A Distributional Analysis of the Return to Marriage

Esfandiar Maasoumi
Southern Methodist University

Daniel L. Millimet*
Southern Methodist University

Dipanwita Sarkar
Southern Methodist University

December 2004

Abstract

The phenomenon that married men earn a higher wage on average than unmarried men, the so-called marriage premium, is rather well established. However, the robustness of the marriage premium across the wage distribution and the underlying cause of the marriage premium are less well known. Focusing on the entire wage distribution, we employ recently developed nonparametric tests for stochastic dominance. Our findings question the current conception of the marriage premium, calling instead for the introduction of a broader second order concept. This broader notion is consistent with the fact that (i) the (unconditional) marriage premium exists across the entire distribution using PSID and CPS data, (ii) there is some evidence that the marriage premium may not be uniform across the wage distribution, implying a need to incorporate wage 'dispersion' into the notion of the marriage premium, (iii) the majority of the premium is explained by selection, but there is a small role for 'causal' explanations, and (iv) conclusions regarding the uniformity of the marriage premium across the distribution, but not the portion of the premium attributable to causal-based explanations, are sensitive to the removal of time invariant and time-varying unobservables which are correlated with marital status and labor market performance and the manner by which potentialy confounding observables are controlled.

JEL Classifications: C14, C33, D31, J12
Key words: Marriage premium, stochastic dominance, difference-in-differences, instrumental variable, nonparametric, selection

*Corresponding author: Daniel Millimet, Department of Economics, Southern Methodist University, Box 750496, Dallas, TX 75275-0496, USA; Email: millimet@mail.smu.edu; Tel. (214) 768-3269; Fax: (214) 768-1821.

1 Introduction

Married men, *on average*, earn more than single men in the labor market even after controlling for observable attributes; a fact commonly referred to as the *marriage premium*. But is an average measure of 'the premium' consistent with a uniform ranking of these two groups across reasonable classes of utility functions and wage distributions? If some individuals do not benefit, or benefit to different degrees, how does one assess the overall gain or loss? The current conception of 'the premium' is so well-established that Loh (1996, p. 566) states: "Virtually all cross-sectional wage studies find that currently married men typically earn a higher wage rate than their unmarried counterparts in the labor market." Cornwell and Rupert (1997, p. 285) similarly note: "Married men earn more than unmarried men. This fact is unassailable and is robust across data sets and over time." While the existence of a 'marriage premium' may not be controversial, the *extent*, *uniformity* across the wage distribution, and the *underlying sources* of 'the premium' remain heavily disputed. Cornwell and Rupert (1997, p. 285) continue: "While there is compelling evidence that married men earn more than unmarried men, the source of this premium remains unsettled." More recently, Stratton (2002, p. 199) states: "Research has failed as yet to reach a consensus regarding the nature of these differentials."

Given this lack of consensus, labor economists continue to seek the underlying sources of 'the premium' – typically on the order of a 10 to 40 percent *average* wage differential – for three main reasons. First, knowledge of the premium's source(s) contributes to our understanding of the general process of wage determination. Second, understanding the marriage premium furthers our knowledge of the role played by gender in the labor market as the premium constitutes about one-third of the entire gender wage gap (Korenman and Neumark 1991). Finally, if the marriage premium reflects true productivity differences, then changes in marital trends in the US and elsewhere may foreshadow changes in future productivity.

Several hypotheses for the marriage premium have been put forth, and these may be loosely classified into three categories: (i) causal explanations, (ii) unobserved covariates/selection explanations, and (iii) reverse causation explanations. Hypotheses that are based on a causal effect of marriage on wages center predominately on the productivity-enhancing impact of marriage on men arising from intrahousehold specialization. Under this argument – originating in the work by Becker (1973, 1974, 1991) – marriage enables men to specialize in labor market activities (due to their comparative advantage in market work), while women specialize in home production. A second, causal explanation attributes the marriage premium to employer discrimination in favor of married men.[1,2] Unobserved covariates/selection explanations note the

[1] According to this reasoning, employers may view married men as more reliable, more honest, less mobile, etc.

[2] An additional causal explanation rests on the theory of compensating differentials, asserting that married men tend to forego non-monetary work benefits (e.g., flexible hours) for greater monetary compensation. Duncan and Holmund (1983) and

selective nature of marriage, and focus on the possibility that marriage may be correlated with unobservable attributes that are valued in both the labor and marriage markets (e.g., interpersonal skills, integrity, reliability, work ethic, etc.). Finally, arguments based on reverse causation center on the possibility that single women may seek out high-earning men as potential partners (e.g., Ginther and Zavodny 2001).

In previous research, several studies document evidence in favor of a small, productitivity-enhancing effect of marriage after controlling for self-selection, typically by employing parametric, fixed effects panel methods (Korenman and Neumark 1991; Daniel 1995; Stratton 2002).[3] Antonovics and Town (2004) find a large, causal effect of marriage – and no evidence favoring the selection hypothesis – upon estimating a fixed effects model using data on monozygotic twins. In further support of the specialization hypothesis, Jacobsen and Rayack (1996), Gray (1997), Chun and Lee (2001), among others, find that the marriage premium declines with wife's labor supply.[4] Moreover, several studies have documented a decline in the marriage premium over time (Blackburn and Korenman 1994; Loh 1996; Gray 1997; Cohen 2002), which may be additional evidence in favor of the specialization hypothesis given the rise in female labor force participation in the US. In terms of the employer discrimination hypothesis, Jacobsen and Rayack (1996) and Loh (1996) test for the existence of the marriage premium amongst self-employed workers, and find an insignificant or even negative marriage premium, perhaps lending some support to the hypothesis.[5] Finally, Ginther and Zavodny (2001) use 'shotgun weddings' to circumvent the selection issue, finding that selection accounts for less than ten percent of the marriage premium.[6] Conversely, several researchers conclude that self-selection is the primary explanation for the existence of the marriage premium (Nakosteen and Zimmer 1987; Cornwell and Rupert 1997), particularly in the 1990s (Gray 1997).[7] Finally, little systematic evidence exists *per se* to support the reverse causation explanation. However, consonant with Becker (1976),

Hersch (1991) find little support for such claims.

[3] Fixed effects methods control for selection based on time invariant wage levels, but not selection into marriage based on wage growth.

[4] Loh (1996), on the other hand, fails to uncover evidence of a consistent relationship between wife's labor supply and husband's wages. Hersch and Stratton (2000) find that while the use of individual fixed effects does not substantially reduce the marriage premium, controlling for time spent in household production has little impact on the magnitude of the marriage premium. The authors, therefore, conclude that selection plays a minimal role, but reject the specialization as the underlying source of the causal relationship.

[5] As indicated by the authors, such conclusions should be interpreted cautiously, given the difficulties that arise in econometric studies of the self-employed (e.g., measurement error in wages, self-selection, pooling self-employed and non-self-employed workers, etc.).

[6] A 'shotgun wedding' is defined in Ginther and Zavodny (2001) as a marriage that is followed by the birth of a child within the subsequent seven months.

[7] Gray (1997) finds that the marriage premium represents a productivity effect in the late 1970s, but is attributable to selection in the early 1990s.

Cornwell and Rupert (1997) find that men who are 'to-be-married,' on average, earn wages comparable to married men, and Nakosteen and Zimmer (1997) and Ginther and Zavodny (2001) note that earnings are positively correlated with the likelihood of marriage.

In this paper, we add to this literature in two important ways. First, we revisit the claim that the marriage premium is robust by assessing the impact of marriage on the *distribution* of wages. As alluded to above, every empirical analysis of the average marriage premium (to our knowledge) utilizes parametric regression analysis and focuses on only the (conditional) mean of the wage distribution. Such analyses ignore a great deal of *available information* and lacks a broadly accepted *welfare underpinning*. Specifically, the *implicit* welfare functions in these types of analyses are linear in wages, neglecting other characteristics of the wage distribution as well as the potentially heterogeneous effects of marriage across subgroups of men. The gain from summary measures, such as a regression coefficient, is that they produce complete, strong rankings of male wages across marital states that are easily interpretable. More sophisticated techniques are needed, however, for weaker, yet uniform assessment of wage effects arising from marriage over large classes of welfare functions. Empirical examination of such *uniform rankings*, based on the notion of stochastic dominance (SD), is the primary goal of the current paper.[8] SD analysis provides welfare comparisons across a wide range of criterion and is an important companion to standard regression analysis.[9] As a result, inferring a dominance relation implies that comparisons based on multiple specific indices are either unnecessary, or only helpful for quantification and cardinal monitoring of change. On the other hand, the inability to find a dominance relation is equally valuable and informative, indicating that any (implicit) welfare ordering based on a particular measure (such as the average) is subjective and will not apply to all segments of the population; different measures can yield different substantive conclusions, implying that the 'marriage premium' may not be a premium for all! Moreover, recently developed nonparametric tests enable us to assess such relations to a degree of statistical certainty. Second, we revisit and distinguish the underlying explanations of the return to marriage within this broader *distributional* definition of the

[8] As we shall refer to the concept of *uniform rankings* often, we wish to be explicit. Uniform rankings imply that all social welfare functions within a particular class would rank distributions (from highest to lowest welfare) in the identical order.

[9] The richness of the SD analysis has led to their growing application. For example, Maasoumi and Millimet (2003) examine changes in US pollution distributions over time and across regions at a point in time. Maasoumi and Heshmati (2000) analyze changes in the Swedish income distribution over time as well as across different population subgroups. Fisher et al. (1998) compare the distribution of returns to different length US Treasury Bills. Particularly relevant to the analysis at hand are previous applications of SD to the analysis of treatment effects. For instance, Amin et al. (2003) analyze the effect of a micro-credit program in Bangladesh on the distribution of consumption of participants versus non-participants. Abadie (2002) analyzes the impact of veteran status on the distribution of civilian earnings. Bishop et al. (2000) compare the distribution of nutrition levels across populations exposed to two different types of food stamp programs. Anderson (1996) compares pre- and post-tax income distributions in Canada over several years.

marriage premium.

To perform the analysis, we begin by using panel data from the Current Population Survey (CPS) for 1992 – 2001. We offer the following comparisons of *unconditional* distributions: (i) single versus 'to-be-married' men (men who marry within the next year), (ii) single versus newly married men, and (iii) the distribution of wage changes for single versus newly married men. Next, we re-examine the same pairwise comparisons after purging wages of a host of observable covariates that may be correlated with both marital status and/or labor market performance. This two-part strategy enables us to (i) support (or refute) the existence of a *uniform* marriage premium, and (ii) comment on the underlying causes of our broader definition of the marriage premium (assuming it exists). Specifically, if the marriage premium represents a causal relationship (with marriage affecting wages), then there should be no difference in the distribution of wages between single and 'to-be-married' men. But the distribution of married men should 'dominate' that of single men. On the other hand, a 'dominant' distribution of wages amongst 'to-be-married' men would suggest a role for both selection and reverse causation explanations. If there is some validity to all the proposed explanations, then we might observe a modest disparity in the distribution of wages favoring 'to-be-married' men (versus singles), followed by an even greater disparity after marriage. Thus, our SD tests can help assess the relative role of causal versus correlation-based explanations of the marriage premium *across the entire wage distribution.*

The above strategy separates the causal and selection components of the marriage premium through comparisons of wage changes before and after marriage, thereby eliminating time invariant unobservables as in fixed effects methods. An alternative strategy is to use an instrumental variable (IV) for marital status. The advantage of employing an IV strategy is that it controls for time-varying unobservables that may be correlated with both marital status and wages. Abadie (2002) shows how one can utilize a binary instrument to perform tests for SD. Thus, for comparison, we use parental marital status as an instrument for one's own marital status and re-examine the return to marriage. However, since information on parent's marital status is unavailable in the CPS, for this analysis we switch to the Panel Study of Income Dynamics (PSID), and utilize available data from 1994.

The results are striking. In particular, we reach four conclusions. First, the marriage premium persists even at the distributional level in the PSID and CPS data in the 1990s. Second, there is some evidence that the return to marriage may not be uniform across the wage distribution; in particular, the gains from marriage may be larger for those in the *bottom third* of the wage distribution. This implies that uniform welfare rankings – rankings robust to the choice of specific preference function – may only be obtained if one broadens the concept of the marriage premium to incorporate *wage dispersion.* Third, the majority of the premium is explained by selection, but there is a small role for causal-based explanations.

Finally, conclusions regarding the uniformity of the return to marriage across the wage distribution, but not the portion of the premium attributable to causal-based explanations, are affected by the removal of time invariant and time-varying unobservables correlated with marital status and labor market performance and the manner by which potentially confounding observables are controlled. This implies that typical fixed effects methods alone – which are prominent in the marriage premium literature – may be insufficient to give a complete picture of the returns to marriage. The remainder of the paper is organized as follows. Section 2 details the difference-in-differences SD methodology and the CPS data, and discusses the corresponding results. Section 3 presents the IV SD methodology and PSID data, and discusses the results. Section 4 concludes.

2 Difference-in-Differences Methodology

2.1 Test Statistics

Several tests for SD have been proposed in the literature; Maasoumi and Heshmati (2000) provide a brief review of the development of alternative tests. The approach herein is based on a generalized Kolmogorov-Smirnov test. To begin, let X and Y denote two wage variables for individuals in two different marital status categories (e.g., X (Y) might denote wages of single (married) men). $\{x_i\}_{i=1}^{N}$ is a vector of N strictly stationary, α-mixing, possibly dependent observations of X; $\{y_i\}_{i=1}^{M}$ is an analogous vector of realizations of Y. In the spirit of the historical development of such two-sample tests, $\{x_i\}_{i=1}^{N}$ and $\{y_i\}_{i=1}^{M}$ each constitute one sample. Thus, we refer to dependence between x_i and x_j, $i \neq j$, as *within-sample dependence* (similarly for observations of Y), and dependence between X and Y as *between-sample dependence.*

With nothing further to assume than general von Neumann-Morgenstern conditions, let $\mathcal{U}_1$ denote the class of (increasing) utility functions u such that utility is increasing in wages (i.e. $u' \geq 0$), and $\mathcal{U}_2$ the class of social welfare functions in $\mathcal{U}_1$ such that $u'' \leq 0$ (i.e. concavity). Concavity represents an aversion to higher dispersion of wages across individuals; a high concentration of earnings is undesirable. Let $F(x)$ and $G(y)$ represent the cumulative density functions (CDF) of X and Y, respectively, which are assumed to be continuous and differentiable.

Under this notation, X First Order Stochastically Dominates Y (denoted X FSD Y) *iff* $\mathrm{E}[u(X)] \geq \mathrm{E}[u(Y)]$ for all $u \in \mathcal{U}_1$, with strict inequality for some u. Equivalently,

$$F(z) \leq G(z) \qquad \forall z \in \mathcal{Z}, \text{ with strict inequality for some } z. \tag{1}$$

where $\mathcal{Z}$ denotes the union of the supports of X and Y. If X FSD Y, then the expected welfare from X is at least as great as that from Y for all increasing welfare functions, with strict inequality holding for some utility function(s) in the class.

The distribution of X Second Order Stochastically Dominates Y (denoted as X SSD Y) *iff* $\mathrm{E}[u(X)] \geq \mathrm{E}[u(Y)]$ for all $u \in \mathcal{U}_2$, with strict inequality for some u. Equivalently,

$$\int_{-\infty}^{z} F(v)dv \leq \int_{-\infty}^{z} G(v)dv \qquad \forall z \in \mathcal{Z}, \text{ with strict inequality for some } z. \tag{2}$$

If X SSD Y, then the expected social welfare from X is at least as great as that from Y for all increasing and concave utility functions in the class $\mathcal{U}_2$, with strict inequality holding for some utility function(s) in the class. Evidently, FSD implies SSD and higher orders. Higher order dominance rankings are based on more restricted classes of utility functions which reflect aversion to asymmetry, kurtosis, and higher order moments.

Now define the following generalizations of the Kolmogorov-Smirnov test criteria:

$$d = \sqrt{\frac{NM}{N+M}} \min \sup_{z \in \mathcal{Z}} [F(z) - G(z)] \tag{3}$$

$$s = \sqrt{\frac{NM}{N+M}} \min \sup_{z \in \mathcal{Z}} \int_{-\infty}^{z} [F(u) - G(u)]\, du \tag{4}$$

where min is taken over $F - G$ and $G - F$, in effect performing two tests in order to leave no ambiguity between the 'equal' and 'unrankable' cases. The null hypothesis of FSD (SSD) implies $d \leq 0$ ($s \leq 0$).

Define the empirical CDF for X as

$$\widehat{F}_N(x) = \frac{1}{N} \sum_{i=1}^{N} \mathrm{I}\,(X \leq x)$$

where $\mathrm{I}(\cdot)$ is an indicator function. $\widehat{G}_M(y)$ is defined similarly for Y. Our nonparametric tests for FSD and SSD are based on the empirical counterparts of d and s using the empirical CDFs. Specifically, the test for FSD requires:

(i) computing the values of $\widehat{F}(z_q)$ and $\widehat{G}(z_q)$ for z_q, $q = 1, ..., Q$, where Q denotes the number of points in the support $\mathcal{Z}$ that are utilized ($Q = 500$ in the application),

(ii) computing the differences $d_1(z_q) = \widehat{F}(z_q) - \widehat{G}(z_q)$ and $d_2(z_q) = \widehat{G}(z_q) - \widehat{F}(z_q)$, and

(iii) finding $\widehat{d} = \sqrt{\frac{NM}{N+M}} \min\{\max\{d_1\}, \max\{d_2\}\}$.[10]

[10] In the empirical implementation, the support points are chosen to be equally-spaced, beginning at the first percentile and ending at the 99^{th} percentile of the empirical support, $\mathcal{Z}$. This process focuses attention away from extreme outliers. In practice, this 'trimming' process is of little consequence.

If $\widehat{d} \leq 0$ (to a degree of statistical certainty), then the null hypothesis of FSD is not rejected. Furthermore, if $\widehat{d} \leq 0$ and $\max\{d_1\} < 0$, then X FSD Y as the value of the CDF for distribution X is at most as large as the corresponding value for distribution Y at all z_q, $q = 1, ..., Q$. On the other hand, if $\widehat{d} \leq 0$ and $\max\{d_2\} < 0$, then Y FSD X.[11] The analogous test for SSD requires the following additional steps:

(i) calculating the sums $s_{1q} = \sum_{j=1}^{q} d_1(z_j)$ and $s_{2q} = \sum_{j=1}^{q} d_2(z_j)$, $q = 1, ..., Q$, and

(ii) finding $\widehat{s} = \sqrt{\frac{NM}{N+M}} \min\{\max\{s_{1q}\}, \max\{s_{2q}\}\}$.

