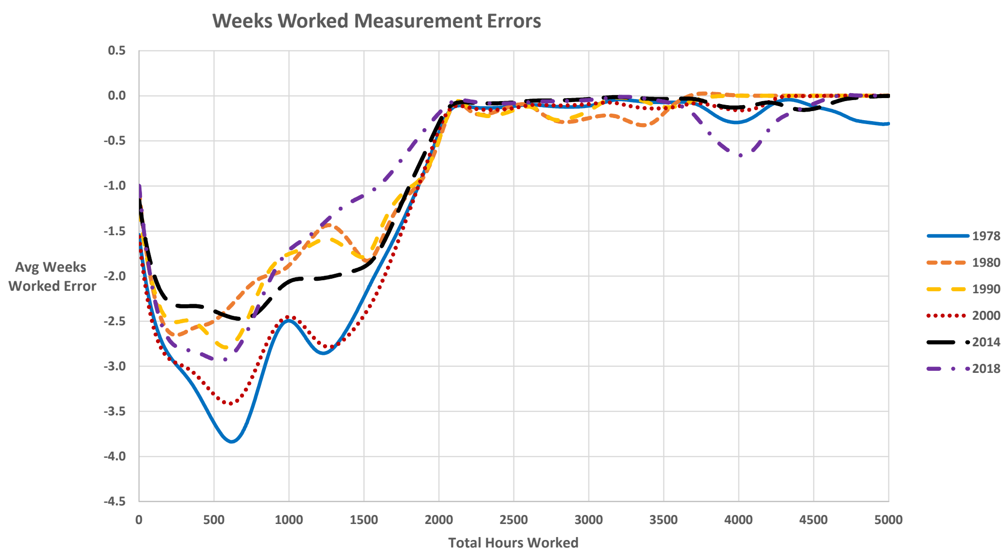
T is the number of sample years,

$$\begin{pmatrix} \widehat{\mu}_0(h) \\ \widehat{\mu}_1(h) \end{pmatrix} = \left[\sum_i K\left(\frac{h_{it}-h}{b}\right) \begin{pmatrix} 1 \\ h_{it} \end{pmatrix} \begin{pmatrix} 1 & h_{it} \end{pmatrix} \right]^{-1} \left[\sum_i K\left(\frac{h_{it}-h}{b}\right) \begin{pmatrix} 1 \\ h_{it} \end{pmatrix} m_{it} \right]$$

and

$$m_{it} = w_{it}^c - w_{it}^d$$

is the measurement error. Then, for $t < 1978$, we replace the *weeks worked* data with the error-adjusted *weeks worked* interpolating with h-points close to that consistent with w_{it}^d.



References

[1] Abowd, John and David Card (1989). "On the Covariance Structure of Earnings and Hours Changes." *Econometrica*. 57(2): 411-445.

[2] Arabmazar, Abbas and Peter Schmidt (1982). "An Investigation of the Robustness of the Tobit Estimator to Non-Normality." *Econometrica*. 50(4): 1055-1063.

[3] Arum, Richard, Josipa Roksa, and Michelle Budig (2008). "The Romance of College Attendance: Higher Education, Stratification and Mate Selection." *Research in Social Stratification and Mobility*. 26: 107–121.

[4] Asadi, Majid and Somayeh Zarezadeh (2020). "A Unified Approach to Constructing Correlation Coefficients between Random Variables." *Metrika*. 83(6): 657-676.

[5] Ashworth, Jared, V. Joseph Hotz, Arnaud Maurel, and Tyler Ransom (2020). "Changes across Cohorts in Wage Returns to Schooling and Early Work Experiences." *Journal of Labor Economics*. forthcoming.

[6] Ashworth, Jared and Tyler Ransom (2019). "Has the College Wage Premium Continued to Rise? Evidence from Multiple U.S. Surveys." *Economics of Education Review*. 69(1): 149-154.

[7] Bailey, Martha, John DiNardo, and Bryan Stuart (2020). "The Economic Impact of a High National Minimum Wage: Evidence from the 1966 Fair Labor Standards Act." Unpublished manuscript.

[8] Bailey, Martha, Brad Hershbein, and Amalia Miller (2012). "The Opt-In Revolution? Contraception and the Gender Gap in Wages." *American Economic Journal: Applied Economics*. 4(3): 225-254.

[9] Baker, Michael (1997). "Growth-Rate Heterogeneity and the Covariance Structure of Life-Cycle Earnings." *Journal of Labor Economics*. 15(2): 338-375.

[10] Baker, Michael and Gary Solon (2003). "Earnings Dynamics and Inequality among Canadien Men, 1976-1992: Evidence from Longitudinal Income Tax Records." *Journal of Labor Economics*. 21(2): 267-288.

[11] Barnett, Vic (1980). "Some Bivariate Uniform Distributions." *Communications in Statistics - Theory and Methods*. 9(4): 453-461.

[12] Becker, Gary (1973). "A Theory of Marriage: Part 1." *Journal of Political Economy*. 81(4): 813-846.

[13] Becker, Gary (1985). "Human Capital, Effort, and the Sexual Division of Labor." *Journal of Labor Economics*. 3(1, part 2): S33-S58.

[14] Becker, Gary (1991). *A Treatise on the Family.* Cambridge, MA: Harvard University Press.

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

[16] Bianchi, Suzanne, John Robinson, and Melissa Milkie (2006). *Changing Rhythms of American Family Life*. Russell Sage Foundation: New York.

[17] Blackburn, McKinley and David Bloom (1995). "Changes in the Structure of Family Income Inequality in the United States and Other Industrialized Nations during the 1980s." *Research in Labor Economics*. 141–170.

[18] Blau, Francine (1998). "Trends in the Well-Being of American Women, 1970-1995." *Journal of Economic Literature*. 36(1): 112-165.

[19] Blau, Francine and Lawrence Kahn (2007). "Changes in the Labor Supply Behavior of Married Women: 1980-2000." *Journal of Labor Economics*. 25(3): 393-438.

[20] Blau, Francine and Lawrence Kahn (2017). "The Gender Wage Gap: Extent, Trends, and Explanations." *Journal of Economic Literature*. 55(3): 789-865.

[21] Blundell, Richard, Pierre-André Chiappori, Thierry Magnac, and Costas Meghir (2005). "Collective Labour Supply: Heterogeneity and Nonparticipation." *Review of Economic Studies*. 74(2): 417-445.

[22] Blundell, R., A. Gosling, H. Ichimura, and C. Meghir (2007). "Changes in the Distribution of Male and Female Wages Accounting for Employment Composition Using Bounds." *Econometrica*. 75(2): 323–363.

[23] Borjas, George (2007). *Mexican Immigration to the United States*. Chicago: University of Chicago Press.

[24] Bredemeier, Christian and Falko Juessen (2013). "Assortative Mating and Female Labor Supply." *Journal of Labor Economics*. 31(3): 603-631.

[25] Burdett, Ken and Melvyn Coles (1997). "Marriage and Class." *Quarterly Journal of Economics*. 112(1): 141-168.

[26] Cancian, Maria, Sheldon Danziger, and Peter Gottschalk (1993). "Working Wives and Family Income Inequality among Married Couples." *Uneven Tides: Rising Inequality in America*. (eds.) Sheldon Danziger and Peter Gottschalk. Russell Sage Foundation: 195–221.

[27] Cancian, Maria and Deborah Reed (1999). "The Impact of Wives' Earnings on Income Inequality: Issues and Estimates." *Demography*. 36(174): 173-184.

[28] Card, David (2001). "Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems." *Econometrica*. 69(5): 1127-1160.

[29] Cattaneo, Matias and Michael Jansson (2018). "Kernel-Based Semiparametric Estimators: Small Bandwidth Asymptotics and Bootstrap Consistency." *Econometrica*. 86(3): 955-995.

[30] Chang, S. (2011). "Simulation Estimation of Two-Tiered Dynamic Panel Tobit Models with an Application to the Labor Supply of Married Women." *Journal of Applied Econometrics*. 26(5): 854–871.

[31] Charles, Kerwin, Erik Hurst, Alexandra Killewald (2013). "Marital Sorting and Parental Wealth." *Demography*. 50: 51-70.

[32] Chiappori, Pierre-André (1988). "Rational Household Labor Supply." *Econometrica*. 56(1): 63-90.

[33] Chiappori, Pierre-André (1992). "Collective Labor Supply and Welfare." *Journal of Political Economy*. 100(3): 437-467.