If $\widehat{s} \leq 0$ (to a degree of statistical certainty), then the null hypothesis of SSD is not rejected. Moreover, if $\widehat{s} \leq 0$ and $\max\{s_{1q}\} < 0$, then X SSD Y as the cumulative value of the CDF (or integrated CDF) for distribution Y exceeds the corresponding value for distribution X at all z_q; otherwise, if $\max\{s_{2q}\} < 0$, then Y SSD X.

Given the two-period panel structure of the data (discussed below), we classify the sample into two marital status groups: (i) men who are single in both periods of the survey (henceforth referred to as 'single') and (ii) men who are single in the first survey period, and are married in the second survey period (henceforth referred to as 'to-be-married' ('married') in the first (second) period).[12] This data structure lends itself to three distributional comparisons: (i) single men versus to-be-married men observed in the first period, $t = 1$; (ii) single men versus married men observed in the second period, $t = 2$; and, (iii) the change in wages from period one to two for single men versus to-be-married/married men. The first two comparisons, while insightful, do not lend themselves to causal conclusions if there are individual attributes correlated with both marital status and wages. The third approach extends the standard difference-in-differences approach used to analyze 'average' effects to distributional comparisons since it purges wages of all time invariant, individual-specific characteristics (observable or unobservable).

The differencing of unconditional wages, however, only controls for time invariant attributes that may be correlated with both wages and marital status. Since to-be-married and married men may possess time-varying characteristics that are also valued in the marriage and labor market, our next step is to compare X and Y computed as *residual* wage distributions. To obtain these residuals, we control for a host of observable attributes that may generate a spurious correlation between marital status and wages and conduct dominance tests on the distributions of wages purged of the effects of these attributes. The conditioning covariates (discussed below) represent individual characteristics (such as investment in human

[11] If $\widehat{d} = \max\{d_1\} = \max\{d_2\} = 0$, then the (estimated) distributions of X and Y are identical.

[12] Note, we do not utilize information on men who are married (single) in the first (second) period. This contrasts with some parametric fixed effects studies that utilize information on transitions into and out of marriage to identify the marriage premium. In addition, the 'to-be-married' sample contains never married individuals; thus, we focus only on first marriages.

capital, age and race), institutional factors (such as union status and occupation), family background variables (such as the presence of children and other dependents), housing variables (such as whether one rents or owns), as well as locational factors pertaining to residence in a particular region (such as urban and regional location).

To proceed, we estimate separate wage functions for males in each marital status category and for each survey period, obtain the intercept–adjusted residuals, and perform tests for SD on these residuals.[13] Specifically, in the first-stage, we estimate via Ordinary Least Squares (OLS)

$$\ln(w_{it}^k) = \alpha_t^k + h_{it}^k \beta_t^k + \widetilde{\epsilon}_{it}^k, \quad t = 1, 2 \ ; \quad k = \text{single, (to-be-)married} \tag{5}$$

where w_{it}^k is the hourly wage income of individual i in the survey round t in marital status category k, h is a vector of individual, family, institutional and geographical attributes, and a set of year dummies (discussed below), and ϵ is the error term.[14] In the second-stage, we analyze the distributions of the intercept–adjusted residuals, $\widehat{\epsilon}_{it}^k \equiv \widehat{\alpha}_t^k + \widehat{\widetilde{\epsilon}}_{it}^k$, utilizing the same set of three comparisons discussed previously. In other words, we compare: (i) single men versus to-be-married men observed in $t = 1$; (ii) single men versus married men observed in $t = 2$; and, (iii) the change in $\widehat{\epsilon}_{it}^k$ from round one to two for single men versus to-be-married/married men.[15]

At this point, a few comments are warranted. First, for comparisons based on (i) and (ii) to yield meaningful causal inference regarding the impact of marital status on wages, (5) must not omit any variables that are correlated with both marital status and wages; thus, these are *selection on observable* estimators. For comparisons based on (iii) to be meaningful from a causal perspective, (5) is only required to not omit any time-varying variables that are correlated with both marital status and wages.[16]

Second, the intercept-adjusted residuals, $\widehat{\epsilon}_{it}^k$, reflect wages net of a linear projection of observable characteristics evaluated at the marital state–specific returns, β_t^k. Since this approach nets out wage differences due to observables as well as the marital state-specific returns to such observables, we refer to these tests as being based on 'Partial Residuals' (PR). As an alternative approach, we also conduct tests

[13] The intercepts are included as part of the residuals, otherwise the conditional distributions will all be mean zero. This is also done since we do not wish to claim the models are perfectly specified.

[14] A critical aspect of this procedure to be noted is that dummies for marital status are omitted from (5), thereby allowing the error term to capture the residual effect of marriage not captured by the included regressors.

[15] In constructing the change in the intercept-adjusted residuals across the two survey periods for each individual in marital group k, $\triangle\widehat{\epsilon}_i^k = \left(\widehat{\alpha}_2^k + \widehat{\widetilde{\epsilon}}_{i2}^k\right) - \left(\widehat{\alpha}_1^k + \widehat{\widetilde{\epsilon}}_{i1}^k\right)$, we allow the first-stage coefficients to vary by t and k. In other words, $\triangle\widehat{\epsilon}_i^k = \left(\ln(w_{i2}^k) - h_{i2}^k \widehat{\beta}_2^k\right) - \left(\ln(w_{i1}^k) - h_{i1}^k \widehat{\beta}_1^k\right)$ for $k =$ single, (to-be-)married.

[16] In addition, selection into marriage based on wage growth will lead to false inference concerning the nature of the marriage premium. However, little evidence exists to support such selection. For example, Gray (1997) finds at best weak evidence that wage growth is positively related to the probability of marriage. See also Korenman and Neumark (1991).

based on the 'Full Residuals' (FR), where we denote the full residuals as inclusive of differences in the return to observables. This approach incorporates the differences in the returns into the intercept-adjusted residual distributions, analogous to the standard Oaxaca-Blinder decomposition. To show how this is done, we re-write the first-stage regression (5) for $k =$ (to-be-)married as

$$\begin{aligned} \ln(w_{it}^m) &= \alpha_t^m + h_{it}^m \beta_t^m + \widetilde{\epsilon}_{it}^m \\ &= \alpha_t^m + h_{it}^m \beta_t^m + \widetilde{\epsilon}_{it}^m + (h_{it}^m \beta_t^s - h_{it}^m \beta_t^s) \\ &= \alpha_t^m + h_{it}^m \beta_t^s + h_{it}^m (\beta_t^m - \beta_t^s) + \widetilde{\epsilon}_{it}^m, \qquad t = 1, 2 \end{aligned} \tag{6}$$

where the single group is implicitly treated as the 'dominant' category (Neuman and Oaxaca 2003). Consequently, we amend the residual tests to compare the previous intercept-adjusted residual distribution of $\widehat{\epsilon}_{it}^s$ with $\widehat{\epsilon}_{it}^{k,ob} \equiv \left(\widehat{\alpha}_t^k + h_{it}^k(\widehat{\beta}_t^k - \widehat{\beta}_t^s) + \widehat{\widetilde{\epsilon}}_{it}^k\right)$, $k =$ to-be-married or married (depending on if the comparison is for $t = 1$ or 2).[17]

In order to help with the interpretation of the residual dominance results, we also discuss the results from the standard Oaxaca-Blinder parametric decompositions. Specifically, the mean wage differential between the two groups, single and (to-be-)married, in period t, may be expressed as

$$\overline{\ln(w_t^m)} - \overline{\ln(w_t^s)} = \underbrace{(\alpha_t^m - \alpha_t^s)}_{U} + \underbrace{(\overline{h}_t^m - \overline{h}_t^s)\beta_t^s}_{E} + \underbrace{\overline{h}_t^m(\beta_t^m - \beta_t^s)}_{C} \tag{7}$$

where the single group is treated as the 'dominant' category. If the difference in returns (term C) in (7) is large in absolute value, then the two sets of residual tests may be expected to yield disparate results. Moreover, if the three sets of results (two sets of residual tests and one set of unconditional tests) offer different inferences, one may infer a 'significant' association between marital status, the distribution of wages, the set of conditioning variables, and the returns to the conditioning variables.

2.2 Alternative Methods of Controlling for Covariates

The use of regression analysis – parametric or otherwise – to purge the effects of potentially confounding observable covariates is a very common projection method in empirical sciences. It may be criticized, however, on the grounds that it may only remove the *mean* effects from the conditional mean. There are no simple solutions to this problem, as it raises all of the identification issues that are still under examination in the treatment effects literature. Anything short of knowledge of the appropriate counterfactual distribution

[17] Under the Oaxaca-Blinder-type FR approach, the tests based on the first-differenced residuals are based on the distributions of $\triangle\widehat{\epsilon}_i^s$ and $\triangle\widehat{\epsilon}_i^{m,ob}$, where $\triangle\widehat{\epsilon}_i^s = \left(\widehat{\alpha}_2^s + \widehat{\widetilde{\epsilon}}_{i2}^s\right) - \left(\widehat{\alpha}_1^s + \widehat{\widetilde{\epsilon}}_{i1}^s\right)$ and $\triangle\widehat{\epsilon}_i^{m,ob} = \left(\widehat{\alpha}_2^m + h_{i2}^m(\widehat{\beta}_2^m - \widehat{\beta}_2^s) + \widehat{\widetilde{\epsilon}}_{i2}^m\right) - \left(\widehat{\alpha}_1^m + h_{i1}^m(\widehat{\beta}_1^m - \widehat{\beta}_1^s) + \widehat{\widetilde{\epsilon}}_{i1}^m\right)$.

is less than ideal. Nonetheless, we assess the robustness of our results using two alternative strategies to control for observables. First, we adopt the now familiar *propensity score matching* (PSM) technique and conduct our dominance tests on pairs of observations matched on the basis of the probability of marriage. Specifically, we estimate the propensity score via a probit model using the same set of individual, family, occupational and location variables as in (5). We then match each individual from the married sample to the single individual with closest propensity score (so-called single nearest neighbor matching).[18] However, unlike the usual matching literature on evaluating treatment effects (see, e.g., Heckman et al. 1997) which is focused on the difference in average outcomes, we compare the CDFs of the outcome. The identifying assumption is that treatment participation (marriage) and outcomes are independent conditional on the vector of observable attributes, commonly known as the conditional independence assumption (Heckman and Robb 1985). This is identical to the selection on observables assumption discussed above. One benefit of this approach is that matching on the propensity score implicitly controls nonparametrically for interactions between the covariates (Blundell et al. 2002; Bratberg et al. 2002). Also, one uses only the observations from the sample of single men deemed most 'similar' to the sample of married men.

An alternative matching approach exemplified in this paper for the first time (to our knowledge), looks at the notion of 'similarity' more closely. Adopting aggregation ideas first proposed in Maasoumi (1986), we obtain a *composite characteristic* of each individual based on any set of desired and observed attributes. The criterion that guides the choice of the *composite index* is consistent with our central philosophy of emphasizing entire distributions. We will select a composite index which will have a distribution that is most similar to the distributions of the many covariates which it aggregates. Natural criteria for similarity of entire distributions are available in the field of information theory. Based on Generalized Entropy as a measure of expected information divergence, Maasoumi (1986) derived such composite indices as $S_i \propto \left(\sum_{j=1}^{M} \delta_j h_{ij}^{-\beta}\right)^{-1/\beta}$, where $\delta_j = \alpha_j / \sum_j \alpha_j$ and h is the vector of controls from (5). Rewriting $\beta = (1/\sigma) - 1$, where σ denotes the constant elasticity of substitution (CES), S_i provides a positive interpretation of many popular utility functions, namely the CES, the Cobb-Douglas ($\beta = 0$) and the linear ($\beta = -1$). Assuming weights proportional to the number of covariates (i.e. $\delta_j = 1/M$), we match to-be-married (married) men to single men in the first (second) period on the basis of nearest S_i values utilizing different values for β, where $\beta \in [-1, 0)$.[19] Interestingly, the propensity score can also be viewed as a *composite characteristic* of individuals and as a special case of our more general functionals. Employing

[18] To be clear, when comparing single men versus to-be-married (married) men observed in $t = 1$ ($t = 2$), the propensity score is obtained by fitting a probit model using the sample of observations from the first (second) period only. When comparing the change in wages from the first to the seond period, the propensity score is estimated using the sample of observations from the second period only and, after matching, the CDFs of wage changes from period one to period two are compared.

[19] In other words, as with the propensity score, we calculate S_i seperately for each period.

the same covariates, PSM fixes $\beta = -1$, and employs likelihood based estimates of the unknown coefficients δ_j. These parameter values, as well as others (OLS, etc.), are 'optimal' relative to the estimation criterion imposed. However, such criteria may or may not be 'optimal' when it comes to the matching of individuals by their entire distributions! Our analysis exposes the subjectivity of the many choices being made, and suggests the need for sensitivity analysis with respect to parameter values and functional forms.[20]

2.3 Inference

It is well known that the asymptotic distribution of our test statistics depend on the unknown underlying distributions. McFadden (1989) and Klecan et al. (1991) examine the asymptotic distribution and a Monte Carlo implementation of these nonparametric tests for FSD and SSD. McFadden (1989) assumes *iid* observations and independent variates. Klecan et al. (1991) allows for general weak (within sample) dependence over time between observations, and only general exchangeability between the variables (distributions) being ranked. Barrett and Donald (2003) also assume *iid* observations and independent variates in deriving a supremum version of the tests. Linton et al. (2003) utilize subsampling and recentered bootstrap techniques and allow for general dependence both between the variables and within samples.

In the analysis below, we first approximate the empirical distribution of the test statistics using *simple bootstrap* techniques (Maasoumi and Heshmati 2000; Maasoumi and Millimet 2003). Specifying the null in terms of (weak) inequality in a particular direction implies that for any pairwise comparison between distributions, dominance relations in both directions must be tested in order to avoid ambiguity between unrankable and equal distributions.

To evaluate the null $\text{H}_o : d \leq 0$, we first report in our tables whether the observed empirical distributions are *seemingly* rankable by FSD or SSD. We present the sample values of $\max\{d_1\}$, $\max\{d_2\}$, $\widehat{d}$, $\max\{s_1\}$, $\max\{s_2\}$, and $\widehat{s}$. We then obtain simple bootstrap estimates of the probability that d lies in the non-positive interval (i.e. $\Pr\{d \leq 0\}$) using the relative frequency of $\{\widehat{d}^* \leq 0\}$, where $\widehat{d}^*$ is the bootstrap estimate of d (500 repetitions are used in the analysis below).[21] If this interval has a large probability, say 0.90 or higher, and $\widehat{d} \leq 0$, we may infer dominance to a desirable degree of confidence. If this interval has a low probability, say 0.10 or smaller, and $\widehat{d} > 0$, we may infer the presence of significant crossings of the empirical CDFs, implying an inability to rank the outcomes. Finally, if the probability lies in the intermediate range, say between 0.10 and 0.90, there is insufficient evidence to distinguish between equal

[20] Separate extensive examination of these alternative matching techniques is in progress and is curtailed here in the interest of brevity.

[21] Note, we also report simple bootstrap estimates of the $\Pr\{d \geq \widehat{d}\}$. These are provided to facilitate visualization of the simple bootstrap distribution.

and unrankable distributions. This is a classic confidence interval test; specifically, we are assessing the likelihood that the event $d \leq 0$ has occurred. Similarly, we estimate $\Pr\{s \leq 0\}$ to evaluate the second order dominance proposition given by $H_o : s \leq 0$.

We do not impose and test the Least Favorable Case (LFC) of equality of the distributions. This could be done by combining the data on X and Y and bootstrapping from the combined sample (e.g., Abadie 2002). Our bootstrap samples still contain N (M) observations from X (Y). As argued in Linton et al. (2003), working under LFC has some undesirable power consequences as it can produce biased tests that are not similar on the boundary of the null. This happens when the boundary of the null itself is composite. Following Maasoumi and Heshmati (2000) and Maasoumi and Millimet (2003), we are also reporting the *maximum test sizes associated with our (conservative) critical value of zero*, which is clearly on the boundary of the null that includes the LFC. Thus, the bootstrap probabilities reported represent the critical levels associated with this non-rejection region. Such critical levels can be shown to be conservative since, in the limit, they are at least as large as the corresponding levels for the asymptotic test on the boundary (Linton et al. 2003).

As an alternative, we also evaluate the less decisive dominance proposition $H_o : d = 0$ via the Linton et al. (2003) *recentered bootstrap* procedure. Linton et al. (2003) show that the recentered bootstrap technique provides a consistent test and competes rather well with their subsampling technique in terms of power. It is known that $\widehat{d}$ converges to d under general conditions (likewise for the SSD statistics). Under the null $H_o : d = 0$, centering of computations around their corresponding sample values introduces second order errors that are negligible for first order (asymptotic) approximations, but is desirable for removing some uncertainties due to estimation of unknown parameters and distributions.[22] This is the source of improvement in bootstrap power gained from recentering, especially for the residual dominance tests discussed above. The other source of improvement arising from recentering pertains to the technique's robustness to within-sample dependence.

To proceed, we obtain recentered bootstrap p-values in the classical sense as the relative frequency of $\{\widehat{d}^{**} > \widehat{d}\}$, where $\widehat{d}^{**}$ is the recentered bootstrap estimate of d. The recentering algorithm requires:

(i) generating bootstrap samples of size N (M) from X (Y),

(ii) computing the values of $\widehat{F}^*(z_q)$ and $\widehat{G}^*(z_q)$ for z_q, $q = 1, ..., Q$, where the values of z_q used to analyze the original sample are again utilized,

[22] In the residual tests, we also account for parameter uncertainty by re-estimating the first-stage regressions for each bootstrap resample. To accomplish this, we employ a nonparametric bootstrap where resamples of $\{w_{it}^{k*}, h_{it}^{k*}\}$ are drawn. This procedure is utilized in the simple bootstrap case as well.

(iii) computing the differences $d_1^c(z_q) = \left[\widehat{F}^*(z_q) - \widehat{G}^*(z_q)\right] - \left[\widehat{F}(z_q) - \widehat{G}(z_q)\right]$ and $d_2^c(z_q) = \left[\widehat{G}^*(z_q) - \widehat{F}^*(z_q)\right] - \left[\widehat{G}(z_q) - \widehat{F}(z_q)\right]$, and

(iv) finding $\widehat{d}^{**} = \sqrt{\frac{NM}{N+M}} \min\left\{\max\{d_1^c\}, \max\{d_2^c\}\right\}$.

If the $\Pr\{\widehat{d}^{**} > \widehat{d}\}$ – the p-value – is low, say 0.10 or smaller, we reject the null $\text{H}_o : d = 0$; if this p-value is greater than 0.10, we fail to reject the null. We also report the $\Pr\{\widehat{d}^{**} \leq 0\}$ in the tables. This allows the reader to see the significance level (size) of the test associated with the special critical value 'zero.' In our tables, these are obtained simply as $\Pr\{\widehat{d}^{**} > 0\} = 1 - \Pr\{\widehat{d}^{**} \leq 0\}$.