[34] Chiappori, Pierre-André, Monica Costa Dias, Sam Crossman, and Costas Meghir (2020). "Changes in Assortative Matching and Inequality in Income: Evidence for the UK." Unpublished manuscript.

[35] Chiappori, Pierre-André, Monica Costa Dias, and Costas Meghir (2020). "Changes in Assortative Matching: Theory and Evidence for the US." Unpublished manuscript.

[36] Chiappori, Pierre-André, Bernard Fortin, and Guy Lacroix (2002). "Marriage Market, Divorce Legislation, and Household Labor Supply." *Journal of Political Economy*. 110(1): 37-72.

[37] Chiappori, Pierre-André, Murat Iyigun, and Yoram Weiss (2009). "Investment in Schooling and the Marriage Market." *American Economic Review*. 99(5): 1689-1713.

[38] Chiappori, Pierre-André, Bernard Salanié, and Yoram Weiss (2017). "Partner Choice, Investment in Children, and the Marital College Premium." *American Economic Review*. 107(8): 2109–67.

[39] Choo, Eugene and Aloysius Siow (2006a). "Estimating a Marriage Matching Model with Spillover Effects." *Demography*. 43(3): 463–490.

[40] Choo, Eugene and Aloysius Siow (2006b). "Who Marries Whom and Why." *Journal of Political Economy*. 114(1): 175–201.

[41] Coslett, Steven (1991). "Semiparametric Estimation of a Regression Model with Sample Selectivity." *Nonparametric and Semiparametric Methods in Econometrics and Statistics*. (eds.) William Barnett, James Powell, and George Tauchen. Cambridge University Press, Cambridge: 175-197.

[42] Dagsvik, John (1994). "Discrete and Continuous Choice, Max-Stable Processes, and Independence from Irrelevant Attributes." *Econometrica*. 62(5): 1179-1205.

[43] Deaton, Angus and Christina Paxson (1994). "Intertemporal Choice and Inequality." *Journal of Political Economy*. 102(3): 384-394.

[44] Derenoncourt, Ellora and Claire Montialoux (2020). "Minimum Wages and Racial Inequality." *Quarterly Journal of Economics*. qjaa031, https://doi.org/10.1093/qje/qjaa031.

[45] Duflo, Esther (2003). "Grandmothers and Granddaughters: Old Age Pension and Intra-Household Allocation in South Africa." *World Bank Economic Review*. 17(1): 1-25.

[46] Duflo, Esther and Christopher Udry (2004). "Intrahousehold Resource Allocation in Côte d'Ivoire: Social Norms, Separate Accounts and Consumption Choices." National Bureau of Economics Research Working Paper No. 10498.

[47] Duncan, Gregory and Duane Leigh (1985). "The Endogeneity of Union Status: An Empirical Test." *Journal of Labor Economics*. 3(3): 385-402.

[48] Dupuy, Arnaud and Alfred Galichon (2014). "Personality Traits and the Marriage Market." *Journal of Political Economy*. 122(6): 1271–1319.

[49] Eckstein, Zvi and Éva Nagypál (2004). "The Evolution of U.S. Earnings Inequality: 1961-2002." *Federal Reserve Bank of Minneapolis Quarterly Review*. 28(2): 10-29.

[50] Eika, Lasse, Magne Mogstad, and Basit Zafar (2017). "Educational Assortative Mating and Household Income Inequality." Unpublished manuscript.

[51] Escanciano, Juan Carlos, David Jacho-Chavez, and Arthur Lewbel (2016). "Identification and Estimation of Semiparametric Two-Step Models." *Quantitative Economics*. 7: 561-589.

[52] Fan, Jianqing and Irene Gijbels (1996). *Local Polynomial Modelling and its Applications: Monographs on Statistics and Applied Probability. 66*. CRC Press.

[53] Fernandez, Raquel, Nezih Guner, and John Knowles (2005). "Love and Money: A Theoretical and Empirical Analysis of Household Sorting and Inequality." *Quarterly Journal of Economics*. 120(1): 273–344.

[54] Friedberg, Leora and Steven Stern (2014). "Marriage, Divorce, and Asymmetric Information." *International Economic Review*. 55(4): 1155-1199.

[55] Friedberg, Leora and Anthony Webb (2005). "The Chore Wars: Household Bargaining and Leisure Time." Unpublished manuscript.

[56] Gayle, George, Limor Golan and Mehmet Soytas (2020). "What Accounts for the Racial Gap in Time Allocation and Intergenerational Transmission of Human Capital?." *Journal of Political Economy*. forthcoming.

[57] Ge, Suqin (2011). "Women's College Decisions: How Much does Marriage Matter?" *Journal of Labor Economics*. 29(4): 773-818.

[58] Gihleb, Rania (2013). "Educational Assortative Mating for Blacks and Whites since the 1960s." Unpublished manuscript.

[59] Gihleb, Rania and Osnat Lifshitz (2013). "Dynamic Effects of Educational Assortative Mating on Labor Supply." Unpublished manuscript.

[60] Gihleb, Rania and Kevin Lang (2020). "Educational Homogamy and Assortative Mating Have not Increased." *Change at Home, in the Labor Market, and On the Job*, *Research in Labor Economics*. 48. (eds.) Polachek, S. and K. Tatsiramos. 1-26.

[61] Goldin, Claudia (1990). *Understanding the Gender Gap: An Economic History of American Women*. Oxford University Press, Oxford.

[62] Goldin, Claudia (2006). "The 'Quiet Revolution That Transformed Women's Employment, Education, and Family." *American Economic Review: Papers and Proceedings*. 96(2): 1–21.

[63] Gourieroux, Christian, Alberto Holly, and Alain Monfort (1982). "Likelihood Ratio Test, Wald Test, and Kuhn-Tucker Test in Linear Models with Inequality Constraints on the Regression Parameters." *Econometrica*. 50(1): 63-80.

[64] Gottschalk, Peter and Robert Moffitt (1995). "Trends in the Covariance Structure of Earnings in the US: 1969-1987." Unpublished manuscript.

[65] Gray, Jerry (1997). "The Fall in Men's Return to Marriage: Declining Productivity Effects or Changing Selection?" *Journal of Human Resources*. 32(3): 481-504.

[66] Greenwood, Jeremy and Nezih Guner (2008). "Marriage and Divorce Since World War II: Analyzing the Role of Technological Progress on the Formation of Households." *NBER Macroeconomics Annual*. 23(1): 231–276.

[67] Greenwood, Jeremy, Nezih Guner, Georgi Kocharkov, and Cezar Santos (2014). "Marry Your Like: Assortative Mating and Income Inequality." *American Economic Review*. 104(5): 348-353.

[68] Greenwood, Jeremy, Nezih Guner, Georgi Kocharkov, and Cezar Santos (2016). "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.

[69] Gustman, Alan and Thomas Steinmeier. (1986). "A Structural Retirement Model." *Econometrica*. 54(3): 555-584.

[70] Guvenen, Fatih (2009). "An Empirical Investigation of Labor Income Processes." Review of Economic Dynamics." 12: 58-79.

[71] Haddad, Lawrence and John Hoddinott (1994). "Women's Income and Boy-Girl Anthropometric Status in the Cóte d'Ivoire." *World Development*. 22(4): 543-553.

[72] Hansen, Bruce (2021). *Econometrics.* Unpublished manuscript. https://www.ssc.wisc.edu/~bhansen/econometrics/.

[73] Härdle, Wolfgang (1992). *Applied Nonparametric Regression*. Cambridge University Press.

[74] Heckman, James (1979). "Sample Selection Bias as a Specification Error." *Econometrica*. 47(1): 153-161.

[75] Hersch, Joni (1991). "Male-Female Differences in Hourly Wages: The Role of Human Capital, Working Conditions, and Housework." *Industrial and Labor Relations Review*. 44(4): 746-759.

[76] Hou, Feng and John Myles (2008). "The Changing Role of Education in the Marriage Market: Assortative Marriage in Canada and the United States since the 1970s." *Canadian Journal of Sociology*. 33(2): 337-366.

[77] Huber, Martin and Giovanni Mellace (2011). "Testing Instrument Validity in Sample Selection Models." Unpublished manuscript.

[78] Hyslop, Dean (2001). "Rising U.S. Earnings Inequality and Family Labor Supply: The Covariance Structure of Intrafamily Earnings." *American Economic Review*. 91(4): 755–777.