It is important to emphasize that while rejection of the null provides valuable insight in the recentered bootstrap case, failure to reject the null provides less information. If we reject the null and $\widehat{d} < 0$, we may infer dominance to a desirable degree of confidence. Conversely, if we reject the null and $\widehat{d} > 0$, we may infer 'unrankable' distributions. These are both strong findings, as the former (latter) indicates that all (not all) increasing social welfare functions will concur on the relative rankings of the distributions in question. On the other hand, failure to reject the null merely implies that we cannot eliminate the possibility that $F = G$; strict dominance also cannot be ruled out to some degree of confidence. Seen in this light, the recentered bootstrap is a conservative test. In the discussion of the results, we focus more heavily on the more decisive simple bootstrap for inference, as in Maasoumi and Heshmati (2000) and Maasoumi and Millimet (2003). Similarly, we report the relative frequency of $\{\widehat{s}^{**} > \widehat{s}\}$ and $\{\widehat{s}^{**} > 0\}$ to evaluate the null $\text{H}_o : s = 0$.

A final, necessary comment pertains to inference in the FR tests (i.e., those incorporating the Oaxaca-Blinder decomposition). Due to the usage of a common set of coefficient estimates in obtaining both residual distributions being compared, there *necessarily* exists between-sample dependence. For example, the FR test using data on single and married men in period two compares the distributions of $\widehat{\epsilon}_{i2}^s$ and $\widehat{\epsilon}_{i2}^{m,ob}$. The former depends on $\{w_{i2}^s, h_{i2}^s, \beta^s(w_2^s, h_2^s)\}$, where w_2^s and h_2^s represent the full data vector for w and h for the sample of singles in period two; the latter, $\widehat{\epsilon}_{i2}^{m,ob} = \ln(w_{i2}^m) - h_{i2}^m \beta_2^s$, depends on $\{w_{i2}^m, h_{i2}^m, \beta^s(w_2^s, h_2^s)\}$. This source of dependence is atypical. Between-sample dependence usually arises when the same individuals appear in the two samples being compared (e.g., distributions of pre- and post-tax incomes for a sample of individuals). To handle this more common type of between-sample dependence, pairwise (or 'clustered') bootstrap samples are drawn in order to maintain the dependence in the resampled data (Linton et al. 2003). In the current situation, the between-sample dependence is maintained by re-estimating the first-stage equations (5) and (6) on each bootstrap resample.[23] Specifically, by resampling N observa-

[23] Similarly, when utilizing the propensity score or the S_i index to control for observables, we estimate the propensity score and S_i anew during each bootstrap repetition.

tions $\{w_{i2}^{s*}, h_{i2}^{s*}\}$ and M observations $\{w_{i2}^{m*}, h_{i2}^{m*}\}$ nonparametrically and re-estimating (5), we obtain the resampled distributions of $\widehat{\epsilon}_{i2}^{s*}$ and $\widehat{\epsilon}_{i2}^{m,ob*}$, where the former depends on $\{w_{i2}^{s*}, h_{i2}^{s*}, \beta^{s*}(w_2^{s*}, h_2^{s*})\}$ and the latter depends on $\{w_{i2}^{m*}, h_{i2}^{m*}, \beta^{s*}(w_2^{s*}, h_2^{s*})\}$. Thus, as in the usual pairwise bootstrap case, the source of between-sample dependence is maintained in the resampling procedure.

2.4 Data

The data used to implement the difference-in-differences SD tests are obtained from the CPS March Supplement. The CPS interviews households who do not change residence for two consecutive periods.[24] For the comparisons we wish to make, we require males who were single in both survey periods, as well as males who were single in the first period and married in the second period. The difficulty is that many males who marry in between survey years change residences, and therefore drop out of the CPS sample. Consequently, the to-be-married/married males we do observe may not be a random sample; we shall return to this below. To circumvent the small sample size issue, we utilize matched pairs of individuals over the period $1992 - 2001$.[25] We restrict the sample to include only employed, native born males between ages 25 and 65; omitting those in school, enrolled in the military, employed in agriculture, disabled, and self-employed.

The outcome of interest is the hourly wage, which is constructed using data on annual wages and salaries, usual number of hours worked per week, and number of weeks worked last year. All wages were converted to (1982) real dollars using the CPI deflator for the respective year. To eliminate the effects of outliers, we drop observations with wages below \$1/hr and above \$100/hr (Loh 1996).[26] The final sample contains 11,154 observations: 4861 (5733) single individuals in period one (two), and 263 (297) to-be-married (married) individuals in period one (two). The balanced sample utilized in the first-difference comparisons contains 4786 (261) single (married) individuals.

To obtain the partial and full residual wages, we utilize an extensive set of individual, family, occupational, and location variables available in the CPS. Specifically, the vector h in (5), (6), (7) includes the following variables (in addition to a constant term): racial dummies (White, Black, American-Indian, Asian and Other), age, age squared, educational attainment dummies (less than high school, some college

[24]The CPS uses the 4-8-4 rotation process for interviewing households. A particular household is interviewed for four consecutive months, then kept out for the following eight months and re-interviewed for another four months. Each year there is an incoming rotation group and an outgoing rotation group.

[25]For further details on the data construction, see Millimet et al. (2003).

[26]We use two other methods to trim outliers: (i) drop men earning wages outside the 5^{th} and 95^{th} percentiles, and (ii) substitute $q_w(5)$ for wages below the 5^{th} percentile and $q_w(95)$ for wages above the 95^{th} percentile. Results do not qualitatively change.

without degree, college and above), number of own children younger than six, dummies for type of housing (own or rent), occupation dummies, full-time employee indicator, dummy for membership in a labor union or similar employee association, class of worker dummies (private, federal, and state and local government), urban and year dummies, and nine regional dummies. Note that unlike typical methods such as standard regression analysis, the variable of interest – marital status – is not included in the regressions estimated herein. Summary statistics are provided in Table 1. Interestingly, to-be-married men earn approximately two percent more per hour than single men; married men earn nearly 18 percent more.

As noted above, given the nature of the CPS, there are concerns about the representativeness of our sample. To investigate, we also display in Table 1 summary statistics for all married and single males in the incoming rotation group of the 1996 CPS (chosen at random). In addition, we provide p-values from t-tests testing the equality of means across our single (married) sample and the 1996 complete incoming rotation group of single (married) men. The three most obvious differences between our sample and the 1996 full sample is that our married sample (i) is younger (32.77 years versus 42.10 years), (ii) has fewer children under age six (0.09 versus 0.38), and (iii) is much more likely to rent, rather than own, a home (55% ownership versus 79%). The majority of other differences are either statistically insignificant, or, where statistically significant, are minor in magnitude. To further assess the impact of the selection criteria utilized, Table 2 reports the results of OLS and fixed effect (FE) regressions using our sample. In addition, benchmark results from Korenman Neumark (1991) are displayed for comparison. While the set of controls are slightly different, we obtain a marriage premium of 0.151 (0.217) using OLS and our period two (pooled) sample, versus 0.11 for Korenman and Neumark (1991). Estimating a fixed effects model, we obtain a premium of 0.068, versus 0.06 for Korenman and Neumark (1991). Thus, in practice, our sample does not appear to be overly selective.[27]

2.5 Results

2.5.1 Unconditional Tests

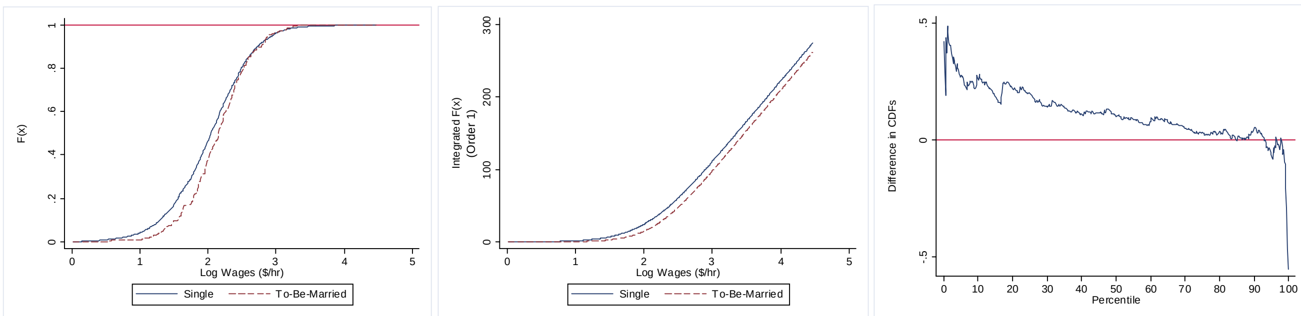
Figure 1 plots the raw distributions (CDFs and integrated CDFs) of (log) wages, as well as their differences at each percentile. The top row contains the first period comparisons (i.e. single versus to-be-married); the second row contains the second period comparisons (i.e. single versus married). The bottom row displays

[27] Neumark and Kawaguchi (2001) document (in a parametric regression framework) that sample attrition in the matched CPS data may understate the marriage premium as the magnitude of the premium appears to be larger for men who change residences after marriage. Given the comparability of our findings to previous studies utilizing other panel data sets (such as the various NLS data sets or PSID), the size of any bias appears small at best. Furthermore, in our residual SD tests, we condition on housing type to further reduce any bias arising from non-random attrition.

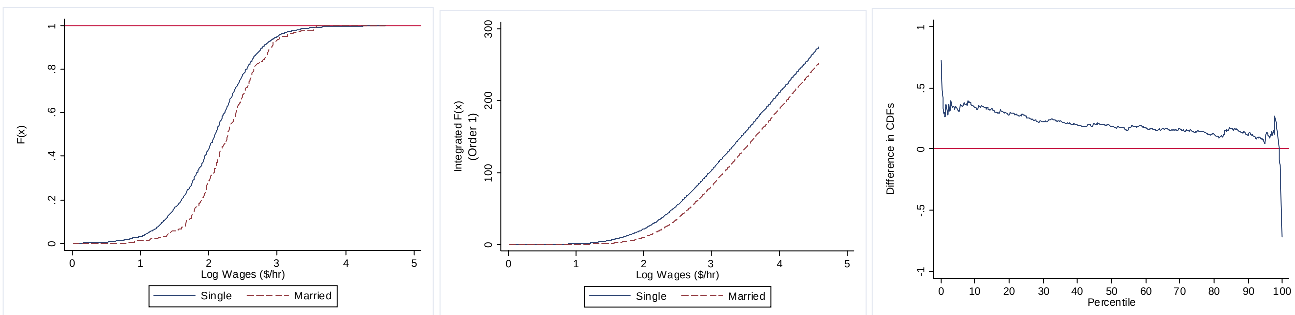
the distributions of wage changes from period one to two for single and married men. The corresponding SD results are displayed in Table 3.

Comparing the distributions of wages in period one (Panel A), it *appears* that the wage distribution for the to-be-married group second order dominates the wage distribution for single men; all increasing and concave welfare functions with an aversion to dispersion (inequality) would favor the to-be-married group. Indeed, referring to the plots in Figure 1, we see that if not for the crossings at above the 85^{th} percentile, we would have observed first order dominance. However, some single men earn very high wage rates. Moreover, the difference in the CDFs is largest in the bottom tail of the wage distribution, and declines steadily as wages rise. Thus, while previous research has documented a positive relationship between wages and the likelihood of marriage (e.g. Nakosteen and Zimmer 1997; Ginther and Zavodny 2001), this relationship is not uniform across the wage distribution and is clearly strongest in the bottom tail of the distribution.

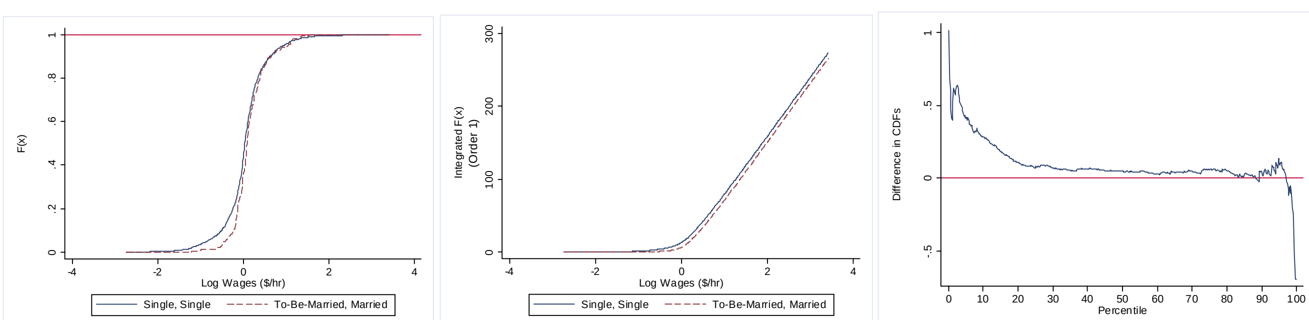
However, simply viewing the graphs is insufficient to draw inference. To assess the statistical significance of this apparent second order relation, we turn to the bootstrap results. The recentered bootstrap yields a p-value of 1.000, with the size of the test at the critical value of 'zero' being 0.562 ($= 1 - 0.438$), while the simple bootstrap yields a probability of $s \leq 0$ of 0.856. The recentered p-value implies that we cannot reject the null $\mathrm{H}_o : s = 0$. The simple bootstrap provides relatively strong evidence of a second order SD ranking, but is not statistically significant at conventional confidence levels. Nonetheless, only when one goes beyond mean wages, taking 'dispersion' considerations into account, does one find a moderately statistically significant (unconditional) wage advantage for to-be-married males that is robust across welfare criteria (i.e. produces a uniform ranking).



Period 1



Period 2



First Differences

Figure 1. CDFs and Integrated CDFs: Unconditional Wages

Note: The difference in CDFs are calculated as 'to-be-married' minus 'single' for Period 1, 'married' minus 'single' for Period 2, 'to-be-married, married' minus 'single, single' for First Differences.

Comparing the distributions of wages in period two (Panel B), the sample wage distribution for married men first order dominates the distribution for single men. Referring to the plots in the second row of Figure 1, we see that the CDFs do, however, cross at above the 99^{th} percentile.[28] In addition, as in period one, the difference in the CDFs is largest in the lower tail of the wage distribution, falling as wages rise until about the 95^{th} percentile. In terms of statistical significance, the recentered bootstrap yields a p-value of 1.000, with the size of the test at the critical value of 'zero' being 0.990, failing to reject the null $H_o : d = 0$. The simple bootstrap, on the other hand, fails to provide strong support for FSD with $\Pr\{d^* \leq 0\} = 0.392$. The simple bootstrap, however, yields $\Pr\{s^* \leq 0\} = 1.000$.[29] As this is stronger evidence of a second order relation than found using unconditional wages in period one, these results suggest that the wage advantage enjoyed by the majority of 'marrying-men' continues, and may well increase, post-marriage. The results also confirm the need to broaden our concept of the marriage premium in order to incorporate wage dispersion into the discussion. Absent the inclusion of dispersion considerations, the (unconditional) marriage premium is not robust across all welfare criteria, failing to give rise to uniform rankings of marital states.

To determine if, in fact, the disparity in wage distributions increases after marriage, we examine the distribution of wage changes (Panel C). Here, we find that the distribution of wage changes for married men second order dominates the respective distribution for single men. Examining the third row of Figure 1, we see that single men are much more likely to suffer large wage decreases from period one to two, the two distributions are fairly equal over the middle range (wage changes close to zero), and that married men are much more likely to enjoy large wage increases from period one to two. Thus, if not for the crossings above the 90^{th} percentile or so, we would have observed first order dominance. In terms of the statistical significance of this second order dominance, the recentered bootstrap method yields p-values of 1.000, once again failing to reject the null $H_o : s = 0$. The simple bootstrap gives a probability of $s \leq 0$ of 1.000, confirming the welfare improvement in the wage distribution post-marriage suggested by the previous results in Panels A and B. Thus, any social welfare function that is increasing and concave in wages would conclude that marriage improves the welfare of individuals. However, individuals possessing different preference functions in the class U_1 can reasonably disagree about whether in fact marriage improves the welfare of men as there are groups of men whose welfare (wages) are higher when unmarried, thus precluding a uniform ranking of distributions over the class of all increasing welfare criteria.

[28] While the CDFs cross in the extreme upper tail, Table 3 reports a finding of FSD because of the 'trimming' procedure utilized; without this 'trimming' procedure, a statistically significant SSD finding is obtained (see footnote 10).

[29] Appendix A contains the four bootstrap distributions of the statistics in Panel B in Table 3 – d^*, s^*, d^{**}, and s^{**} – to illustrate one particular case for the reader. The full set of bootstrap distributions are available from the authors if desired.

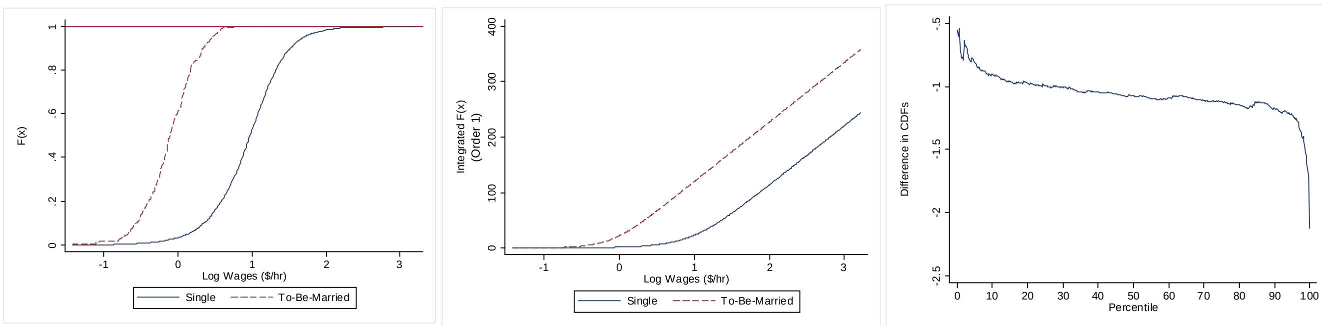
As a whole, the unconditional tests confirm the existence of a marriage premium at the distributional level. The observed (and weakly supported) FSD ranking, in addition to the statistically significant SSD ranking, in period two indicates that the (unconditional) marriage premium is, on the one hand, more robust than previously thought because of the uniform ranking by all increasing and inequality averse welfare criteria. On the other hand, it is less robust than previously documented in that such uniform rankings are only obtained when one moves beyond simple mean wage comparisons and incorporates dispersion into the welfare criteria. The introduction of dispersion into the discussion is necessary given the insignificant differences (or significant crossings) in the wage distributions of single and married men at high wages.