[79] Ichimura, Hidehiko (1993). "Semiparametric Least Squares and Weighted SLS Estimation of Single Index Models." *Journal of Econometrics*. 58(1-2): 71-120.

[80] Ichimura, Hidehiko and Lung Fei Lee (1991): "Semiparametric Least Squares Estimation of Multiple Index Models: Single Equation Estimation." *Nonparametric and Semiparametric Methods in Econometrics and Statistics*. (eds.) William Barnett, James Powell, and George Tauchen. 3–49. Cambridge: Cambridge University Press.

[81] Ichimura, Hidehiko and Petra Todd (2007). "Implementing Nonparametric and Semiparametric Estimators." *Handbook of Econometrics, vol. 6B*. (eds., Robert Engle and Daniel McFadden) Chapter 74. Elsevier.

[82] Jepsen, Lisa and Christopher Jepsen (2002). "An Empirical Analysis of the Matching Patterns of Same-Sex and Opposite-Sex Couples." *Demography*. 39(3): 435-453.

[83] Kang, Hyunjae and Steven Stern (2021). "Alternative Methods to Test for Positive Assortative Mating." Unpublished manuscript.

[84] Kodde, David and Franz Palm (1986). "Wald Criteria for Jointly Testing Equality and Inequality Restrictions." *Econometrica*. 54(5): 1243-1248.

[85] Konrad, Kai and Kjell Lommerud (2000). "The Bargaining Family Revisited." *Canadian Journal of Economics*. 33(2): 471-487.

[86] Kudo, Akio (1962). "A Multivariate Analogue of the One-Sided Test." *Biometrika*. 50(3): 403-418.

[87] Lalonde, Robert (1986). "Evaluating the Econometric Evaluations of Training Programs with Experimental Data." *American Economic Review*. 76(4): 604-620.

[88] Lee, D. (2009). "Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects." *Review of Economic Studies*. 76(3): 1071–1102.

[89] Lee, Lung-Fei (1982). "Some Approaches to the Correction of Selectivity Bias." *Review of Economic Studies*. 49(3): 355-372.

[90] Lundberg, Shelley, Robert Pollak, and Terence Wales (1997). "Do Husbands and Wives Pool Resources: Evidence from the UK Child Benefit." *Journal of Human Resources*. 32(3): 463-480.

[91] Lundberg, Shelley, Richard Startz, and Steven Stillman (2003). "The Retirement Consumption Puzzle: A Marital Bargaining Approach." *Journal of Public Economics.* 87(5-6): 1119-1218.

[92] Macurdy, Thomas (1982). "The Use of Time Series Processes to Model the Error Structure of Earnings in a Longitudinal Data Analysis." *Journal of Econometrics.* 18(1): 82-114.

[93] Maddala, G. (1983). *Limited-Dependent and Qualitative Variables in Econometrics.* Cambridge University Press, Cambridge.

[94] Mare, Robert (1991). "Five Decades of Educational Assortative Mating." *American Sociological Review.* 56(1): 15-32.

[95] Martins, Maria (2001). "Parametric and Semiparametric Estimation of Sample Selection Models: An Empirical Application to the Female Labour Force in Portugal." *Journal of Applied Econometrics.* 16(1): 23–39.

[96] Mazzocco, Maurizio, Claudia Ruiz, and Shintaro Yamaguchi (2007). "Labor Supply, Wealth Dynamics and Marriage Decisions." Unpublished manuscript.

[97] Meghir, Costas and Luigi Pistaferri (2004). "Income Variance Dynamics and Heterogeneity." *Econometrica.* 72(1): 1-32.

[98] Mroz, Thomas (1987). "The Sensitivity of an Empirical Model of Married Women's Hours of Work to Economic and Statistical Assumptions." *Econometrica.* 55(4): 765-799.

[99] Mulligan, Casey and Yona Rubinstein (2008). "Selection, Investment, and Women's Relative Wages over Time." *Quarterly Journal of Economics.* 123(3): 1061-1110.

[100] Nakosteen, Robert and Michael Zimmer (2001). "Spousal Selection and Earnings: Evidence of Marital Sorting." *Economic Inquiry.* 39(2): 201-213.