Moreover, in line with Becker (1976), Cornwell and Rupert (1997), and others, we find that the vast majority of the (unconditional) wage advantage enjoyed by married men predates marriage. In addition, unknown heretofore, the distributional approach demonstrates that the (unconditional) return to marriage is not uniform, but rather favors the lower end of the wage distribution, and that the previously documented positive association between wages and the probability of marriage does not hold in the upper tail of the wage distribution. These findings are merely suggestive, however, as they fail to control for time-varying observables correlated with both marital status and wages. For instance, Table 1 indicates that to-be-married men are better educated, more likely to belong to a union, more likely to work full-time, and are more likely to be white. To determine if the unconditional results hold once we purge wages of such attributes, we now turn to the residual SD tests.

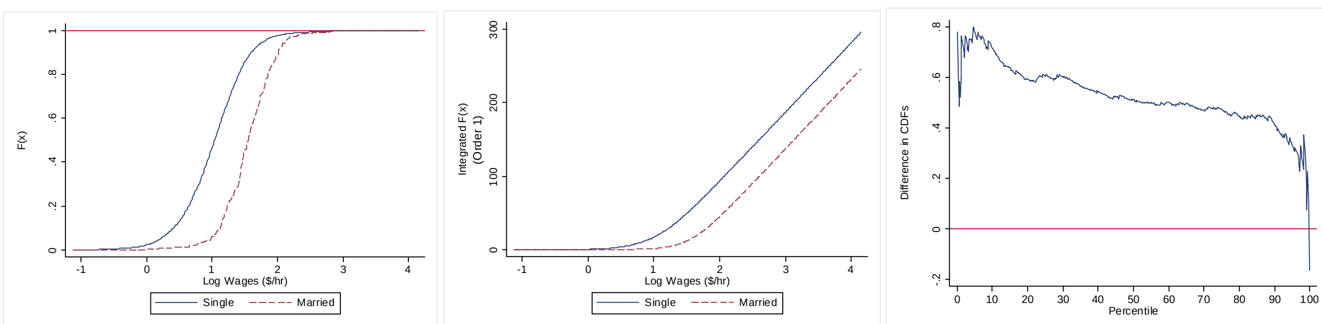
2.5.2 Residual Tests

Partial Residual Tests The PR results are displayed in Table 4; the corresponding plots are contained in Figure 2. Recall, the PR tests compare the distributions of intercept-adjusted log wage residuals, where the intercepts and coefficients used to obtain the residuals are period and marital-state specific.[30] If time-varying observables are associated with both wage outcomes and marital status, these results may differ from the unconditional results discussed previously.

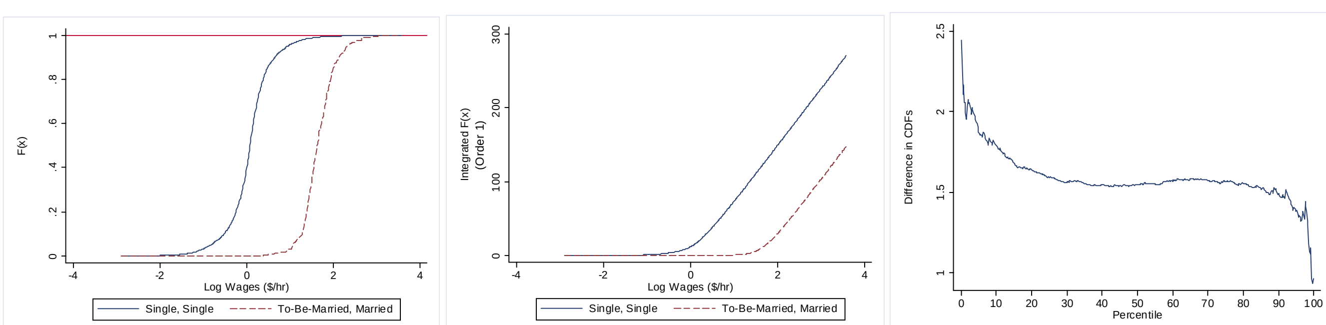
[30] To be clear, four first-stage OLS regressions are estimated: (i) single men in period one, (ii) single men in period two, (iii) to-be-married men in period one, and (iv) married men in period two. Results are not shown, but available upon request.



Period 1



Period 2



First Differences

Figure 2. CDFs and Integrated CDFs: Partial Residuals

Note: The difference in CDFs are calculated as 'to-be-married' minus 'single' for Period 1, 'married' minus 'single' for Period 2, 'to-be-married, married' minus 'single, single' for First Differences.

Comparing the residual distributions in period one (Panel A), we now observe that the distribution for the single group seemingly first order dominates the distribution for to-be-married men (in stark contrast to the unconditional results in Table 3, Panel A). Referring to the top row of Figure 2, we see that the disparity in the distributions increases gradually as PR wages rise, and then increases even further at the upper tail of the distribution. In terms of the statistical significance of this first order relation, the recentered bootstrap fails to reject the null $H_o : d = 0$. The simple bootstrap offers somewhat strong evidence for FSD and SSD ($\Pr\{d \leq 0\} = 0.744; \Pr\{s \leq 0\} = 0.802$). One might expect a third order SD ranking in this situation which would incorporate 'increasing aversion to inequality.' This reversal (or loss) of the rankings from the unconditional to the PR case indicates that time-varying observables which are more prominent in the to-be-married sample are associated with higher wages. Thus, once we purge wages of the effects of these attributes, we no longer find evidence that the distribution of wages of to-be-married men dominates the distribution of wages of single men.

Turning to the PR distributions in period two (Panel B), it now seems (as in the unconditional case) that the married group first order dominates the single men.[31] Consequently, all increasing welfare functions appear to favor the wage distribution for married men. Examining the middle row of Figure 2, we see once again that the disparity in the distributions is largest at the bottom tail of the distribution, and decreases gradually as PR wages rise. Both bootstrap procedures fail to support FSD at conventional levels, but there is relatively strong evidence for SSD of married men over singles with 0.85 level of confidence. However, the fact that the observed ranking reverses from period one (Panel A) to period two (Panel B) leads to a statistically significant result when we examine the distribution of wage changes (Panel C). Specifically, we find that the PR distribution for married men first order dominates the corresponding distribution for single men, and the disparity in the distributions remains greatest in the lower tail of the distribution (Figure 2). In addition, this ranking is statistically significant according to the simple bootstrap ($\Pr\{d \leq 0\} = 0.772$; $\Pr\{s \leq 0\} = 0.950$). The recentered bootstrap fails to reject the null $H_o : d = 0$ or $s = 0$.

The PR tests, in sum, provide evidence that is more favorable to causal explanations (i.e. the specialization hypothesis or employer discrimination) and less supportive of the selection hypothesis relative to the unconditional results. Furthermore, the PR results confirm the non-uniformity of the return to marriage documented in the unconditional results, highlighting the usefulness of the distributional approach and the necessity of including inequality considerations into welfare comparisons of marital states for 'majority' ranking. It seems that causal or selection factors have at least as large an impact on 'dispersion' as on the

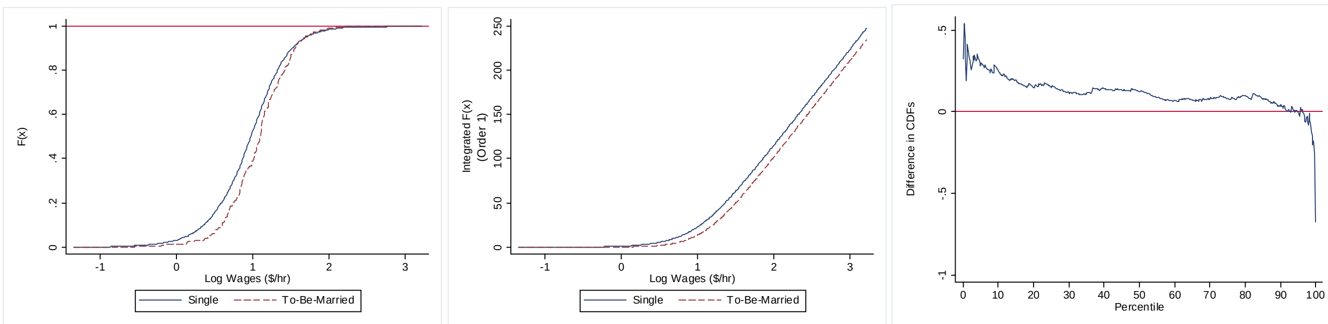
[31] As in the middle row of Figure 1, the CDFs in the middle row of Figure 2 do, in fact, cross in the extreme upper tail (above the 99^{th} percentile). Table 4 reports a finding of FSD because of the 'trimming' procedure utilized; without this 'trimming' procedure, a statistically significant SSD finding is obtained (see footnotes 10, 28).

average wages.

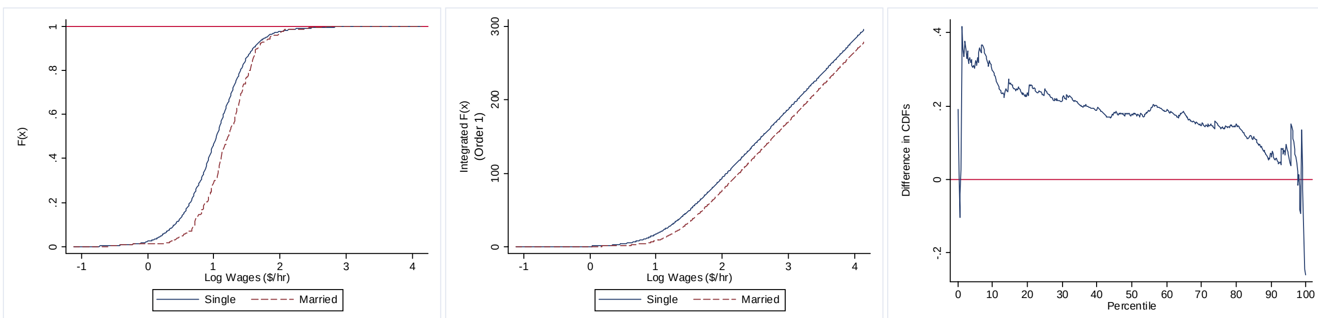
As noted in Section 2.1, however, a potential shortcoming of the PR tests is that differences in the marital-state specific returns to observables are not included in the PR distributions. As a result, any differences in the returns to observable characteristics are not attributable to the effects of marriage, which may yield misleading inferences. For example, if the return to education is higher amongst married men than single men in period two, that may represent evidence further in support of the specialization hypothesis. Alternatively, if the return to education is higher amongst to-be-married men than single men in period one, that may indicate greater innate ability amongst the to-be-married men, providing evidence in favor of the selection hypothesis. To examine the impact of incorporating differences in returns into the residual distributions, we now turn to the FR test results.

Full Residual Tests The FR results are displayed in Table 5; the corresponding plots are presented in Figure 3. Comparing the residual distributions in period one (Panel A), we now observe that the distribution for the to-be-married sample may second order dominate the distribution for single men (in contrast to the PR results, but consonant with the unconditional results). Referring to the top row in Figure 3, we see that the disparity in the distributions is largest at the bottom tail of the distribution, declining gradually until the CDFs eventually cross at around the 95^{th} percentile, confirming the previous results that wages and the likelihood of marriage are uncorrelated in the upper tail of the wage distribution.

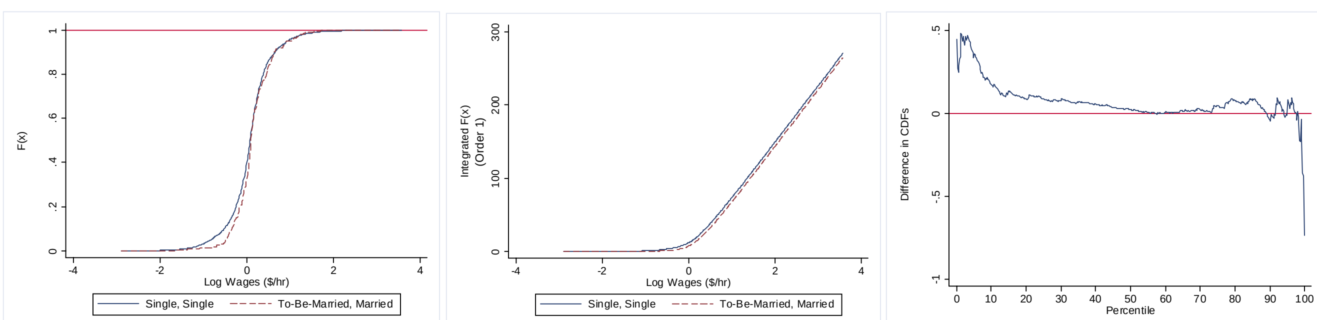
In terms of the statistical significance of this second order relation, the simple bootstrap yields a probability of $s \leq 0$ of 0.928, and the recentered bootstrap yields a large p-value, strongly supporting the observed ranking. The reversal in rankings from the PR to the FR tests indicates that the returns to observable attributes favor to-be-married men. As a result, once we incorporate the differential returns into the residual distributions, we continue to find evidence that the distribution of wages of to-be-married men dominates the distribution of wages of single men *prior* to marriage (consonant with the selection hypothesis). Again, one has to combine considerations of wage levels (averages) and dispersion to arrive at such a uniform ranking.



Period 1



Period 2



First Differences

Figure 3. CDFs and Integrated CDFs: Full Residuals

Note: The difference in CDFs are calculated as 'to-be-married' minus 'single' for Period 1, 'married' minus 'single' for Period 2, 'to-be-married, married' minus 'single, single' for First Differences.

To further assess this finding, we perform the standard Oaxaca-Blinder decomposition of wages in period one using (7). As expected, the difference in coefficients is extremely large and favors the to-be-married sample. Specifically, the mean log wage differential in period one is 0.112; the gap due to the difference in coefficients (part C in (7)) is 1.170.[32] Assuming the differential returns reflect greater unobserved ability in the to-be-married sample (and not other sources, such as greater measurement error in the single sample), this confirms the findings in Cornwell and Rupert (1997) and others that to-be-married men fare well relative to single men prior to marriage.

Comparing the residual distributions in period two (Panel B), we are unable to rank the observed distributions (in either the first- or second-degree sense). The fact that the PR distribution for the married group first order dominates the corresponding distribution for single men (Table 4, Panel B) indicates that the returns to observable attributes favor single men in period two. Indeed, results from the standard Oaxaca-Blinder decomposition of wages in period two using (7) is in conformity with this; the mean log wage differential in period two is 0.198, but the gap due to the difference in coefficients (part C in (7)) is -0.324.[33] While the distributions are unrankable in the first- and second-degree sense despite the advantage in returns favoring single men in period two, the FR distribution for married men would nonetheless first order dominate the corresponding distribution for single men if not for a few crossings at the extreme tails of the distribution (Figure 3, middle row).

Finally, comparing the distributions of FR changes (Panel C), we observe that the married sample second order dominates the single sample, a slightly weaker finding than the observed FSD ranking noted in Panel C of Table 4. Consonant with Figure 2, the bottom row of Figure 3 shows that the largest disparity occurs in the bottom tail of the distribution. In terms of the statistical significance of this second order relation, the recentered bootstrap fails to reject the null $\text{H}_o : s = 0$, and the simple bootstrap offers only modest evidence of a second order ranking, giving a probability of $s \leq 0$ of 0.740.

In the end, we believe the FR tests to best isolate the effects of marriage on the distribution of wages, and the simple bootstrap to be the more informative method of inference. According to this criteria, we conclude that after purging wages of time-varying observables, (i) the distribution of wages

[32] The majority of this difference is due to the greater return to age (proxying for experience) for to-be-married men, although the return to education and the returns to select occupations also favor to-be-married men. The full set of results are available upon request.

[33] The main reason for the reversal from period one to period two is the changes in relative magnitudes of the coefficients on age and age squared. In period one, the age-earnings profile is much steeper for to-be-married than single men (both coefficients are statistically significant at conventional levels, both individually and jointly within each regression). In period two, the age-earnings profile is steeper for single than married men (both coefficients are jointly significant within each regression, but are individually insignificant in the married sample).

favors married men *prior* to marriage (consonant with the selection hypothesis), particularly at the lower tail of the distribution (yielding a marginally significant SSD ranking) and (ii) the distribution of wages favors married men *after* marriage, except at the extreme tails of the distribution (failing to give rise to either a FSD or SSD ranking). Furthermore, after purging wages of time-invariant unobservables via the difference-in-differences distributional approach, we conclude that (iii) the distribution of wage changes favors married men. However, the wage changes are modest, and confined predominantly to the lower tail of the distribution (yielding a statistically insignificant SSD ranking). Thus, we find that the marriage premium persists at the distributional level, and find a *small* role for causal-based explanations of the marriage premium after controlling for selection. However, the marriage premium is not constant across the distribution of wages. Specifically, we find that the premium is largest in the lower tail of the distribution, and negligible in the upper tail. As a result, uniform rankings are only possible when the marriage premium concept is broadened to incorporate wage dispersion into the evaluation of marital states.

Yet, as stated previously, such conclusions may be sensitive to the projection method used to purge wages of potentially confounding covariates, as well as are the presence of time-varying unobservables correlated with both marital status and labor market performance. We assess the validity of the conclusions drawn from the difference-in-differences SD approach to each potential criticism in turn.

2.5.3 Sensitivity Analysis

In the interest of brevity, we only present the graphical results from the analysis utilizing the propensity score and the S_i index to control for observables.[34] Plots obtained utilizing the propensity score are given in Figure B1 in Appendix B. The top panel in the figure plots the CDFs of hourly wages for 'to-be-married' men and their single counterparts matched (with replacement) on the basis of the nearest propensity score; the middle panel corresponds to the wages of married versus matched single men. The bottom panel gives the CDFs of the change in wages from period one to two for individuals matched on the basis of their second period propensity score. The graphs display striking similarity to those obtained from the FR distributional plots in Figure 3. Specifically, the plots reinforce the notion that the wage distributions of 'to-be-married' and single men are less disparate at the upper tail; however, second order dominance cannot be ruled out. The second row shows a large marriage premium for men over much of the distribution, although the CDFs continue to cross in the tails. Finally, the bottom row confirms the observed second order dominance ranking found in Figure 3, although some differences do arise in the extreme upper tail. Overall, then, the analysis continues to support the selection hypothesis and the finding of a small role for causal-based explanations of the marriage premium (confined mostly to the lower tail of the wage distribution) even

[34] The full set of results are available from the authors upon request.

when using the propensity score – rather than OLS – to control for covariates.