[101] Newey, Whitney, James Powell, and James Walker (1990). "Semiparametric Estimation of Selection Models: Some Empirical Results." *American Economic Review.* 80(2): 324-328.

[102] Olsen, Randall (1982). "Distributional Tests for Selectivity Bias and a More Robust Likelihood Estimator." *International Economic Review.* 23(1): 223-240.

[103] Pencavel, John (1998). "Assortative Mating by Schooling and the Work Behavior of Wives and Husbands." *American Economic Review.* 88(2): 326-329.

[104] Phipps, Shelley and Peter Burton (1998). "What's Mine is Yours? The Influence of Male and Female Incomes on Patterns of Household Expenditure." *Economica.* 65(26): 599-613.

[105] Pollak, Robert (2005). "Bargaining Power in Marriage: Earnings, Wage Rates and Household Production." NBER Working Papers 11239, National Bureau of Economic Research.

[106] Powell, James, James Stock, and Thomas Stoker (1989). "Semiparametric Estimation of Index Coefficients." *Econometrica.* 57(6): 1403-1430.

[107] Rosenzweig, Mark and Kenneth Wolpin (1980). "Life-cycle Labor Supply and Fertility: Causal Inference from Household Models." *Journal of Political Economy.* 88(2): 328-348.

[108] Santos, Cezar and David Weiss (2016). "Why Not Settle Down Already?" A Quantitative Analysis of the Delay in Marriage." *International Economic Review.* 57(2): 425–452.

[109] Schucany,William, William Parr, and John Boyer (1978). "Correlation Structure in Farlie-Gumbel-Morgenstern Distributions." *Biometrika.* 65(3): 650–653.

[110] Schultz, T. Paul (1990). "Testing the Neoclassical Model of Family Labor Supply and Fertility." *Journal of Human Resources.* 25(4): 599-634.

[111] Schultz, T. Paul (2010). "Population and Health Policies." *Handbook of Development Economics, Vol. 5.* (eds.) D. Rodrik and M. Rosenzweig. Elsevier, North-Holland: Amsterdam. 4785-4881.

[112] Schwartz, Christine (2010). "Earnings Inequality and the Changing Association between Spouses' Earnings." *American Journal of Sociology.* 115(5): 1524-1557.

[113] Schwartz, Christine and Robert Mare (2005). "Trends in Educational Assortative Marriage from 1940 to 2003." *Demography.* 42(4): 621–646.

[114] Siow, Aloysius (2015). "Testing Becker's Theory of Positive Assortative Matching." *Journal of Labor Economics.* 33(2): 409-441.

[115] Smith, James and Michael Ward (1989). "Women in the Labor Market and in the Family." *Journal of Economic Perspectives.* 3: 9–23.

[116] Smits, Jeroen, Wout Ultee, and Jan Lammers (1998). "Educational Homogamy in 65 Countries: An Explanation of Differences in Openness Using Country-Level Explanatory Variables." *American Sociological Review.* 63(2): 264–285.

[117] Stern, Steven (1996). "Semiparametric Estimates of the Supply and Demand Effects of Disability on Labor Force Participation." *Journal of Econometrics.* 71(1-2): 49-70.

[118] Suen, Wing and Hon-Kwong Lui (1999). "A Direct Test of the Efficient Marriage Market Hypothesis." *Economic Inquiry.* 37(1): 29–46.

[119] Thomas, Duncan (1990). "Intra-Household Resource Allocation: An Inferential Approach." *Journal of Human Resources.* 25(4): 635-664.

[120] Thomas, Duncan (1994). "Like Father, Like Son, Like Mother, Like Daughter: Parental Resources and Child Height." *Journal of Human Resources.* 29(4): 950-988.

[121] Van der Klaauw, Wilbert (1996). "Female Labour Supply and Marital Status Decision: A Life-Cycle Model." *Review of Economic Studies.* 63(2): 199-235.

[122] Zabel, Jeffrey (1993). "The Relationship between Hours of Work and Labor Force Participation in Four Models of Labor Supply Behavior." *Journal of Labor Economics.* 11(2): 387–416.

[123] Zhang, Junsen and Pak-Wai Liu (2003). "Testing Becker's Prediction on Assortative Mating on Spouse's Wages." *Journal of Human Resources.* 38(1): 99-110.