Figure B2 contains the plots after matching individuals using S_i with $\beta = -0.5$; results are unaffected by the choice of β within the range examined. As with the PSM results, the graphs are consonant with the results obtained based on the FR method. Specifically, although some minor differences arise in the behavior of the CDFs in the tails, the bottom panel shows a modest marriage premium at the distributional level after controlling for selection. The only subtle difference between the current results, using S_i, and the previous results based on the FR and propensity score methods is that now the premium is fairly uniform across the distribution; aside from the tails of the distribution, the results are qualitatively the same. To further assess the validity of the conclusions drawn from the difference-in-differences SD approach, we now turn to an alternative method – an IV distributional analysis – that is robust to the presence of time-varying unobservables.

3 Instrumental Variable Methodology

3.1 Test Statistics

To determine if the presence of time-varying unobservables correlated with both marital status and wages preclude us from drawing valid conclusions from the tests in Section 2, we implement the methodology developed in Abadie (2002) to compare the distributions of potential wages for (a subpopulation of) married and single men using an IV method. According to Imbens and Rubin (1997), when a binary instrumental variable is available, the potential distributions of the outcome variable are identified for the subpopulation (referred to as *compliers*) whose treatment assignment (in this case, marital status) is potentially affected by variation in the instrument.

Define W_0 and W_1 as the distribution of potential outcomes (wages) for the untreated (single) and treated (married), with $w_i(0)$ and $w_i(1)$ representing specific values for observation i, $i = 1, ..., N + M$, from the respective distribution. Let D_i be a binary variable equal to zero (one) if the individual is single (married), and Z_i be a binary instrument (discussed below). Denote $D_i(0)$ the value of D_i if $Z_i = 0$; similarly for $D_i(1)$. Thus, D_i may be written as

$$D_i = \begin{cases} D_i(0) & \text{if} \quad Z_i = 0 \\ D_i(1) & \text{if} \quad Z_i = 1 \end{cases} \tag{8}$$

Given this setup, for any individual i, the pair of marital status indicators $\{D_i(0), D_i(1)\}$ and the pair of potential wages $\{w_i(0), w_i(1)\}$ are not both observed since only one state of the world – $Z_i = 0$ or $Z_i = 1$ – is observed. Instead, the realized treatment assignment $D_i = D_i(1)Z_i + D_i(0)(1 - Z_i)$ and the realized potential outcome $w_i = w_i(1)D_i + w_i(0)(1 - D_i)$ are observed.

Let $F^c(w)$ and $G^c(w)$ represent the CDFs of potential wages for single and married 'compliers,' which are defined as follows:

$$\begin{aligned} F^c(w) &= \mathrm{E}\left[\mathrm{I}\{w_i(0) \leq w\}|D_i(1) = 1, D_i(0) = 0\right] \\ G^c(w) &= \mathrm{E}\left[\mathrm{I}\{w_i(1) \leq w\}|D_i(1) = 1, D_i(0) = 0\right] \end{aligned} \tag{9}$$

If Z_i satisfies the following three assumptions:

(i) Independence: $\{w_i(0), w_i(1), D_i(0), D_i(1)\} \perp Z_i$

(ii) Correlation: $\Pr(Z_i = 1) \in (0, 1)$ and $\Pr(D_i(0) = 1) < \Pr(D_i(1) = 1)$

(iii) Monotonicity: $\Pr(D_i(0) \leq D_i(1)) = 1$,

then the dominance tests defined in the previous section conducted on the distributions $F^c(w)$ and $G^c(w)$ identify the causal effect of marriage for the subpopulation of compliers, even if there exist time-varying, individual-specific attributes correlated with both marital status and wages (Imbens and Angrist 1994; Angrist et al. 1996). Moreover, as shown in Abadie (2002), SD tests conducted on the distributions $F^c(w)$ and $G^c(w)$ are equivalent to tests conducted on the distributions $F(w)$ and $G(w)$, where F (G) represents the distribution of wages for individuals with $Z_i = 0$ $(Z_i = 1)$. Thus, the test statistics in (3) and (4) are obtained by replacing F and G with their empirical counterparts:

$$\widehat{F}_{N_0}(w) = \frac{1}{N_o}\sum_{i=1}^{N_0}\mathrm{I}\left(W \leq w\right) \tag{10}$$

$$\widehat{G}_{N_1}(w) = \frac{1}{N_1}\sum_{i=1}^{N_1}\mathrm{I}\left(W \leq w\right) \tag{11}$$

where N_0 (N_1) is the size of the sample with $Z_i = 0$ $(Z_i = 1)$.

In the analysis, the IV, Z_i, is an indicator of whether the marriage of the individual's parents remained intact. A vast sociology literature is divided on the impact of family structure on offspring marital timing. While some find that parental divorce makes marriage more likely, others show that it delays or deters marriage (e.g., Kobrin and Waite 1984; Goldscheider and Waite 1986, 1991; Avery et al. 1992; Li and Wojtkiewicz 1994; South 2001). Much of this ambiguity, however, has been shown to relate to the age and cohort of the individuals under study. In particular, Wolfinger (2003) finds that parental divorce greatly increased the probability of marriage in the early 1970s, but lowered the likelihood of marriage in the mid-1990s. Furthermore, Wolfinger (2003) documents that parental divorce raises the likelihood of teenage marriage, but reduces the probability of marriage conditional on remaining single until age 20. Given that our sample (discussed below) contains individuals over age 20 from the mid-1990s, we suspect that coming from an intact family should raise the likelihood of marriage in our sample.

Aside from the issues of correlation and monotonicity, the instrument must also satisfy the independence assumption. Although untestable, Manski et al. (1992, p. 35) provide evidence that our "exogeneity assumption is not far off the mark." Manski et al. (1992) analyze the impact of parental marital status on children's educational outcomes, concluding that previous research on the topic that assumed the exogeneity of the parents' marital status yielded fairly accurate inference. Given the fact that the marital decisions of children tend to occur later in life than educational decisions (Manski et al. (1992) focus on high school completion), the exogeneity assumption seems even less problematic in the current context.

Finally, as in the previous section, in the empirical analysis the outcome being compared may be the unconditional wage, the partial residual, or the full residual. The simple and recentered bootstrap techniques are utilized for inference. For robustness, we also utilize the propensity score and the S_i index to control for observables.[35]

3.2 Data

To conduct the Abadie (2002) IV tests, we use cross-sectional data from the PSID from 1994. First, we switch from the CPS to the PSID since the latter has information on an individual's parents' marital history (if the individual is a child of an original sample member), whereas the former does not.[36] Second, we choose the 1994 wave for no particular reason other than that it is fairly centrally spaced with respect to the time range of the CPS sample. Upon dropping observations according to the same criteria as applied to the CPS data, we obtain a sample of 1057 male family heads, of which 770 are currently married and 287 are never married. We utilize an individual's parents' marital status as the binary instrument.[37] Specifically, we trace the parents' marital history for every male head and record the parent's marital status from the time of the son's birth onwards. We define Z_i as zero if there is an incidence of parental divorce or separation during the son's lifetime, and a value of unity if the parents' marital union remained intact through 1994. Hourly wages are constructed and trimmed as before, and the same set of conditioning variables are used to control for individual, family, labor market, and locational attributes. Summary statistics are provided

[35] The only difference that arises compared with the application of these alternative strategies in the previous section is that now the treatment and control group are defined on the basis of the instrument, Z, rather than actual marital status.

[36] The CPS provides information only for those parents who reside with their children, which introduces additional sample selection issues.

[37] For robustness, we also attempted to utilize an IV based on the presence of unilateral divorce laws in the state of residence (Gruber 2000). However, the IV was not statistically significant in the first-stage. Other variables that have been utilized as IVs in the past, such as religiosity and parents' education, have been shown to be correlated with earnings as well.

in Table 6, and results from an OLS regression reveal an estimated marriage premium of 0.209 (see Table 2), consonant with the previous literature.

Before discussing the actual results, we note that the instrument is statistically significant in a probit regression with marriage as the dependent variable ($\beta = 0.345$, $s.e. = 0.093$; marginal effect = 0.114); thus, concerns associated with weak instruments do not seem warranted. This finding is consonant with Wolfinger (2003), who documents a negative association between parental divorce and children's marital probability during the 1990s for individuals over age 20.

3.3 Results

3.3.1 Unconditional Tests

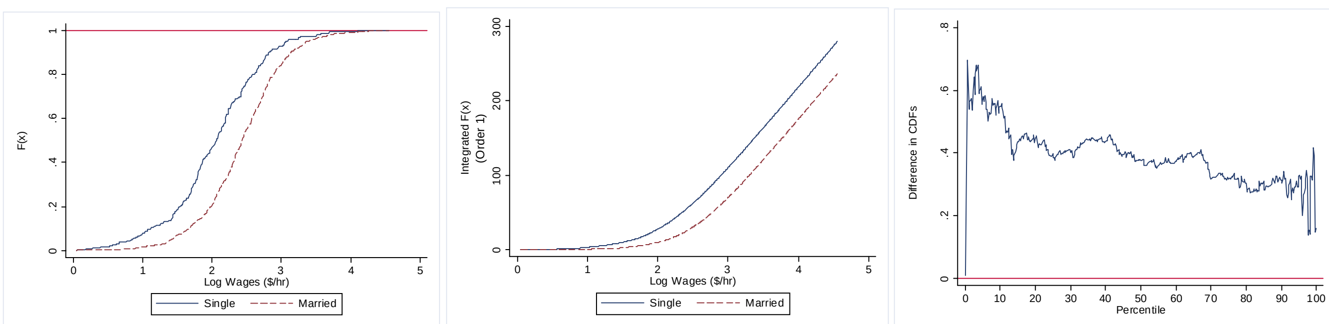
Figure 4 plots the unconditional distributions (CDFs and integrated CDFs) of (log) wages, as well as the differences in the unconditional CDFs at each percentile. The top row splits the sample according to actual marital status (i.e. not utilizing the IV); the bottom row partitions the sample according to the instrument.[38] The corresponding SD results are displayed in Table 7, with Panel A displaying the results based on actual marital status.(labelled 'Without Instrument') and Panel B containing the results using the instrument (labelled 'With Instrument').

The non-IV results in Panel A are remarkably similar to the unconditional period two results using the CPS sample (Table 3, Panel B). Moreover, the plots in the top row of Figure 4 are virtually identical to those in the middle row of Figure 1. As in Panel B of Table 3, we observe a ranking of first order dominance (married over single), but the ranking is statistically significant only in the second-degree sense according to the simple bootstrap ($\Pr\{d \leq 0\} = 0.572$; $\Pr\{s \leq 0\} = 0.928$). The recentered bootstrap fails to reject the null $H_o : d = 0$ or $s = 0$ (FSD: p-value > 0.99; SSD: p-value = 0.88). Furthermore, as in the CPS sample, we find that the disparity in distributions is largest in the bottom tail. Thus, we continue to document the need to consider both wage levels and wage dispersion in order to produce uniform rankings of (unconditional) wages across marital states.

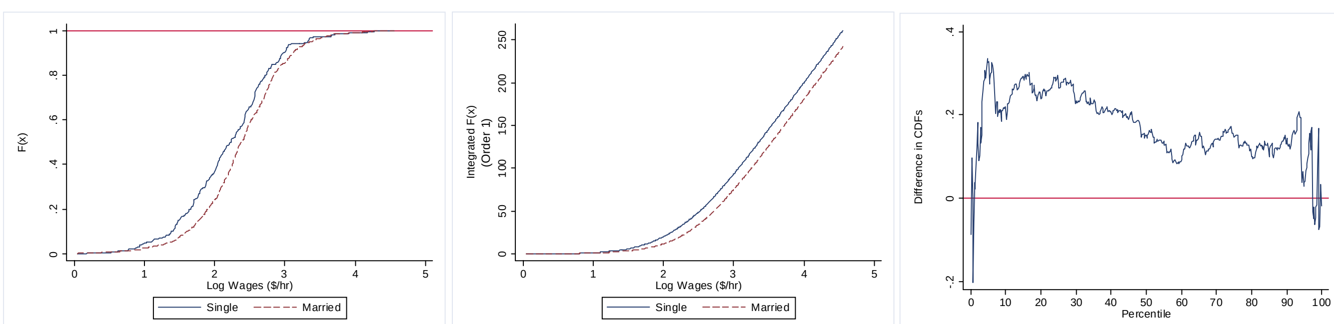
According to the IV results in Panel B, however, the wage distributions are unrankable (in the first- or second-degree sense). This suggests that if parents' marital status constitutes an exogenous source of variation in men's marital status, then the marriage premium is no longer robust to the choice among increasing and inequality averse welfare criteria. This statement is a bit strong, though, once we examine the bottom row of Figure 4. Plotting the CDFs shows that the crossings occur only in the extreme tails. Moreover, the crossing at the first percentile precludes a finding of even second order dominance.

[38] Thus, the sample denoted 'married' ('single') in the plots is the sample whose parents' marriage remained intact (dissolved).

Nonetheless, over a wide range of the unconditional distribution, the IV results indicate that the marriage premium clearly still exists, is larger at the bottom tail of the wage distribution, and contains a large causal component. Stated differently, there exists a rich class of cardinal welfare functions that will rank married men higher than singles. This conclusion rests on the validity of the IV. Since parents' marital history may be correlated with other potential wage determinants, in addition to marriage, such as education, purging wages of observable attributes is required for inference.[39]



Without Instrument



With Instrument

Figure 4. CDFs and Integrated CDFs: Unconditional Wages

Note: The difference in CDFs is calculated as 'married' minus 'single'.

[39]For example, the correlation between the IV and a college degree is of the same magnitude as the correlation between the IV and marital status.

3.3.2 Residual Tests

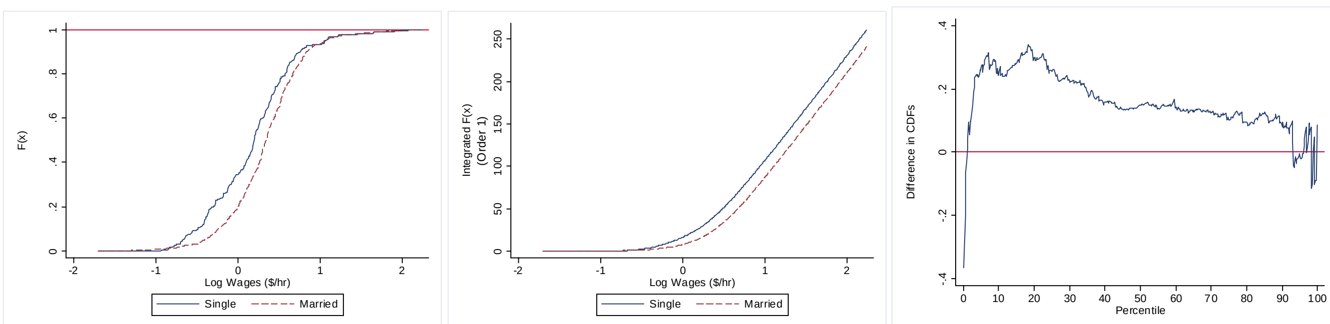
Partial Residual Tests The PR results are displayed in Table 8; the corresponding plots are contained in Figure 5. In both Panels A (without the instrument) and B (with the instrument), the observed distributions are not rankable in the first- or second-degree sense. However, the top row of Figure 5 shows that the CDFs only cross in the extreme lower and upper tails; over the majority of the distribution, there remains evidence of a marriage premium, and the return to marriage remains largest at the lower tail of the wage distribution. The bottom row of Figure 5, on the other hand, reveals that the wage distribution for married men lies to the right (left) of the wage distribution for single men over the lower (upper) portion of the distribution. This suggests that selection explains the entire marriage premium at the top of the wage distribution, but not at the bottom, according to the PR tests. This is consonant with the previous CPS results, revealing a larger marriage premium at the bottom of the wage distribution. To see if this conclusion stands once we incorporate differences in the returns into the residual distributions, we turn to the preferred FR tests.

Full Residual Tests The FR results are displayed in Table 9; the corresponding plots are presented in Figure 6. As with the PR tests, the observed distributions are not rankable in the first- or second-degree sense with (Panel B) or without (Panel A) utilizing the IV. Examining the top row of Figure 6, however, now reveals only a single crossing at the lower tail of the distribution; over the remainder of the distribution, the marriage premium is substantial and virtually *uniform*. The bottom row of Figure 6, on the other hand, indicates a subtle difference between the PR and FR tests. According to the PR test (Figure 5), the wage distribution for married men lies to left of the corresponding distribution for single men over much of the (upper portion of the) support. The FR test (Figure 6) indicates that the wage distribution for married men lies just to right of the corresponding distribution for single men over virtually the entire support, yielding a small, but *uniform*, causal return to marriage.[40]

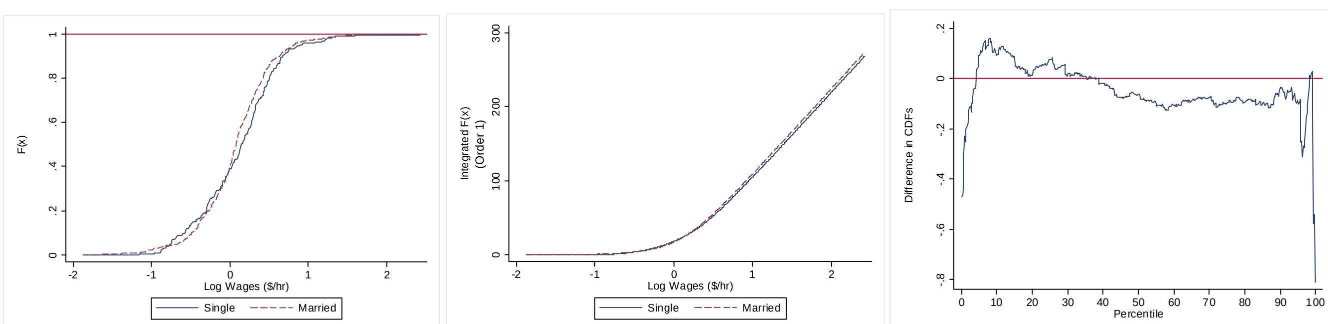
This reversal, in combination with the comparison of the non-IV-FR and IV-FR tests, as well as the previous results from the CPS sample, suggests that (i) a marriage premium exists at the distributional level in the PSID and CPS data, (ii) the majority of the premium is explained by selection, but there is a *small* role for causal-based explanations (i.e. specialization or employer discrimination), (iii) conclusions

[40] In general, the shift to the right of (at least) portions of the wage distribution for married men indicates that the difference in coefficients favors married men. Performing the standard Oaxaca-Blinder decomposition of wages using (7), we find (results not shown) a mean log wage differential of 0.390 without using the IV; the gap due to the difference in coefficients (part C in (7)) is 0.069. Utilizing the IV, we find (results not shown) a mean log wage differential of 0.175; the gap due to the difference in coefficients (part C in (7)) is 0.093.

regarding the portion of the premium not explained by selection are qualitatively similar across the two methods employed (first-differences and IV) and two data sets (CPS and PSID), and (iv) in contrast to the previous FR and propensity score matching results from the CPS, but consonant with the S_i based CPS results, the IV-FR results from the PSID indicate that the marriage premium is fairly uniform across the wage distribution.



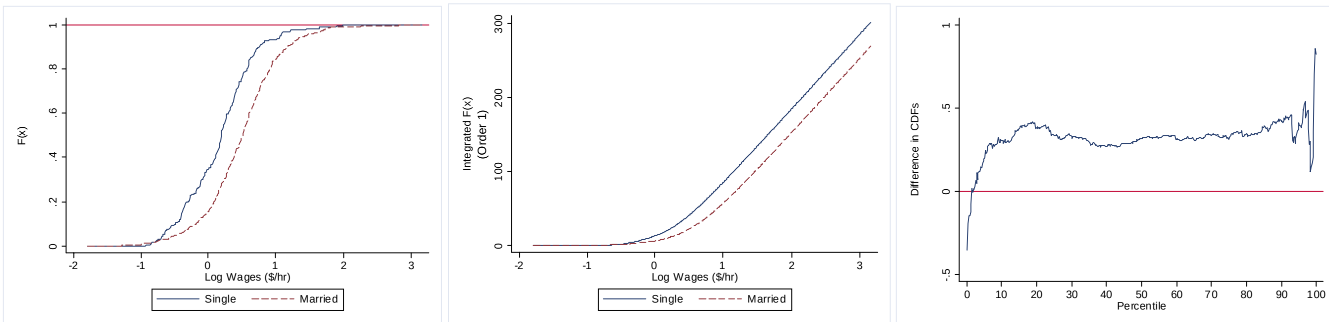
Without Instrument



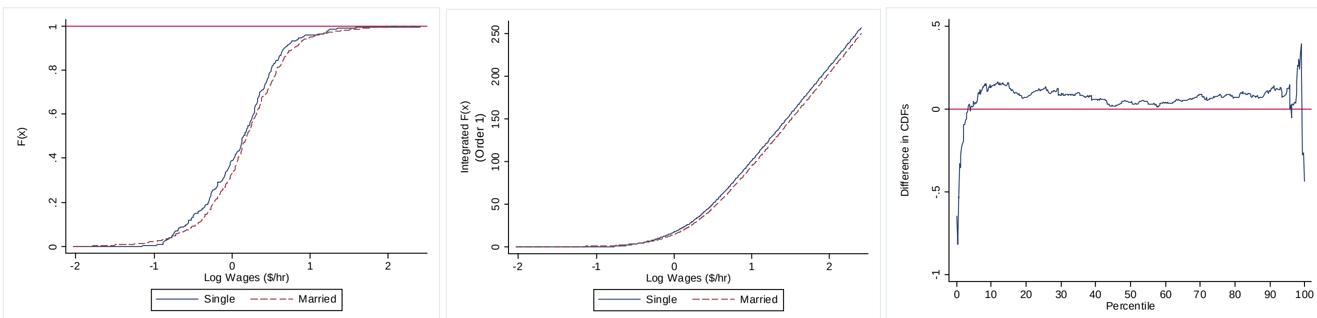
With Instrument

Figure 5. CDFs and Integrated CDFs: Partial Residuals

Note: The difference in CDFs is calculated as 'married' minus 'single'.



Without Instrument



With Instrument

Figure 6. CDFs and Integrated CDFs: Full Residuals

Note: The difference in CDFs is calculated as 'married' minus 'single'.

3.3.3 Sensitivity Analysis

In the interest of brevity, we once again only present the graphical results from the analysis utilizing the propensity score and the S_i index to control for observables. The results are displayed in Figures B3 and B4 in Appendix B. As in the previous section, the results are invariant to the choice of β; we display the results using $\beta = -0.5$. Moreover, the results continue to be consonant with the FR results shown in Figure 6. Specifically, both figures reveal a larger premium when the IV is not used (suggesting a substantial role for selection) and the IV-S_i results indicate a fairly uniform marriage premium across the distribution after controlling for selection; the propensity score method, on the other hand, suggests that the premium may be confined to only those below the median. Nonetheless, the basic conclusions about the relative merit of selection versus causal-based explanations of the marriage premium are robust across methodologies.

4 Conclusion

The existence of a return to marriage for men in the labor market is one of the many long-standing, stylized facts in labor economics. While the marriage premium has been well studied, all such studies (to our knowledge) focus on only one descriptive measure of the distribution of wages: the (conditional) mean. Such a narrow focus may be misleading as it may mask differences in the impact of marriage across the wage distribution and it lacks a solid welfare underpinning.

Utilizing recently developed nonparametric tests for stochastic dominance, we reach four conclusions. First, the marriage premium persists even at the distributional level in the PSID and CPS data in the 1990s. Thus, the premium is more robust than previously believed and there is no need to focus on subjective indices, such as averages, that fail to give rise to uniform welfare orderings. Second, there is some evidence that the return to marriage may not be uniform across the wage distribution; in particular, the gains from marriage may be larger for those in the bottom third of the wage distribution. This implies that uniform welfare rankings may only be obtained if one broadens the concept of the marriage premium to incorporate wage dispersion. Third, the majority of the premium is explained by selection, but there is a small role for causal-based explanations. Finally, conclusions regarding the uniformity of the return to marriage across the wage distribution, but not the portion of the premium attributable to causal-based explanations, are affected by removal of time invariant and time-varying unobservables correlated with marital status and labor market performance, as well as the method used to control for potentially confounding observable covariates. In particular, our findings suggest that typical fixed effects methods alone – which are prominent in the marriage premium literature – are insufficient to give a complete picture of the returns to marriage.

There are two potential limitations to these findings. First, the additional results obtained from the application of Abadie's (2002) IV methodology – the greater role of selection and the uniformity across the distribution of the remaining causal portion of the marriage premium – hinge on the validity of the instrument utilized. If parents' marital status is correlated with time-varying unobservables that affect wages, some of the preceding conclusions may be suspect. Validation of these findings remains the goal of future work as new instruments are uncovered. Second, the nature of the CPS data led us to focus solely on the return to the first year of marriage, while the nature of the IV estimation (and sample size) led us to pool all married men together (regardless of length of marriage). There is some prior research that suggests that the return to marriage increases with the length of marriage (typically viewed as evidence in favor of the specialization hypothesis); see, e.g., Korenman and Neumark (1991) and Stratton (2002). However, Cornwell and Rupert (1997) conclude that the marriage premium represents a one-time intercept shift, as the authors fail to find a significant effect of marital duration on wages. Nonetheless, application of the

SD tests to comparisons of distributions differentiated by years of marriage may also prove fruitful in the future, particularly if an exogenous source of variation in marital duration is found (perhaps the gender of children as there is some evidence that divorce is less frequent when a couple has a son). In any event, given recent developments of stochastic dominance techniques and their potential for identifying *majority* preferences, future work should continue at the distributional level given the additional insights offered and the broader concept of the marriage premium that is defined – one inclusive of dispersion considerations.

References

[1] Abadie, A. (2002), "Bootstrap Tests for Distributional Treatment Effects in Instrumental Variable Models," *Journal of the American Statistical Association*, 97, 284-292.

[2] Amin, S., A.S. Rai, and G. Topa (2003), "Does Microcredit Reach the Poor and Vulnerable? Evidence From Northern Bangladesh," *Journal of Development Economics*, 70, 59-82.

[3] Anderson, G. (1996), "Nonparametric Tests of Stochastic Dominance in Income Distributions," *Econometrica*, 64, 1183-1193.

[4] Angrist, J.D., G.W. Imbens, and D.B. Rubin (1996), "Identi cation of Causal Effects Using Instrumental Variables," *Journal of the American Statistical Association*, 91, 444-472.

[5] Antonovics, K. and R. Town (2004), "Are all the Good Men Married? Uncovering the Source of the Marital Wage Premium," *American Economic Review*, 94, 317-321.

[6] Avery, R., F. Goldscheider, and A. Speare Jr. (1992), "Feathered Nest/Gilded Cage: Parental Income and Leaving Home in the Transition to Adulthood," *Demography*, 29, 375-388.

[7] Barrett, G. and S. Donald (2003), "Consistent Tests for Stochastic Dominance," *Econometrica*, 71, 71-104.

[8] Becker, G.S. (1973), "A Theory of the Marriage, Part I," *Journal of Political Economy*, 81, 813-846.

[9] Becker, G.S. (1974), "A Theory of the Marriage, Part II," *Journal of Political Economy*, 81, S11-S26.

[10] Becker, G.S. (1974), *The Economic Approach to Human Behavior*, Chicago: University of Chicago Press.

[11] Becker, G.S. (1991), *A Treatise on the Family*, enlarged edition Cambridge, MA: Harvard University Press.

[12] Bishop, J.A., J.P. Formby, and L.A. Zeager (2000), "The Effect of Food Stamp Cashout on Undernutrition," *Economics Letters*, 67, 75-85.

[13] Blackburn, M. and S. Korenman (1994), "The Declining Marital Status Earnings Differential," *Journal of Population Economics*, 7, 249-270.

[14] Blundell, R., L. Dearden, and B. Sianesi, (2002), "Evaluating the Impact of Education on Earnings in the UK: Models, Methods and Results from the NCDS," *Journal of the Royal Statistical Society*, forthcoming.

[15] Bratberg, E., A. Grasdal, and A.E. Risa, (2002), "Evaluating Social Policy by Experimental and Nonexperimental Methods," *Scandanavian Journal of Economics*, 104, 147-171.

[16] Chun, H. and I. Lee (2001), "Why Do Married Men Earn More: Productivity of Marriage Selection?" *Economic Inquiry*, 39, 307-319.

[17] Cohen, P.N. (2002), "Cohabitation and the Declining Marriage Premium for Men," *Work and Occupations*, 29, 34-363.

[18] Cornwell, C. and P. Rupert (1997), "Unobservable Individual Effects, Marriage and the Earnings of Young Men," *Economic Inquiry*, 35, 285-294.

[19] Daniel, K. (1995), "The Marriage Premium," in M. Tommasi and K. Ierulli (eds.) *The New Economics of Human Behavior*, Cambridge, England: Cambridge University Press.

[20] Duncan, G. and B. Holmund (1983), "Was Adam Smith Right After All? Another Test of the Theory of Compensating Wage Differentials," *Journal of Labor Economics*, 1, 366-374.

[21] Fisher, G., D. Wilson, and K. Xu (1998), "An Empirical Analysis of Term Premiums Using Significance Tests for Stochastic Dominance," *Economics Letters*, 60, 195-203.

[22] Ginther, D.K. and M. Zavodny (2001), "Is the Male Marriage Premium Due to Selection? The Effect of Shotgun Weddings on the Return to Marriage," *Journal of Population Economics*, 14, 313-328.

[23] Goldscheider, F.K. and L.J. Waite (1986), "Sex Differences in the Entry into Marriage," *American Journal of Sociology*, 92, 91-109.

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

[25] Gruber, J. (2000), "Is Making Divorce Easier Bad for Children? The Long Run Implications of Unilateral Divorce," NBER Working Paper 7968.

[26] Heckman, J.J., H. Ichimura, and P.E. Todd (1997), "Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Program," *Review of Economic Studies*, 64, 605-654.

[27] Heckman, J. and R. Robb (1985), "Alternative Methods for Evaluating the Impact of Interventions" in J. Heckman and B. Singer (eds.) *Longitudinal Analysis of Labor Market Data*, Econometric Society Monograph No. 10, Cambridge, U.K.: Cambridge University Press, 63-113.

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

[29] Hersch, J. and L.S. Stratton (2000), "Household Specialization and the Male Marriage Wage Premium," *Industrial and Labor Relations Review*, 54, 78-94.

[30] Horowitz, J.L. (2001), "Response to Comments of Esfandiar Maasoumi," *Journal of Econometrics*, 100, 97-98.

[31] Imbens, G.W and J.D. Angrist (1994), "Identification and Estimation of Local Average Treatment Effects," *Econometrica*, 62, 467-476.

[32] Imbens, G.W and D.B. Rubin (1997), "Estimating Outcome Distributions for Compliers in Instrumental Variable Models," *Review of Economic Studies*, 64, 555-574.

[33] Kaur, A., B.L.S. Prakasa Rao, and H. Singh (1994), "Testing for Second-Order Stochastic Dominance of Two Distributions," *Econometric Theory*, 10, 849-866.

[34] Klecan, L., R. McFadden, and D. McFadden (1991), "A Robust Test for Stochastic Dominance," unpublished manuscript, Department of Economics, MIT.

[35] Kobrin, F.A. and L.J. Waite (1984), " ," *Journal of Marriage and the Family*, 46, 807-816.

[36] Korenman, S. and D. Neumark (1991), "Does Marriage Really Make Men More Productive?" *Journal of Human Resources*, 26, 282-307.

[37] Li, J.H. and R.A. Wojtkiewicz (1994), "Childhood Family Structure and Entry into First Marriage," *Sociological Quarterly*, 35, 247-268.

[38] Light, A. (2004), "Gender Differences in the Marriage and Cohabitation Income Premium," *Demography*, 41, 263-284.

[39] Linton, O., E. Maasoumi, and Y.J. Whang (2003), "Consistent Testing for Stochastic Dominance: A Subsampling Approach," *Review Of Economic Studies*, forthcoming.

[40] Loh, E.S. (1996), "Productivity Differences and the Marriage Wage Premium for White Males," *Journal of Human Resources*, 31, 566-589.

[41] Maasoumi, E. (1986), "The Measurement and Decomposition of Multi-Dimensional Inequality," *Econometrica*, 54, 991-998.

[42] Maasoumi, E. (1998), "Empirical Analyses of Inequality and Welfare," Chapter 5 in P. Schmidt and M.H. Pesaran (eds.) *Handbook of Applied Econometrics*, Vol. II: Microeconomics, Basil Blackwell.

[43] Maasoumi, E. (2001), "On the Relevance of First Order Asymptotic Theory to Economics," *Journal of Econometrics*, 100, 83-86.

[44] Maasoumi, E. and A. Heshmati (2000), "Stochastic Dominance Amongst Swedish Income Distributions," *Econometric Reviews*, 19, 287-320.

[45] Maasoumi, E. and D.L. Millimet (2003), "Robust Inference Concerning Recent Trends in U.S. Environmental Quality," *Journal of Applied Econometrics*, forthcoming.

[46] Manski, C.F., G.D. Sandefur, S. McLanahan, and D. Powers (1992), "Alternative Estimates of the Effect of Family Structure During Adolescence on High School Graduation," *Journal of the American Statistical Association*, 87, 25-37.

[47] McFadden, D. (1989), "Testing for Stochastic Dominance," in Part II of T. Fomby and T.K. Seo (eds.) *Studies in the Economics of Uncertainty* (in honor of J. Hadar), Springer-Verlag.

[48] Millimet, D.L., M. Nieswiadomy, H. Ryu, and D.J. Slottje (2003), "Estimating Worklife Expectancies: An Econometric Approach," *Journal of Econometrics*, 113, 83-113.

[49] Nakosteen, R.A. and M.A. Zimmer (1987), "Marital Status and Earnings of Young Men: A Model With Endogenous Selection," *Journal of Human Resources*, 22, 248-268.

[50] Nakosteen, R.A. and M.A. Zimmer (1997), "Men, Money, and Marriage: Are High Earners More Prone than Low Earners to Marry?" *Social Science Quarterly*, 78, 66-82.

[51] Neumark, D. and D. Kawaguchi (2001), "Attrition Bias in Economic Relationships Estimated with Matched CPS Files," NBER Working Paper 8663.

[52] South, S.J. (2001), "The Variable Effects of Family Background on the Timing of First Marriage: United States, 1969–1993," *Social Science Research*, 30, 606-626.

[53] Stratton, L.S. (2002), "Examining the Wage Differential for Married and Cohabitating Men," *Economic Inquiry*, 40, 199-212.

[54] Wolfinger, N.H. (2003), "Parental Divorce and Offspring Marriage: Early or Late?," *Social Forces*, 82, 337-353.

A Appendix: Bootstrap Distribution Examples

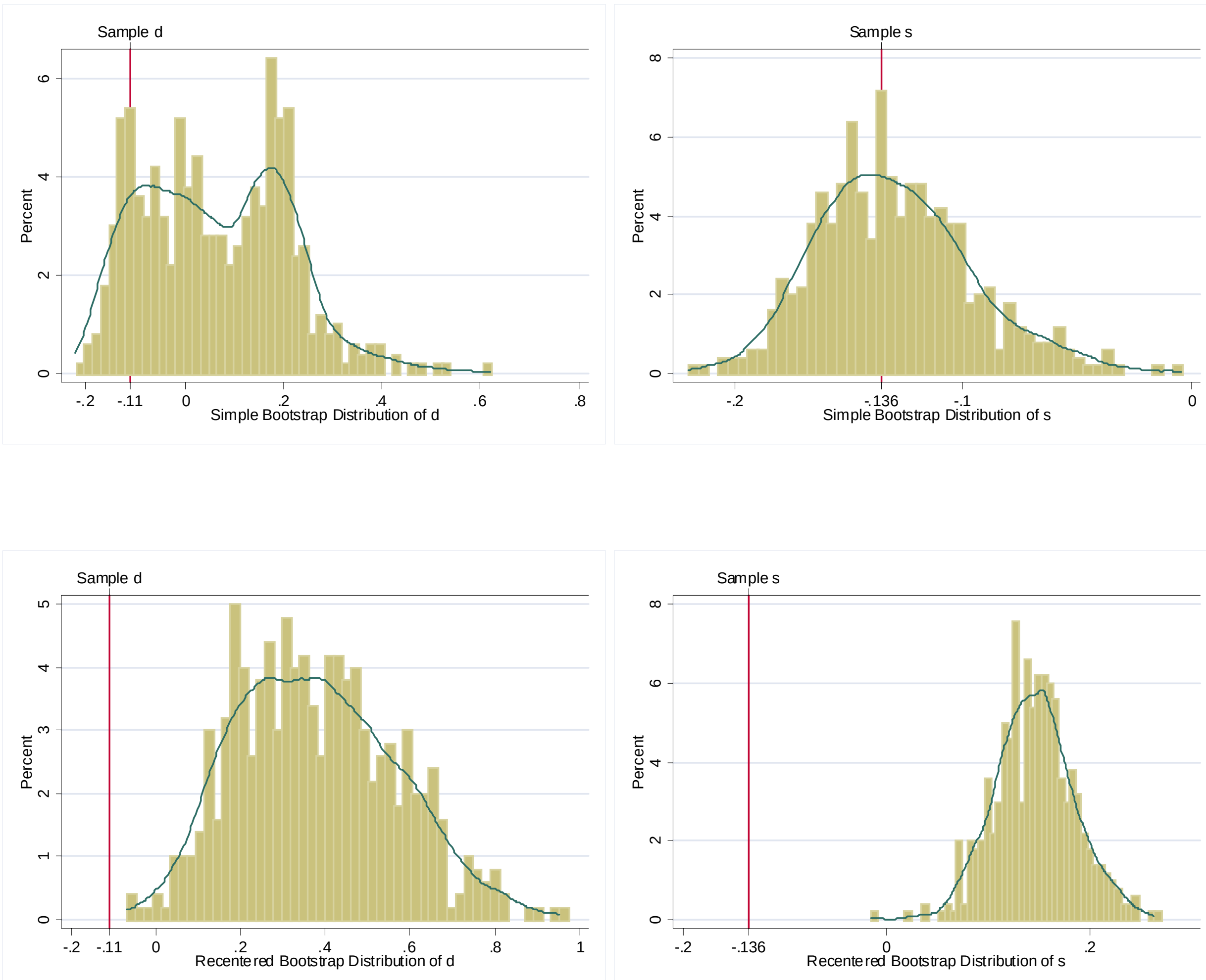
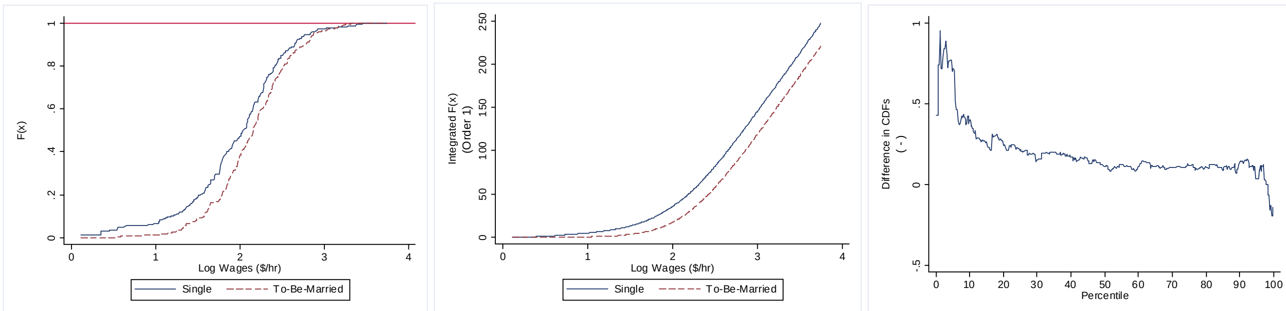


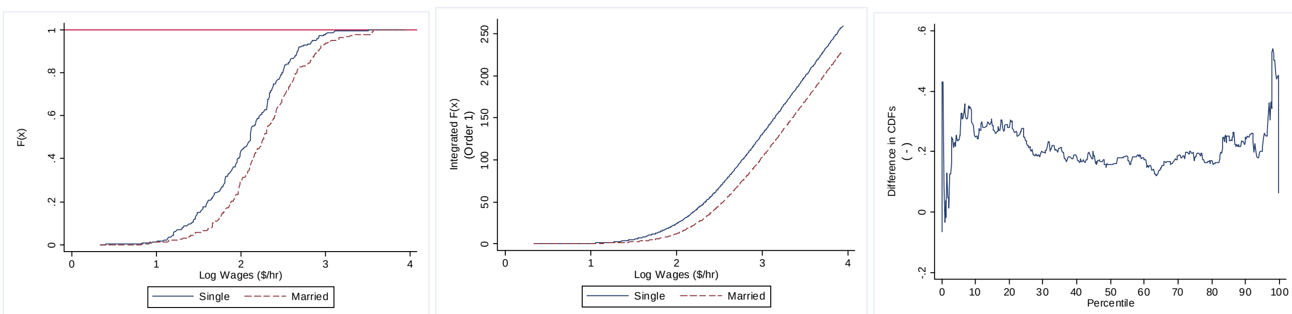
Figure A1. Bootstrap Distributions: Unconditional Wages in Period Two (CPS)

Notes: Kernel density overlaid; Epanechnikov kernel and Silverman's rule of thumb bandwidth utilized. See Table 3, Panel B.

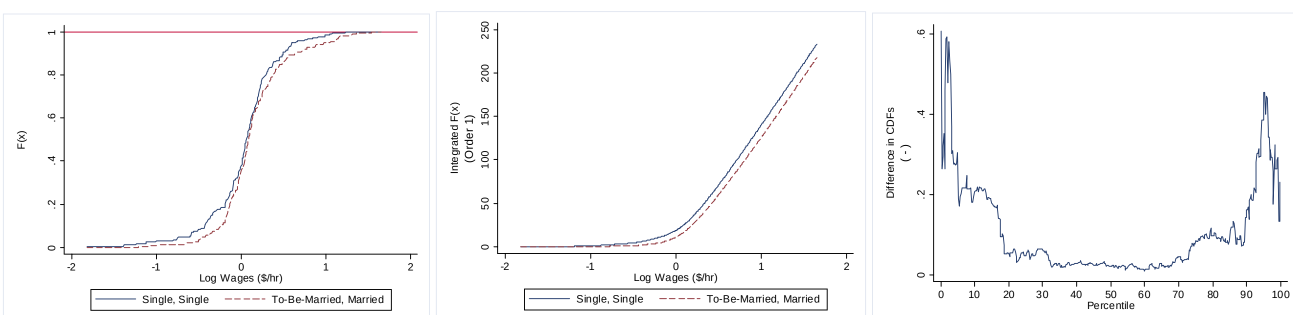
B Appendix: Sensitivity Results



Period 1



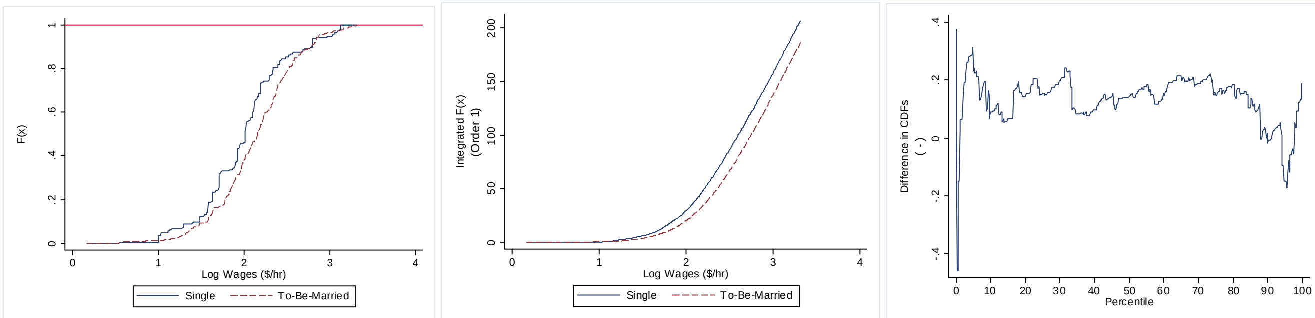
Period 2



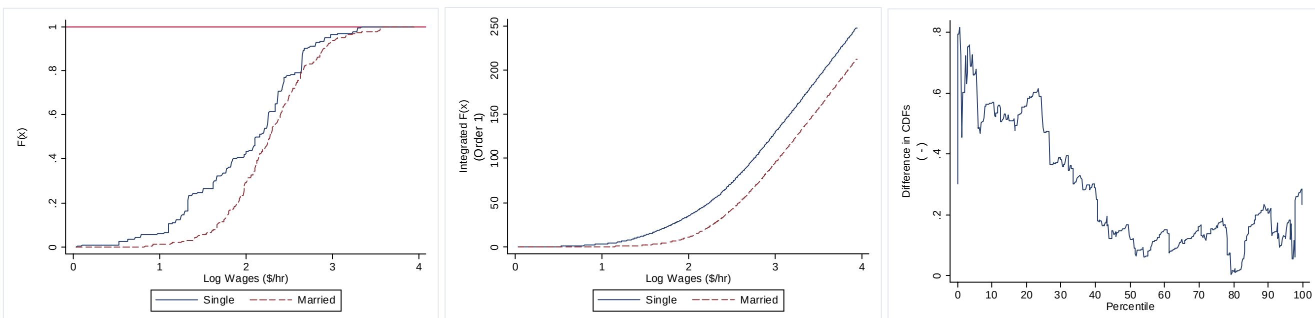
First Differences

Figure B1. CDFs and Integrated CDFs: Matching using Propensity Score (CPS Sample)

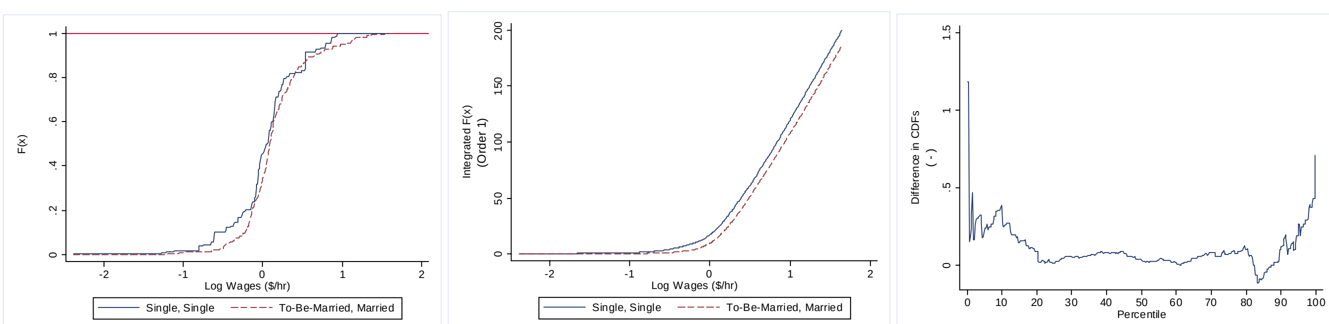
Note: The difference in CDFs are calculated as 'to-be-married' minus 'single' for Period 1, 'married' minus 'single' for Period 2, 'to-be-married, married' minus 'single, single' for First Differences.



Period 1



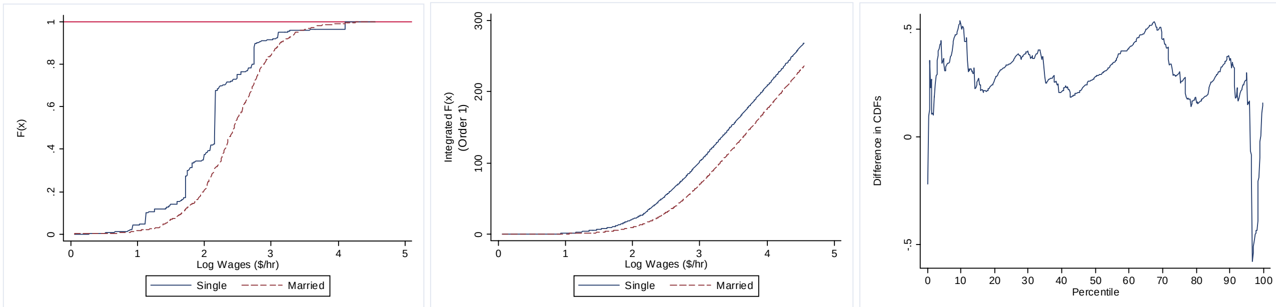
Period 2



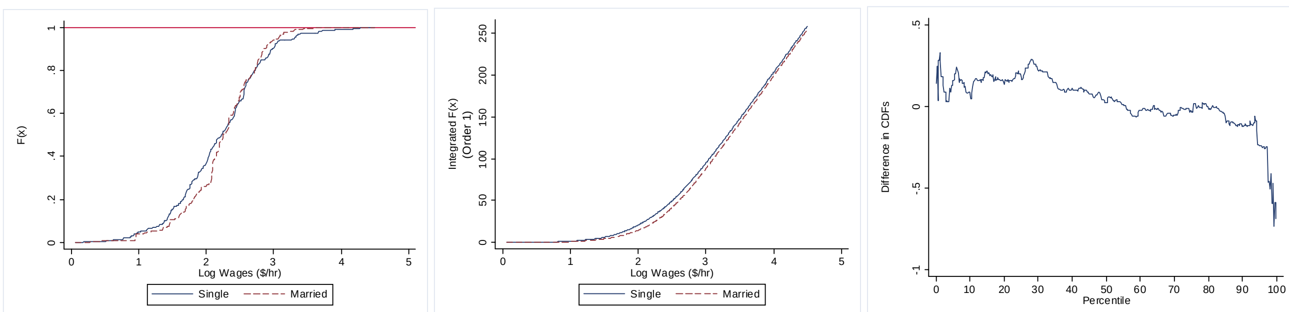
First Differences

Figure B2. CDFs and Integrated CDFs: Matching using Minimum Entropy Distance (CPS Sample)

Note: S_i matching based on $\beta = -0.5$. The difference in CDFs are calculated as 'to-be-married' minus 'single' for Period 1, 'married' minus 'single' for Period 2, 'to-be-married, married' minus 'single, single' for First Differences.



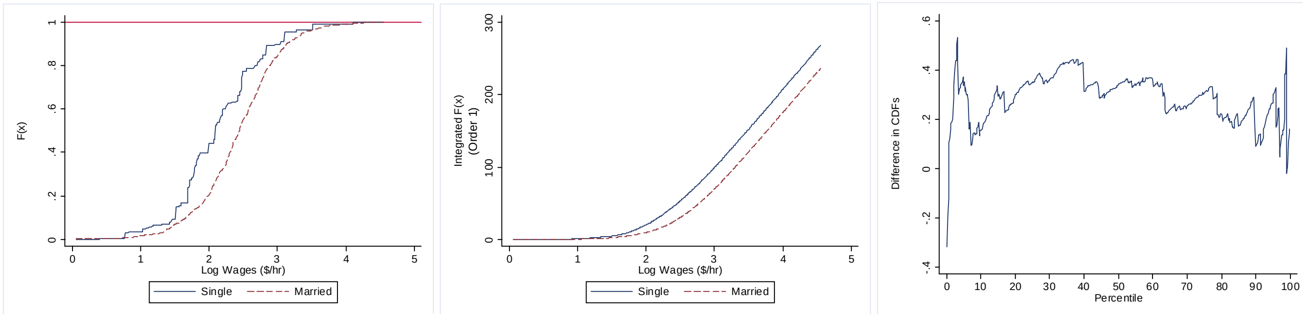
Without Instrument



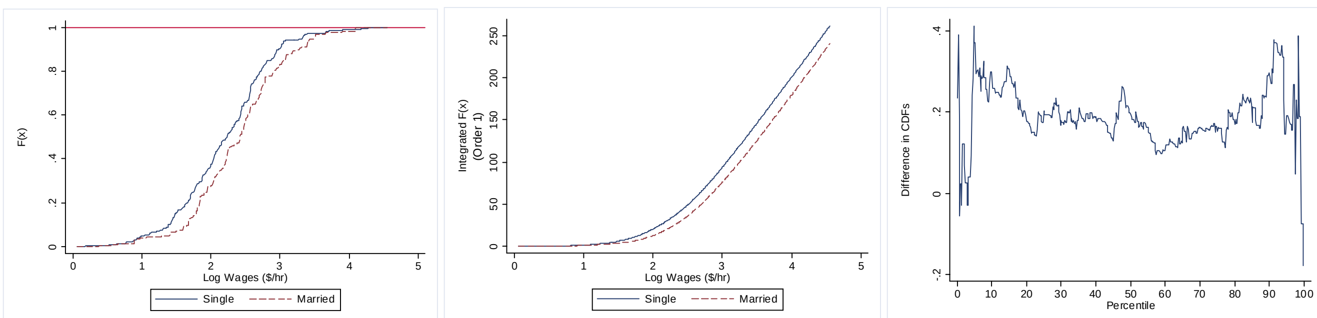
With Instrument

Figure B3. CDFs and Integrated CDFs: Matching using Propensity Score (PSID Sample)

Note: The difference in CDFs are calculated as 'married' minus 'single'.



Without Instrument



With Instrument

Figure B4. CDFs and Integrated CDFs: Matching using Minimum Entropy Distance (PSID Sample)

Note: S_i matching based on $\beta = -0.5$. The difference in CDFs are calculated as 'married' minus 'single'.

Table 1. Summary Statistics: CPS.

	Actual Sample						CPS 1996 Comparison				Test of Equality	
	Single		To-Be-Married		Married		Single		Married		Single	Married
Variable	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	p-value	
Hourly Wage	9.29	6.56	9.48	4.58	11.11	6.05	7.99	5.99	11.77	9.56	0.00	0.24
Labor Supply (hrs/wk)	42.27	8.45	44.25	8.37	42.91	7.00	41.66	9.19	44.31	9.15	0.00	0.01
Education												
< High School	0.07	0.26	0.06	0.23	0.06	0.25	0.12	0.32	0.11	0.32	0.00	0.01
High School	0.57	0.50	0.54	0.50	0.53	0.50	0.57	0.50	0.58	0.49	1.00	0.08
College +	0.35	0.48	0.40	0.49	0.40	0.49	0.31	0.46	0.31	0.46	0.00	0.00
Age	35.72	8.34	31.61	6.40	32.77	6.74	33.14	7.66	42.10	9.94	0.00	0.00
Class of Worker												
Private	0.85	0.36	0.85	0.36	0.83	0.37	0.87	0.33	0.84	0.37	0.00	0.65
Federal	0.03	0.18	0.02	0.14	0.04	0.19	0.02	0.15	0.04	0.20	0.00	1.00
State Govt.	0.04	0.20	0.06	0.23	0.05	0.23	0.04	0.20	0.05	0.21	1.00	1.00
Local Govt.	0.07	0.26	0.08	0.26	0.08	0.27	0.06	0.24	0.08	0.26	0.04	1.00
Union Member	0.04	0.21	0.07	0.26	0.07	0.26	0.03	0.18	0.05	0.21	0.01	0.12
Full Time Status	0.91	0.29	0.96	0.21	0.96	0.19	0.91	0.29	0.97	0.18	1.00	0.34
(> 35 hrs/wk)												
Metropolitan Status												
Urban	0.86	0.35	0.82	0.38	0.82	0.38	0.88	0.32	0.81	0.39	0.00	0.66
Rural	0.14	0.35	0.17	0.38	0.18	0.38	0.12	0.32	0.19	0.39	0.00	0.66
Race												
White	0.79	0.41	0.82	0.39	0.82	0.39	0.78	0.41	0.88	0.32	0.20	0.00
Black	0.17	0.37	0.11	0.32	0.11	0.31	0.16	0.37	0.08	0.27	0.15	0.06
American Indian	0.01	0.09	0.01	0.08	0.01	0.11	0.01	0.08	0.01	0.07	1.00	1.00
Asian or Pacific Isles	0.04	0.19	0.06	0.23	0.06	0.23	0.05	0.22	0.04	0.19	0.01	0.07
Region												
New England	0.07	0.25	0.06	0.24	0.05	0.23	0.06	0.24	0.05	0.22	0.03	1.00
Middle Atlantic	0.17	0.38	0.16	0.37	0.17	0.37	0.17	0.38	0.14	0.35	1.00	0.14
East North Central	0.18	0.38	0.15	0.36	0.15	0.36	0.16	0.37	0.18	0.38	0.01	0.18
West North Central	0.08	0.26	0.10	0.30	0.09	0.28	0.07	0.25	0.07	0.26	0.04	0.19
South Atlantic	0.17	0.38	0.20	0.40	0.19	0.40	0.16	0.36	0.18	0.38	0.16	0.65
East South Central	0.04	0.19	0.03	0.18	0.04	0.19	0.04	0.20	0.06	0.24	1.00	0.15
West South Central	0.08	0.26	0.08	0.28	0.08	0.27	0.09	0.28	0.11	0.31	0.05	0.10
Mountain	0.06	0.23	0.04	0.21	0.06	0.24	0.06	0.23	0.06	0.24	1.00	1.00
Pacific	0.17	0.37	0.18	0.38	0.17	0.38	0.20	0.40	0.15	0.36	0.00	0.34
Home Ownership												
Own	0.58	0.49	0.55	0.50	0.55	0.50	0.50	0.50	0.79	0.41	0.00	0.00
Rent	0.40	0.49	0.44	0.50	0.45	0.50	0.50	0.50	0.20	0.40	0.00	0.00
Number of Own Children												
Under 6	0.04	0.23	0.09	0.31	0.24	0.54	0.05	0.28	0.38	0.68	0.03	0.00
Number of Observations	10594		263		297		3782		15692			

Notes: 13 occupational categories are also utilized, but are not displayed, due to space considerations. Sample restricted to those with an hourly wage between $1 - 100. CPS Comparison Sample is obtained from the 1996 incoming rotation group. Wages are in 1982 dollars. Appropriate sample weights utilized.

Table 2. OLS and Fixed Effects Estimates Using CPS and PSID Samples.

	CPS						Korenman-Neumark				PSID	
	OLS				FE		OLS		FE		OLS	
	Pooled Sample		Period Two Only		Pooled Sample							
Variable	**Coeff**	**Std Error**	**Coeff**	**Std Error**	**Coeff**	**Std Error**	**Coeff**	**Std Error**	**Coeff**	**Std Error**	**Coeff**	**Std Error**
Married (1 = Yes)	0.217***	0.091	0.151***	0.034	0.068*	0.038	0.11***	0.02	0.06**	0.03	0.209***	0.048
Other Covariates												
Age, Age Squared	Yes						No				Yes	
Exper., Exper. Squared	No						Yes				No	
Region	9 region dummies						Dummy for South				9 region dummies	
Urban	Urban/Rural dummies						Urban/Rural dummies				Urban/Rural dummies	
Union	Union/Non-Union member dummies						Union/Non-Union member dummies				Union/Non-Union member dummies	
Occupation	13 three-digit occupational categories						8 single-digit occupational categories				11 Three-digit occupational categories	
Industry	20 industry dummies						11 industry dummies				13 industry dummies	
Year	10 Year dummies						3 Year dummies				N/A	
Non-Spouse Dependents	Number of own children under 6						Dummy for Non-Spouse Dependents				Number of children under 18	
Schooling	3 dummies for <HS, HS, College+						5 dummies				3 dummies for <HS, HS, College+	
Race	5 dummies						No				7 dummies	
Full-time Status	FT/PT status dummies						No				FT/PT status dummies	
Own/Rent Home	2 dummies						No				No	
Class of Worker	4 dummies for State, Local, Federal Govt. & Private						No				4 dummies for State, Local, Federal Govt. & Private	

Notes: ***/**/* denotes significance at the 1%, 5%, and 10% level, respectively. Korenman-Neumark results are taken from Korenman and Neumark (1991). N/A = not applicable.

Table 3. Unconditional Stochastic Dominance Tests: CPS.

	A. Wages in Period One		*B. Wages in Period Two*		*C. First-Difference in Wages*	
	Simple Bootstrap	**Recentered Bootstrap**	**Simple Bootstrap**	**Recentered Bootstrap**	**Simple Bootstrap**	**Recentered Bootstrap**
	X Single / ***Y*** To-Be Married	***X*** Single / ***Y*** To-Be Married	***X*** Single / ***Y*** Married	***X*** Single / ***Y*** Married	***X*** Single / ***Y*** Married	***X*** Single / ***Y*** Married
Observed Ranking	***Y SSD X***		***Y FSD X***		***Y SSD X***	
$d_{1,MAX}$	1.969		2.940		1.797	
$d_{2,MAX}$	0.273		-0.110		0.182	
d	0.273		-0.110		0.182	
$Pr\{d_1^* \leq 0\}$	0.000	0.002	0.000	0.004	0.000	0.016
$Pr\{d_2^* \leq 0\}$	0.006	0.002	0.392	0.006	0.004	0.002
$Pr\{d^* \leq 0\}$	0.006	0.004	0.392	0.010	0.004	0.018
$Pr\{d_1^* \geq d_1\}$	0.676	0.004	0.714	0.000	0.724	0.028
$Pr\{d_2^* \geq d_2\}$	0.626	0.832	0.856	1.000	0.842	0.910
$Pr\{d^* \geq d\}$	0.626	0.686	0.856	1.000	0.842	0.792
$s_{1,MAX}$	329.996		591.733		269.923	
$s_{2,MAX}$	-0.187		-0.136		-0.181	
s	-0.187		-0.136		-0.181	
$Pr\{s_1^* \leq 0\}$	0.000	0.000	0.000	0.000	0.000	0.000
$Pr\{s_2^* \leq 0\}$	0.856	0.438	1.000	0.002	1.000	0.000
$Pr\{s^* \leq 0\}$	0.856	0.438	1.000	0.002	1.000	0.000
$Pr\{s_1^* \geq s_1\}$	0.558	0.540	0.522	0.514	0.530	0.540
$Pr\{s_2^* \geq s_2\}$	0.798	1.000	0.538	1.000	0.638	1.000
$Pr\{s^* \geq s\}$	0.798	1.000	0.538	1.000	0.638	1.000

Notes: Bootstrap results based on 500 repetitions. Appropriate sample weights utilized. See text for further details.

Table 4. Partial Residual Stochastic Dominance Tests: CPS.

	A. Wages in Period One				*B. Wages in Period Two*				*C. First-Difference in Wages*			
	Simple Bootstrap		**Recentered Bootstrap**		**Simple Bootstrap**		**Recentered Bootstrap**		**Simple Bootstrap**		**Recentered Bootstrap**	
	X	***Y***	***X***	***Y***	***X***	***Y***	***X***	***Y***	***X***	***Y***	***X***	***Y***
	Single	**To-Be Married**	**Single**	**To-Be Married**	**Single**	**Married**	**Single**	**Married**	**Single**	**Married**	**Single**	**Married**
Observed Ranking	***X FSD Y***				***Y FSD X***				***Y FSD X***			
$d_{1,MAX}$	-0.152				8.427				14.383			
$d_{2,MAX}$	12.745				-0.128				-0.168			
d	-0.152				-0.128				-0.168			
$Pr\{d_1^* \leq 0\}$	0.720		0.180		0.104		0.322		0.004		0.362	
$Pr\{d_2^* \leq 0\}$	0.024		0.230		0.508		0.156		0.768		0.000	
$Pr\{d^* \leq 0\}$	0.744		0.410		0.612		0.478		0.772		0.362	
$Pr\{d_1^* \geq d_1\}$	0.636		0.946		0.500		0.264		0.316		0.000	
$Pr\{d_2^* \geq d_2\}$	0.516		0.116		0.542		0.968		0.666		1.000	
$Pr\{d^* \geq d\}$	0.616		0.928		0.470		0.810		0.664		1.000	
$s_{1,MAX}$	-0.948				1658.642				3479.133			
$s_{2,MAX}$	2937.248				-0.159				-0.168			
s	-0.948				-0.159				-0.168			
$Pr\{s_1^* \leq 0\}$	0.730		0.442		0.112		0.130		0.004		0.002	
$Pr\{s_2^* \leq 0\}$	0.072		0.090		0.742		0.054		0.946		0.004	
$Pr\{s^* \leq 0\}$	0.802		0.532		0.854		0.184		0.950		0.006	
$Pr\{s_1^* \geq s_1\}$	0.534		0.638		0.474		0.468		0.302		0.016	
$Pr\{s_2^* \geq s_2\}$	0.496		0.472		0.666		0.990		0.540		1.000	
$Pr\{s^* \geq s\}$	0.534		0.548		0.588		0.862		0.538		0.998	

Notes: First-stage regressions include controls for: age, age squared, number of own children under 6, race, education, occupation, class of worker, full-time status, union membership, housing type, urban, and region. See Table 3 and text for further details.

Table 5. Full Residual Stochastic Dominance Tests: CPS.

	A. Wages in Period One		*B. Wages in Period Two*		*C. First-Difference in Wages*	
	Simple Bootstrap	**Recentered Bootstrap**	**Simple Bootstrap**	**Recentered Bootstrap**	**Simple Bootstrap**	**Recentered Bootstrap**
	X Single, ***Y*** To-Be Married	***X*** Single, ***Y*** To-Be Married	***X*** Single, ***Y*** Married	***X*** Single, ***Y*** Married	***X*** Single, ***Y*** Married	***X*** Single, ***Y*** Married
Observed Ranking	***Y SSD X***		***None***		***Y SSD X***	
$d_{1,MAX}$	2.274		3.257		1.372	
$d_{2,MAX}$	0.163		0.068		0.165	
d	0.163		0.068		0.165	
$Pr\{d_1^* \leq 0\}$	0.000	0.000	0.000	0.000	0.000	0.052
$Pr\{d_2^* \leq 0\}$	0.010	0.000	0.114	0.000	0.004	0.000
$Pr\{d^* \leq 0\}$	0.010	0.000	0.114	0.000	0.004	0.052
$Pr\{d_1^* \geq d_1\}$	0.544	0.008	0.620	0.008	0.632	0.370
$Pr\{d_2^* \geq d_2\}$	0.726	0.998	0.724	1.000	0.836	1.000
$Pr\{d^* \geq d\}$	0.726	0.994	0.724	1.000	0.836	0.828
$s_{1,MAX}$	374.495		583.895		196.161	
$s_{2,MAX}$	-0.114		0.020		-0.084	
s	-0.114		0.020		-0.084	
$Pr\{s_1^* \leq 0\}$	0.000	0.000	0.000	0.000	0.000	0.000
$Pr\{s_2^* \leq 0\}$	0.928	0.546	0.518	0.540	0.740	0.006
$Pr\{s^* \leq 0\}$	0.928	0.546	0.518	0.540	0.740	0.006
$Pr\{s_1^* \geq s_1\}$	0.534	0.500	0.480	0.502	0.604	0.554
$Pr\{s_2^* \geq s_2\}$	0.476	0.896	0.436	0.444	0.524	0.998
$Pr\{s^* \geq s\}$	0.476	0.896	0.436	0.444	0.524	0.998

Notes: See Tables 3 and 4, and text for further details.

Table 6. Summary Statistics: PSID.

	Single		Married	
Variable	**Mean**	**SD**	**Mean**	**SD**
Hourly Wage	9.64	7.54	13.61	9.38
Labor Supply (hrs/wk)	44.33	9.82	45.71	9.23
Parent's Marital Status	0.65	0.48	0.77	0.42
(1 = Remained Intact)				
Education				
< High School	0.14	0.35	0.12	0.32
High School	0.51	0.50	0.54	0.50
College +	0.35	0.48	0.35	0.48
Age	32.21	5.79	37.33	6.56
Class of Worker				
Private	0.67	0.47	0.71	0.46
Federal	0.05	0.22	0.05	0.21
State Govt.	0.05	0.22	0.06	0.23
Local Govt.	0.04	0.20	0.05	0.22
Union Member	0.17	0.38	0.17	0.37
Full Time Status	0.87	0.33	0.94	0.23
(> 35 hrs/wk)				
Current Labor Market Status	0.91	0.29	0.97	0.16
(1 = Currently Employed)				
Race				
White	0.76	0.43	0.93	0.25
Black	0.19	0.39	0.05	0.22
American Indian	0.00	0.02	0.01	0.07
Asian or Pacific Isles	0.01	0.12	0.00	0.01
Latino	0.01	0.08	0.00	0.06
Region				
New England	0.06	0.23	0.07	0.26
Middle Atlantic	0.17	0.38	0.14	0.35
East North Central	0.17	0.38	0.17	0.38
West North Central	0.09	0.29	0.12	0.32
South Atlantic	0.17	0.38	0.16	0.37
East South Central	0.05	0.22	0.08	0.28
West South Central	0.05	0.22	0.08	0.27
Mountain	0.07	0.26	0.05	0.23
Pacific	0.17	0.38	0.13	0.34
Own Children under 18	0.23	0.71	1.48	1.17
Number of Observations	287		770	

Notes: Data from 1994 wave. Appropriate sample weights utilized.

Table 7. Unconditional Stochastic Dominance Tests: PSID.

	A. Without Instrument				*B. With Instrument*			
	Simple Bootstrap		**Recentered Bootstrap**		**Simple Bootstrap**		**Recentered Bootstrap**	
	X	***Y***	***X***	***Y***	***X***	***Y***	***X***	***Y***
	Single	**Married**	**Single**	**Married**	**Single**	**Married**	**Single**	**Married**
Observed Ranking	***Y FSD X***				***None***			
$d_{1,MAX}$	4.260				2.357			
$d_{2,MAX}$	-0.101				0.103			
d	-0.101				0.103			
$Pr\{d_1^*\leq 0\}$	0.000		0.004		0.000		0.006	
$Pr\{d_2^*\leq 0\}$	0.572		0.004		0.028		0.004	
$Pr\{d^*\leq 0\}$	0.572		0.008		0.028		0.010	
$Pr\{d_1^*\geq d_1\}$	0.646		0.000		0.650		0.000	
$Pr\{d_2^*\geq d_2\}$	0.752		0.998		0.766		0.978	
$Pr\{d^*\geq d\}$	0.752		0.998		0.766		0.950	
$s_{1,MAX}$	797.441				398.374			
$s_{2,MAX}$	-0.189				0.996			
s	-0.189				0.996			
$Pr\{s_1^*\leq 0\}$	0.002		0.214		0.000		0.176	
$Pr\{s_2^*\leq 0\}$	0.928		0.206		0.276		0.148	
$Pr\{s^*\leq 0\}$	0.928		0.420		0.276		0.324	
$Pr\{s_1^*\geq s_1\}$	0.580		0.000		0.518		0.000	
$Pr\{s_2^*\geq s_2\}$	0.602		0.930		0.574		0.814	
$Pr\{s^*\geq s\}$	0.602		0.880		0.574		0.598	

Notes: Bootstrap results based on 500 repetitions. Appropriate sample weights utilized. See text for further details.

Table 8. Partial Residual Stochastic Dominance Tests: PSID.

	A. Without Instrument				B. With Instrument			
	Simple Bootstrap		Recentered Bootstrap		Simple Bootstrap		Recentered Bootstrap	
	X	*Y*	*X*	*Y*	*X*	*Y*	*X*	*Y*
	Single	Married	Single	Married	Single	Married	Single	Married
Observed Ranking	*None*				*None*			
$d_{1,MAX}$	2.703				0.764			
$d_{2,MAX}$	0.149				1.628			
d	0.149				0.764			
$Pr\{d_1^*\leq 0\}$	0.424		0.236		0.394		0.208	
$Pr\{d_2^*\leq 0\}$	0.432		0.288		0.434		0.322	
$Pr\{d^*\leq 0\}$	0.856		0.524		0.828		0.530	
$Pr\{d_1^*\geq d_1\}$	0.496		0.520		0.552		0.632	
$Pr\{d_2^*\geq d_2\}$	0.518		0.676		0.452		0.452	
$Pr\{d^*\geq d\}$	0.076		0.268		0.022		0.118	
$s_{1,MAX}$	437.531				39.929			
$s_{2,MAX}$	1.054				110.562			
s	1.054				39.929			
$Pr\{s_1^*\leq 0\}$	0.460		0.290		0.444		0.290	
$Pr\{s_2^*\leq 0\}$	0.470		0.446		0.456		0.480	
$Pr\{s^*\leq 0\}$	0.930		0.736		0.900		0.770	
$Pr\{s_1^*\geq s_1\}$	0.488		0.480		0.546		0.554	
$Pr\{s_2^*\geq s_2\}$	0.508		0.554		0.448		0.450	
$Pr\{s^*\geq s\}$	0.048		0.132		0.008		0.028	

Notes: First-stage regressions include controls for: age, age squared, number of own children under 18, race, education, occupation, class of worker, full-time status, current labor market status, union membership, and region. See Table 7 and text for further details.

Table 9. Full Residual Stochastic Dominance Tests: PSID.

	A. Without Instrument				*B. With Instrument*			
	Simple Bootstrap		**Recentered Bootstrap**		**Simple Bootstrap**		**Recentered Bootstrap**	
	X	***Y***	***X***	***Y***	***X***	***Y***	***X***	***Y***
	Single	**Married**	**Single**	**Married**	**Single**	**Married**	**Single**	**Married**
Observed Ranking	***None***				***None***			
$d_{1,MAX}$	4.128				1.143			
$d_{2,MAX}$	0.173				0.288			
d	0.173				0.288			
Pr{d_1*£0}	0.000		0.000		0.000		0.000	
Pr{d_2*£0}	0.062		0.006		0.000		0.000	
Pr{d*£0}	0.062		0.006		0.000		0.000	
Pr{$d_1^* \geq d_1$}	0.630		0.250		0.936		0.622	
Pr{$d_2^* \geq d_2$}	0.720		0.988		0.928		0.988	
Pr{$d^* \geq d$}	0.720		0.938		0.928		0.928	
$s_{1,MAX}$	775.586				186.406			
$s_{2,MAX}$	2.787				15.068			
s	2.787				15.068			
Pr{$s_1^* \leq 0$}	0.006		0.052		0.090		0.074	
Pr{$s_2^* \leq 0$}	0.082		0.134		0.000		0.128	
Pr{$s^* \leq 0$}	0.088		0.186		0.090		0.202	
Pr{$s_1^* \geq s_1$}	0.476		0.024		0.540		0.110	
Pr{$s_2^* \geq s_2$}	0.740		0.856		0.862		0.802	
Pr{$s^* \geq s$}	0.734		0.582		0.752		0.414	

Notes: See Tables 7 and 8, and text for further details.